

Simulation of Rapidly Varying and Dry
Bed Flow using the Serre equations
solved by a Finite Element Volume
Method.

Jordan Peter Anthony Pitt

April 2019

A thesis submitted for the degree of Doctor of Philosophy
of the Australian National University



To my mother and father who have provided me with everything.

Declaration

The work in this thesis is my own except where otherwise stated.

Jordan Peter Anthony Pitt

Acknowledgements

I would like to thank my lead supervisor Professor Stephen Roberts for his insight, suggestions and time spent improving my research. Additionally, I would also like to thank Dr Chris Zoppou who put a tremendous amount of time and effort into reading and editing my work. Furthermore, I want to thank the remainder of my supervisory panel Professor Markus Hegland and Professor John Urbas.

I would also like to thank fellow researchers who provided their experimental data:

- Dr David George, Cascades Volcano Observatory, U.S. Geological Survey for providing the digitised data for the negative rectangular wave experiment.
- Professor Sedar Beji, Department of Naval Architecture and Ocean Engineering, Istanbul Technical University for providing the data for the periodic waves over a submerged bar experiment.
- Dr Volker Roeber, Department of Physical Oceanography, University of Hawai‘i at Mānoa for providing the data for the solitary wave over a fringing reef experiment.

Finally, I extend my thanks to my dear friend Joseph Gibson who donated a computer that hosted many numerical experiments.

Abstract

Recent research in numerical wave modelling has focused on developing computational methods for solving non-linear dispersive wave equations as an extension to methods solving the non-linear shallow water wave equations. By including extra terms that allow for dispersion these equations more accurately model water waves than the shallow water wave equations. An interesting example of these non-linear dispersive equations for modelling water waves are the Serre equations.

In this work an efficient and robust numerical method for the one-dimensional Serre equations was developed. This method uses a finite element method to solve an elliptic equation and a finite volume method to solve the remaining conservation equations and is hence termed the Finite Element Volume Method (FEVM). The use of a finite element method and a finite volume method makes the FEVM adaptable to unstructured meshes and parallelisable. The FEVM recovers the lake at rest steady state and accurately simulates flows over dry beds.

The convergence and dispersion properties of the FEVM were determined using a linear analysis. The FEVM was validated against analytic and forced solutions of the Serre equations, demonstrating its convergence and conservation properties. Finally, the method was validated against experimental data for a wide array of physical scenarios, establishing its utility as a realistic model.

All these analyses and validations were also conducted on other methods allowing comparisons between them and the FEVM. The FEVM was found to be the most robust of these methods whilst being adequately accurate. Since the FEVM can be extended to unstructured meshes with parallelised code it is the most promising of the studied methods for solving the Serre equations in two dimensions.

Contents

Acknowledgements	vii
Abstract	ix
1 Introduction	1
1.1 Objectives of the Thesis	3
1.2 Original Contribution of the Thesis	3
1.2.1 Publications	4
1.3 Organisation of the Thesis	9
2 The Serre Equations	11
2.1 The One-Dimensional Serre Equations	12
2.1.1 Alternative Form	14
2.2 Properties	15
2.2.1 Conservation Properties	15
2.2.2 Dispersion Properties	16
2.2.3 Analytic Solutions	17
2.3 Forced Solutions	19
2.4 Behaviour in the Presence of Steep Gradients	21
3 Finite Element Volume Method	27
3.1 Notation for Numerical Grids	28
3.2 Structure Overview	29
3.2.1 Reconstruction	32
3.2.2 Fluid Velocity	34
3.2.3 Flux Across the Cell Interfaces	39
3.2.4 Source Terms	43
3.2.5 Update Cell Averages	44
3.2.6 Second-Order SSP Runge-Kutta Method	44

3.3	CFL condition	45
3.4	Boundary Conditions	45
3.5	Dry Beds	46
4	Linear Analysis	49
4.1	Linearised Serre Equations	51
4.2	Evolution Matrix	51
4.2.1	Reconstruction	52
4.2.2	Fluid Velocity	53
4.2.3	Flux Across the Cell Interfaces	55
4.2.4	Update Cell Averages	58
4.2.5	Second-Order SSP Runge-Kutta Method	59
4.3	Convergence Analysis	60
4.3.1	Stability	60
4.3.2	Consistency	61
4.4	Dispersion Analysis	64
5	Numerical Validation	73
5.1	Measuring Convergence and Conservation	74
5.1.1	Measure of Convergence	74
5.1.2	Measures of Conservation	74
5.2	Solitary Travelling Wave Solution	75
5.3	Lake at Rest Solution	81
5.4	Forced Solutions	85
5.4.1	Results for a Wet Bed	89
5.4.2	Results with a Dry Bed	89
6	Experimental Validation	97
6.1	Evolution of a Negative Rectangular Wave	97
6.1.1	Results for the $0.01m$ Negative Rectangular Wave	99
6.1.2	Results for the $0.03m$ Negative Rectangular Wave	102
6.2	Periodic Waves Over A Submerged Bar	107
6.2.1	Low Frequency Results	108
6.2.2	High Frequency Results	113
6.3	Solitary Wave Over a Fringing Reef	118
6.4	Run-up Experiment	125

7 Conclusion	131
7.1 Future Work	133
7.1.1 Including Wave-breaking	134
7.1.2 Implementation of Different Boundary Conditions	135
7.1.3 Incorporation of Discontinuous Bed Profiles	135
7.1.4 Incorporation of Bed Friction	135
7.1.5 Complete Convergence Analysis	136
7.1.6 Extension to two-dimensional Serre equations on unstructured meshes	136
A Additional Conservation Information	137
A.1 Solitary Travelling Wave Solution	137
A.2 Lake At Rest Solution	139
B Finite Element Method Definitions	141
B.1 Basis Function Definitions	141
B.2 Function Space Definitions	143
C Linear Analysis Results	145
C.1 Finite Difference Volume Methods	147
C.2 Finite Difference Methods	154
C.3 Consistency Results	156
D Publications	163
Bibliography	170

Chapter 1

Introduction

A significant portion of the world's people and critical infrastructure is located near the coast. While the ocean provides many opportunities it also susceptible to natural hazards, particularly extreme ocean waves such as tsunamis and storm surges. Furthermore, the dynamics of ocean waves drives other physical phenomena; such as the break-up of sea-ice and the erosion of beaches. Therefore, accurate modelling of ocean waves is important to society.

The physics of water can be described using Newton's second law and were initially presented by Euler in 1757 [1]. The Euler equations were then extended to include viscosity, producing the full Navier-Stokes equations [2, 3]. Numerical methods [4, 5, 6] have been developed to solve the Euler equations; however due to their complexity these methods cannot accurately resolve fluid behaviour over the scales required to model tsunamis along a coastline.

For this reason the central focus of water wave modelling has been simplified water wave theories that approximate the behaviour of the free surface of water governed by the Euler equations. The most popular class of these approximate water wave theories are the shallow water wave theories where the characteristic water depth h_0 is far smaller than the characteristic wavelength λ_0 , thus $\sigma = h_0/\lambda_0 \ll 1$. For tsunamis and storm surges h_0 is typically 4km far from the coastline and λ_0 can be 100km , therefore $\sigma \ll 1$.

Neglecting all terms of order $\mathcal{O}(\sigma^2)$ the full Euler equations reduce to the Shallow Water Wave Equations (SWWE) [7] which describe fully non-linear non-dispersive waves. Retaining higher powers of σ leads to a class of equations known as 'Boussinesq-type' equations. Boussinesq-type equations are then classified by the powers of σ they retain and their retained non-linearity. The non-linearity $\epsilon = a_0/h_0$ compares the characteristic amplitude of the waves a_0 to the water

depth h_0 .

Water wave models form a spectrum with the SWWE being the simplest and most restrictive model and the Boussinesq-type models retaining the highest powers of σ and allowing the largest ϵ values being the most complex and least restrictive. The Serre equations are one particular Boussinesq-type equation that retains all terms of order $\mathcal{O}(\sigma^4)$ and makes no assumption on the size of ϵ [7]. Since the Serre equations allow arbitrary wave height they are the most appropriate model for water waves for the $\mathcal{O}(\sigma^4)$ class of Boussinesq-type equations. Furthermore, by retaining σ^2 terms the Serre equations more accurately model water waves than the SWWE for intermediate water depths where σ^2 terms can be significant. Intermediate water depths tend to occur as tsunamis and storm surges approach the coastline and interact with the varying bathymetry. Therefore, the Serre equations are well suited to modelling tsunamis and storm surges.

To model the impact of tsunamis and storm surges on a coastline requires modelling a large area over a long period of time. For this reason, most large scale models of tsunamis and storm surges are performed on supercomputers. To efficiently use supercomputers requires computational methods that are efficient, parallelisable and robust. A method is robust if it can stably and accurately model all appropriate physical scenarios. Robustness ensures that multiple simulations are not required because a method encountered a physical situation that lead to instabilities. This thesis will focus on robustness of the methods in the presence of steep gradients in the water surface and for the inundation of beaches.

A variety of numerical methods have been developed to solve the SWWE from finite difference methods [9] to finite volume methods [8, 10] and discontinuous Galerkin methods [11]. When producing large scale, parallelised, efficient and robust computational methods the finite volume methods and discontinuous Galerkin methods have distinct advantages over their finite difference method counterparts; they can be readily extended to unstructured meshes and are robust. Unstructured meshes are attractive because they allow for efficient numerical solutions over complex geometries. Furthermore, finite volume methods conserve mass and momentum in their numerical solutions, reproducing the conservative physics at the foundation of the equations. Discontinuous Galerkin methods can also be made conservative under the right conditions, but due to their longer history and simplicity the most popular numerical methods for modelling tsunamis and storm surges are the finite volume methods [8, 10].

Although great progress has been made in the development of large scale, efficient and robust computational methods for the SWWE [8, 10] these models

suffer from the assumptions made to derive the SWWE. Crucially, the SWWE neglect terms of order $\mathcal{O}(\sigma^2)$ in the Euler equations and are therefore unable to model dispersion. Recent research has highlighted the significant impact of dispersion on the evolution of tsunamis [12, 13], driving the development of models based on Boussinesq-type equations. Among these Boussinesq-type equations the Serre equations are the best placed [7]; retaining high-order σ terms and allowing arbitrary wave amplitude. Hence, the overarching goal of this research is the development of large-scale, efficient and robust computational methods for the Serre equations for the purposes of modelling tsunamis and storm surges.

1.1 Objectives of the Thesis

In view of the overarching goal, the primary motivation of this thesis was the development of a numerical method for solving the one-dimensional Serre equations. This method should be robust to steep gradients in the free surface and the inundation of beaches and be extendable to the two-dimensional Serre equations using unstructured meshes. Since many of these aims were achieved for the SWWE using finite volume methods, the focus of this thesis is the extension of finite volume methods for the SWWE to the Serre equations.

The primary focus of this thesis is the development of the Finite Element Volume Method (FEVM), which satisfies all the initial goals. The FEVM is an improvement of the Finite Difference Volume Methods (FDVM) described by Zoppou [14]. The FEVM can adequately handle dry beds and uses a finite element method instead of a finite difference method, making it suitable for unstructured meshes.

The FEVM was assessed with a linear analysis, a validation against analytic and forced solutions and experimental results. At all stages of this assessment the method is compared to at least one other method to demonstrate its strengths and weaknesses. Overall, the FEVM compared well with the other methods and satisfied all the objectives of the thesis.

1.2 Original Contribution of the Thesis

The research in this thesis made the following original contributions to the field:

- Implementation of the third-order FDVM.

- Observation and justification of a new structure in the solution of the Serre equations in the presence of steep gradients in the free surface.
- Extension of the second-order FDVM to allow for dry beds.
- Development and description of the well-balanced second-order FEVM that is capable of modelling flows over dry beds.
- A linear analysis of convergence for all developed finite volume based methods as well as some finite difference methods.
- A complete linear analysis of the dispersion properties for all developed finite volume based methods as well as some finite difference methods.
- Validation of FEVM and the second-order FDVM using forced solutions where all terms of the Serre equations are present for both wet and dry beds.
- Comparison of the numerical solutions of FEVM and the second-order FDVM with experimental results in the presence of dry beds and with wave breaking.

1.2.1 Publications

The publications generated from this research are compiled here in chronological order. A brief summary, my contribution to the paper and the relevance of the paper to this thesis are also provided. The publication list is reproduced in Appendix D where the abstracts of the publications are also provided.

A Solution of the Conservation Law Form of the Serre Equations

Australia and New Zealand Industrial and Applied Mathematics Journal (2016)

C. Zoppou, S.G. Roberts and J. Pitt

Summary:

A second-order FDVM for the one-dimensional Serre equations with a horizontal bed is described and validated against an analytic solution.

My Contribution:

I reproduced the results of my coauthors with my own implementation of the method. This method was consistent with a SWWE solver allowing the computational cost of solving the Serre equations and the SWWE to be compared.

Relevance to Thesis:

The method described in this paper was extended to allow for varying bathymetry [15] and dry beds, producing the second-order FDVM. The results of a linear analysis and numerical experiments for the second-order FDVM are reported in this thesis. The linear analysis of the convergence and dispersion properties performed in this thesis assume a completely wet horizontal bed, and therefore the results in this thesis apply to the method described in this paper. The linear analysis results can be found in Chapter 4 and Appendix C. The extended version of the method was validated against analytic and forced solutions in Chapter 5 and compared to experimental results in Chapter 6.

Numerical Solution of the Fully Non-Linear Weakly Dispersive Serre Equations for Steep Gradient Flows

Applied Mathematical Modelling (2017)

C. Zoppou, J. Pitt and S.G. Roberts

Summary:

The first-, second-, and third-order FDVM for the one-dimensional Serre equations with variable bathymetry were described. The results of a linear analysis of the dispersion properties of the methods for waves on quiescent water are provided. These methods were then validated against an analytic solution. A well-balanced version of the second-order FDVM is described and further validated against experimental results.

My Contribution:

The methods, linear dispersion analysis and numerical solutions were primarily produced by me in collaboration with my coauthors who wrote the paper.

Relevance to Thesis:

The first-, second- and third-order FDVM described in this paper are the methods whose results are reported in this thesis. The results of a linear analysis of the dispersion and convergence properties of all these methods are provided in Chapter 4 and Appendix C. The linear analysis of dispersion presented in this paper was extended to allow for a mean background flow in the water on which the waves occur. The results of the analytic solution validation in this paper are reproduced in Chapter 5 and are extended by studying the convergence and conservation properties of more quantities. The second-order FDVM was then validated against forced solutions in Chapter 5 and experimental results in Chapter 6. The results of the second-order FDVM for the negative rectangular wave experiment [16] and the periodic waves over a submerged bar experiment [17] in this paper are reproduced in this thesis.

Importance of Dispersion for Shoaling Waves

22nd International Congress on Modelling and Simulation (2017)

J. Pitt, C. Zoppou and S.G. Roberts

Summary:

The numerical solutions of the second-order FDVM [15] for the Serre equations were compared to the numerical solutions of ANUGA [10] for the SWWE. By comparing dispersive and non-dispersive water wave models the influence of dispersion on shoaling waves was studied. Two scenarios were investigated, an artificial example of the shoaling of a solitary wave over a fringing reef and the experimental results of Beji and Battjes [17] studying periodic waves travelling over a submerged bar.

My Contribution:

This paper was primarily produced by me in collaboration with my coauthors, based on research that I primarily undertook.

Relevance to Thesis:

The experimental results of the second-order FDVM for the periodic waves over a submerged bar experiments of Beji and Battjes [17] are reproduced in Chapter 6.

Behaviour of the Serre Equations in the Presence of Steep Gradients Revisited

Wave Motion (2018)

J.P.A. Pitt, C. Zoppou and S.G. Roberts

Summary:

The first-, second- and third-order FDVM as well as two second-order finite difference methods for the Serre equations were used to numerically study the behaviour of steep gradients in the water surface. The convergence and conservation properties of many numerical solutions of all these methods were used to justify the observed behaviours. One such behaviour was not previously reported by the literature and was shown to best represent the solution of the Serre equations in the presence of steep gradients in the water surface over short time spans. The effect of the numerical method, the grid resolution and the smoothing of the initial conditions on the observed behaviours was studied.

My Contribution:

This paper was primarily produced by me in collaboration with my coauthors, based on research that I primarily undertook.

Relevance to Thesis:

The behaviours observed and justified in this paper are summarised in Chapter 2. The results of this paper demonstrate the utility of using a finite volume based method to solve the Serre equations in the presence of steep gradients, hence the further development of these methods in this thesis.

1.3 Organisation of the Thesis

Chapter 2 proceeds by presenting the one-dimensional Serre equations in conservation law form with a source term. The dispersion and conservation properties and known analytic solutions of the Serre equations are also presented. The forced Serre equations and the concept of forced solutions are introduced. Finally, the main results of the investigation into the behaviour of the Serre equations in the presence of steep gradients in the free surface [18] are summarised.

This is followed by Chapter 3 which provides a detailed description of the FEVM. In this thesis the results of other numerical methods are also provided, descriptions of these methods can be found in the literature [15, 18].

Chapter 4 provides a linear analysis of the convergence and dispersion properties of the FEVM in detail. The analysis begins with the linearised Serre equations over a horizontal bed and then derives the evolution matrix; through which the convergence and dispersion properties can be studied. The results of the linear analysis are also provided for all the methods used by Pitt et al. [18] as a comparison.

The convergence and conservation properties of the FEVM and all the numerical methods used by Pitt et al. [18] are then assessed in Chapter 5 using analytic and forced solutions of the Serre equations. This is followed by an experimental validation of the FEVM and the second-order FDVM in Chapter 6.

Chapter 7 summarises the major contributions and findings of this thesis and provides ideas for future work.

Appendix A provides expressions for the total amount of the conserved quantities over any domain for all of the analytic solutions described in this thesis.

The basis functions used by the FEVM and the precise definition of the function spaces mentioned in this thesis are provided in Appendix B.

Appendix C provides the evolution matrix produced by the linear analysis for the first-, second- and third-order FDVM and two second-order finite difference methods. It also includes the consistency results for these methods, as they were omitted from Chapter 4.

Finally the publications I contributed to throughout the research underpinning this thesis and their abstracts are provided in chronological order in Appendix D.

Chapter 2

The Serre Equations

In this chapter the Serre equations are introduced and their relevant properties are presented.

The Serre equations are a system of partial differential equations that describe the free-surface waves of fluids whose motion is dominated by gravitational forces. The Serre equations are an approximation to the Euler equations [1]; describing waves in shallow water where the characteristic depth of the water h_0 is much smaller than the characteristic wavelength of the waves λ_0 thus, the shallowness parameter $\sigma = h_0/\lambda_0 \ll 1$. Typically, water is considered to be shallow when $\sigma \leq 1/20$ [19]. Tsunamis and storm surges are shallow water waves as the typical ocean depth in which they occur is 4km and both can have wavelengths up to 100km long.

The Serre equations for one-dimensional flows over horizontal beds were first derived by Serre [20] using asymptotic expansion. They were later derived using depth integration by Su and Gardner [21] and are equivalent to the Green-Naghdi equations [22] derived using the theory of directed fluid sheets. The Serre equations were extended to spatially varying bathymetry by Seabra-Santos et al. [23].

The Serre equations are fully non-linear and thus applicable across the entire range of wave amplitudes a_0 , which are commonly characterised by the non-linearity parameter $\epsilon = a_0/h_0$. They describe fluids under gravity without assuming a hydrostatic pressure distribution and thus allow for dispersion. Furthermore, the dispersion relationship for linear waves of the Serre equations approximates the linear wave theory for the Euler equations well [24]. Therefore, the Serre equations are considered one of the best approximate water wave models up to wave-breaking [7, 25].

In this chapter the one-dimensional Serre equations and a reformulation of

these equations into conservation law form with a source term are presented. The relevant properties of the Serre equations and forced solutions are introduced. Finally, the contribution of the research underpinning this thesis [18] to the understanding of the behaviour of steep gradients in the free-surface for the Serre equations is summarised.

2.1 The One-Dimensional Serre Equations

This thesis focuses on the Serre equations as derived from the depth-integration approach [21, 23]. A derivation of these equations is not presented here given the extent of the literature available [14, 21, 23]. The one-dimensional Serre equations describe the behaviour of unsteady free surface fluid flow for an inviscid fluid with a constant density ρ , neglecting wave-breaking. Additionally, bottom friction effects are neglected in this thesis.

The primitive variables of the Serre equations are demonstrated in Figure 2.1, they are the height $h(x, t)$ of a column of fluid above the stationary bed profile given by $b(x)$ and the average horizontal velocity $u(x, t)$ of the column of fluid. The stage $w(x, t) = h(x, t) + b(x)$ gives the absolute location of the free surface. Moreover, to relate the primitive variables of the Serre equations to the intuitive variables of the Euler equations the horizontal velocity $u'(x, z, t)$ and the vertical velocity $v'(x, z, t)$ of a fluid particle are introduced.

The derivation of the Serre equations is similar to that of the Shallow Water Wave Equations (SWWE) [26]. Both derivations assume that the horizontal velocity of a fluid particle $u'(x, z, t)$ inside a column of fluid equals the average horizontal velocity of the column of fluid $u(x, t)$. By integrating the conservation of mass equation for an inviscid fluid with a no-slip condition at the bed, the distribution of the vertical velocity of a fluid particle throughout a column of fluid is obtained [14]

$$v'(x, z, t) = u \frac{\partial b}{\partial x} - (z - b) \frac{\partial u}{\partial x}. \quad (2.1)$$

Unlike the SWWE the vertical velocity of a fluid particle in the Serre equations is not zero throughout the depth of water. This non-zero vertical velocity of a fluid particle leads to a non-hydrostatic pressure distribution on the fluid particles [14]

$$p'(x, z, t) = \underbrace{\rho g (h + b - z)}_{\text{hydrostatic pressure}} + \rho (h + b - z) \Psi + \frac{1}{2} \rho (h^2 - [z - b]^2) \Phi \quad (2.2)$$

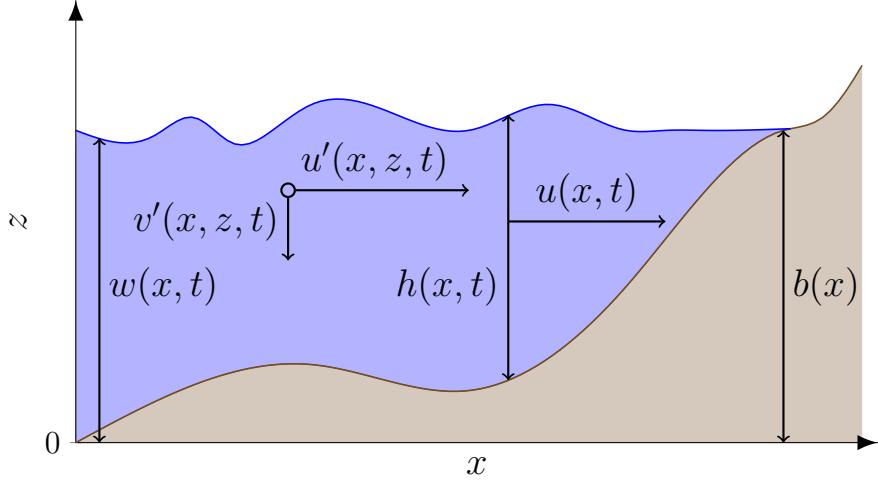


Figure 2.1: Diagram demonstrating a free surface flow (■) over a bed (□) where $w(x, t)$ is the absolute location of the free surface, $v'(x, z, t)$ is the vertical velocity of a fluid particle, $u'(x, z, t)$ is the horizontal velocity of a fluid particle, $h(x, t)$ is the height of a column of fluid, $u(x, t)$ is the average horizontal velocity of a column of fluid and $b(x)$ is the stationary bed profile.

where

$$\Psi = \frac{\partial b}{\partial x} \left(\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} \right) + u^2 \frac{\partial^2 b}{\partial x^2}, \quad (2.3a)$$

$$\Phi = \frac{\partial u}{\partial x} \frac{\partial u}{\partial x} - u \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial x \partial t}. \quad (2.3b)$$

The hydrostatic pressure term $\rho g (h + b - z)$ is the pressure due to the weight of water above the fluid particle. The non-hydrostatic pressure terms are a consequence of the non-zero vertical velocity which modifies the underlying hydrostatic pressure distribution causing dispersion of waves on the free surface.

Integrating the Euler equations [14, 21] over the entire depth with a no-slip condition at the bed, a free surface condition at the free surface, the vertical velocity relation (2.1) and the pressure distribution (2.2) the Serre equations are obtained

$$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} = 0, \quad (2.4a)$$

$$\frac{\partial(uh)}{\partial t} + \frac{\partial}{\partial x} \left(u^2 h + \frac{gh^2}{2} + \frac{h^2}{2} \Psi + \frac{h^3}{3} \Phi \right) + \frac{\partial b}{\partial x} \left(gh + h\Psi + \frac{h^2}{2} \Phi \right) = 0. \quad (2.4b)$$

The Serre equations are depth integrated approximations to the conservation of mass and horizontal momentum equations of the Euler equations. When $\Phi = \Psi = 0$ the Serre equations reduce to the SWWE where the vertical velocity is zero throughout the depth, the pressure distribution is hydrostatic and there is no dispersion of water waves.

Due to the presence of the Φ and Ψ terms the Serre equations are more difficult to solve analytically and numerically than the SWWE. The primary reason for this is that whilst the SWWE are hyperbolic the Serre equations are neither hyperbolic nor parabolic. Furthermore, the Serre equations are not in conservation law form due to the presence of temporal derivatives in Φ and Ψ , although they are derived from conservation equations.

For a horizontal bed $\partial b/\partial x = 0$, $\Psi = 0$ and so the Serre equations reduce to

$$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} = 0, \quad (2.5a)$$

$$\frac{\partial(uh)}{\partial t} + \frac{\partial}{\partial x} \left(u^2 h + \frac{gh^2}{2} + \frac{h^3}{3} \Phi \right) = 0. \quad (2.5b)$$

For a horizontal bed the Serre equations are more challenging to solve analytically and numerically than the SWWE.

2.1.1 Alternative Form

A major hurdle for developing numerical methods for the Serre equations is the presence of the temporal derivative in Φ and Ψ (2.3). Using the quantity G [14, 27, 28] the Serre equations can be rewritten in conservation law form with a source term

$$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} = 0, \quad (2.6a)$$

$$\begin{aligned} \frac{\partial G}{\partial t} + \frac{\partial}{\partial x} \left(uG + \frac{gh^2}{2} - \frac{2}{3}h^3 \left[\frac{\partial u}{\partial x} \right]^2 + h^2 u \frac{\partial u}{\partial x} \frac{\partial b}{\partial x} \right) \\ + \underbrace{\frac{1}{2}h^2 u \frac{\partial u}{\partial x} \frac{\partial^2 b}{\partial x^2} - hu^2 \frac{\partial b}{\partial x} \frac{\partial^2 b}{\partial x^2} + gh \frac{\partial b}{\partial x}}_{\text{source term}} = 0 \end{aligned} \quad (2.6b)$$

where

$$G = uh \left(1 + \frac{\partial h}{\partial x} \frac{\partial b}{\partial x} + \frac{1}{2} h \frac{\partial^2 b}{\partial x^2} + \left[\frac{\partial b}{\partial x} \right]^2 \right) - \frac{\partial}{\partial x} \left(\frac{1}{3} h^3 \frac{\partial u}{\partial x} \right). \quad (2.7)$$

The introduced quantity G resembles h multiplied by the irrotationality [29, 30].

Since the Serre equations can be written in conservation law form with a source term they can be numerically solved with a finite volume method to evolve h and G through time. However, a finite volume method for (2.6) which evolves h and G requires a numerical method to solve (2.7) for the remaining primitive variable u given h , G and b .

For a horizontal bed $\partial b / \partial x = 0$ the conservation law form of the Serre equations using the new quantity G is

$$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} = 0, \quad (2.8a)$$

$$\frac{\partial G}{\partial t} + \frac{\partial}{\partial x} \left(uG + \frac{gh^2}{2} - \frac{2}{3}h^3 \left[\frac{\partial u}{\partial x} \right]^2 \right) = 0 \quad (2.8b)$$

with

$$G = uh - \frac{\partial}{\partial x} \left(\frac{1}{3}h^3 \frac{\partial u}{\partial x} \right). \quad (2.8c)$$

2.2 Properties

The Serre equations possess a number of desirable properties for the modelling of water waves; in particular their conservation of fundamental quantities and their dispersion relation. If a numerical method accurately approximates the Serre equations then the numerical method should reproduce the conservation and dispersion properties of the Serre equations. Furthermore, the numerical method should accurately reproduce the analytic solutions of the Serre equations. In this thesis the conservation and dispersion properties and the analytic solutions of the Serre equations introduced below are used to assess the veracity of numerical methods.

2.2.1 Conservation Properties

A quantity is conserved if the total amount of a quantity q in a closed system remains constant through time.

Definition 2.1. The total amount of a quantity q in a system occurring on the interval $[a, b]$ at time t is

$$\mathcal{C}_q(t) = \int_a^b q(x, t) dx.$$

Using this notation conservation of a quantity q implies that $\mathcal{C}_q(0) = \mathcal{C}_q(t)$ for all t , provided the system is closed and thus has no flux across its boundary.

Integrating the canonical Serre equations (2.4) and the Serre equations in conservation law form with a source term (2.6) for a closed system implies that h , uh and G are conserved by the Serre equations. However, uh and G are only conserved when the contribution of their respective source terms over the domain vanishes. Additionally, the Green-Naghdi equations [22] which are equivalent to the Serre equations for one-dimensional flows were derived by conserving the energy

$$\mathcal{H}(x, t) = \frac{1}{2} \left(gh(h + 2b) + hu^2 + \frac{h^3}{3} \left[\frac{\partial u}{\partial x} \right]^2 + u^2 h \left[\frac{\partial b}{\partial x} \right]^2 - uh^2 \frac{\partial u}{\partial x} \frac{\partial b}{\partial x} \right).$$

Therefore, the one-dimensional Serre equations should also conserve \mathcal{H} . The energy \mathcal{H} is the sum of the gravitational potential energy

$$\frac{1}{2} \int_b^{h+b} gz dx = \frac{1}{2} gh(h + 2b),$$

the horizontal kinetic energy

$$\frac{1}{2} \int_b^{h+b} (u')^2 dx = \frac{1}{2} hu^2$$

and the vertical kinetic energy

$$\frac{1}{2} \int_b^{h+b} (v')^2 dx = \frac{1}{2} \left(\frac{h^3}{3} \left[\frac{\partial u}{\partial x} \right]^2 + u^2 h \left[\frac{\partial b}{\partial x} \right]^2 - uh^2 \frac{\partial u}{\partial x} \frac{\partial b}{\partial x} \right)$$

for a column of fluid. For horizontal beds \mathcal{H} is the Hamiltonian of the Serre equations [31].

2.2.2 Dispersion Properties

The dispersion properties are studied by linearising the Serre equations with a horizontal bed, assuming periodic wave solutions and then deriving a relationship

between the frequency ω and the wave number k of these solutions. For the Serre equations the dispersion relation [28] is

$$\omega^\pm = U k \pm k \sqrt{gH} \sqrt{\frac{3}{(kH)^2 + 3}} \quad (2.9)$$

where U and H are the mean velocity and height of the fluid respectively. For the dispersion relation two wave frequencies ω^+ and ω^- are possible for each wavenumber allowing for downwind and upwind travelling waves respectively. Barthélemy [24] compared this dispersion relation to the dispersion relation given by the linear theory for water waves and demonstrated its utility when k is small. However, as k increases so does the difference between the dispersion relation of the Serre equations and that of the linear water wave theory.

From the dispersion relation (2.9) the phase velocity $v_p^\pm = \omega^\pm/k$ and the group velocity $v_g^\pm = \partial\omega^\pm/\partial k$ can be written in terms of the wave number as

$$v_p^\pm = U \pm \sqrt{gH} \sqrt{\frac{3}{(kH)^2 + 3}},$$

$$v_g^\pm = U \pm \sqrt{gH} \left(\sqrt{\frac{3}{(kH)^2 + 3}} \mp (kH)^2 \sqrt{\frac{3}{([kH]^2 + 3)^3}} \right).$$

Since both the phase and group velocities depend on the wave number, waves of different wavelengths travel at different speeds indicating that the Serre equations describe dispersive waves.

Fortunately, the phase velocity and the group velocity of waves are bounded, since as $k \rightarrow 0$ then v_p^\pm and $v_g^\pm \rightarrow U \pm \sqrt{gH}$ and as $k \rightarrow \infty$ then v_p^\pm and $v_g^\pm \rightarrow U$. Therefore,

$$U - \sqrt{gH} \leq v_p^- \leq U \leq v_p^+ \leq U + \sqrt{gH}, \quad (2.10a)$$

$$U - \sqrt{gH} \leq v_g^- \leq U \leq v_g^+ \leq U + \sqrt{gH} \quad (2.10b)$$

hence, the speed of waves in the Serre equations are bounded by the speed of waves in the SWWE.

2.2.3 Analytic Solutions

Currently, few analytic solutions have been discovered for the Serre equations. There is a travelling wave solution for a horizontal bed [32] and the stationary lake at rest solution for arbitrary bathymetry.

Solitary Travelling Wave Solution

The Serre equations admit a family of travelling wave solutions that propagate at a constant speed without deformation due to a balance between the non-linear and dispersive effects [32]. One particular member of this family is the solitary travelling wave solution, which unlike the Euler equations has a closed form

$$h(x, t) = a_0 + a_1 \operatorname{sech}^2(\kappa [x - ct]), \quad (2.11a)$$

$$u(x, t) = c \left(1 - \frac{a_0}{h(x, t)} \right), \quad (2.11b)$$

$$b(x) = 0 \quad (2.11c)$$

with

$$\kappa = \frac{\sqrt{3a_1}}{2a_0 \sqrt{(a_0 + a_1)}},$$

$$c = \sqrt{g(a_0 + a_1)}.$$

From these equations the total amounts of h (A.1a), uh (A.1b), G (A.1c) and \mathcal{H} (A.1d) at $t = 0s$ can be derived and are provided in Appendix A.

The solitary wave solution has an amplitude of a_1 , an infinite wavelength and propagates on water a_0 deep. In (2.11) it is written as travelling in the positive x direction, but can travel in either direction, depending on the sign of $u(x, t)$. Although, it is a solitary travelling wave this solution is not a true soliton as these solitary waves do not collide elastically [33].

This analytic solution can only be reproduced with the appropriate order of accuracy if all terms of the Serre equations with a horizontal bed (2.8) are adequately approximated by the numerical method. Furthermore, since this solution is maintained by a balance between non-linear and dispersive effects it tests the balance of these effects in the numerical method. Therefore, this analytic solution is a good test for assessing the accuracy of numerical methods for solving the Serre equations with a horizontal bed (2.8).

Lake at Rest

The lake at rest solution is a stationary solution of the Serre equations that exists for all bathymetry $b(x)$. The lake at rest solution is maintained due to the balance

between the hydrostatic pressure distribution and the forcing of the bed slope. The lake at rest solution for an arbitrary bed profile $b(x)$ is

$$h(x, t) = \max \{a_0 - b(x), 0\}, \quad (2.12a)$$

$$u(x, t) = 0, \quad (2.12b)$$

$$G(x, t) = 0. \quad (2.12c)$$

It represents a quiescent body of water with a horizontal water surface or stage $w(x, t) = h(x, t) + b(x)$ over any bathymetry. The maximum function has been included in the water depth to allow for dry regions of the bed when $b(x) > a_0$. The total amounts of h and \mathcal{H} for the lake at rest solution can be calculated by summing their total amounts in wet regions which are given by the integrals (A.2), while the total amounts of uh and G are zero. Expressions for the total amount of all the conserved quantities over any domain are provided in Appendix A.

Since $w(x, t)$ is constant when $h > 0$ then $\partial w / \partial x = \partial h / \partial x + \partial b / \partial x = 0$ and $u = 0$ therefore, the Serre equations (2.6) reduce to

$$\frac{\partial h}{\partial t} = 0, \quad \frac{\partial G}{\partial t} = 0.$$

Hence, h , G and u are constant in time and thus the solution is stationary.

When naive numerical methods are used to solve the Serre equations the hydrostatic pressure terms and the bed slope terms are not appropriately balanced and therefore do not completely cancel. The error produced by the lack of cancellation produces non-physical velocities in initially still lakes, degrading the convergence of the numerical methods to the analytic solution. To ensure that these terms are balanced, modifications are made to the flux and source term approximations, producing a ‘well-balanced’ method. The lake at rest solution provides a test for the effectiveness of these well-balancing modifications to the numerical methods.

2.3 Forced Solutions

The known analytic solutions of the Serre equations provide a stringent test when the bed is horizontal, as all terms in the equations are non-zero and vary in space and time. For varying bathymetry there is only the lake at rest solution where all terms are constant in time and some vanish. Therefore, the accuracy of the

approximations of all terms of the Serre equations in the numerical method is not adequately assessed using only the currently available analytic solutions.

To verify the order of accuracy of the numerical methods for transient solutions with varying bathymetry forced solutions are utilised. To construct a forced solution, select some particular functions for all of the primitive quantities; h , u and b which will be denoted using $*$ as a superscript. To force the functions h^* , u^* and b^* to be solutions of some modified Serre equations the terms S_h and S_G are added to (2.6) to obtain

$$\frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} + S_h = 0, \quad (2.13a)$$

$$\frac{\partial G}{\partial t} + \frac{\partial}{\partial x} \left(uG + \frac{gh^2}{2} - \frac{2}{3}h^3 \left[\frac{\partial u}{\partial x} \right]^2 + h^2u \frac{\partial u}{\partial x} \frac{\partial b}{\partial x} \right) \\ (2.13b)$$

$$+ \frac{1}{2}h^2u \frac{\partial u}{\partial x} \frac{\partial^2 b}{\partial x^2} - hu^2 \frac{\partial b}{\partial x} \frac{\partial^2 b}{\partial x^2} + gh \frac{\partial b}{\partial x} + S_G = 0$$

where

$$S_h = -\frac{\partial h^*}{\partial t} - \frac{\partial(u^*h^*)}{\partial x},$$

$$S_G = -\frac{\partial G^*}{\partial t} - \frac{\partial}{\partial x} \left(u^*G^* + \frac{g[h^*]^2}{2} - \frac{2}{3}[h^*]^3 \left[\frac{\partial u^*}{\partial x} \right]^2 + [h^*]^2 u^* \frac{\partial u^*}{\partial x} \frac{\partial b^*}{\partial x} \right) \\ - \frac{1}{2}[h^*]^2 u^* \frac{\partial u^*}{\partial x} \frac{\partial^2 b^*}{\partial x^2} + h^*[u^*]^2 \frac{\partial b^*}{\partial x} \frac{\partial^2 b^*}{\partial x^2} - gh^* \frac{\partial b^*}{\partial x}.$$

These forced Serre equations are then numerically solved by solving the Serre equations (2.6) with the analytic values of S_h and S_G given h^* , u^* and b^* . Therefore, the only error present in the numerical solutions of the forced Serre equations is the error produced by the numerical methods used to solve the Serre equations.

Note that since the choice of the forced solutions h^* , u^* and b^* is arbitrary the solutions of the forced Serre equations need not be conservative or retain any properties of the underlying Serre equations.

2.4 Behaviour in the Presence of Steep Gradients

To ensure that the developed numerical methods are robust, their capability to handle initial condition value problems with quantities possessing discontinuities must be tested. One group of these initial condition value problems that has been of particular interest to the water wave community is the dam-break problem [27, 32, 34, 35, 36]. The dam-break problem consists of an initially still body of water with a discontinuous jump in its surface between two depth values. The dam-break problem is thus given by

$$h(x, 0) = \begin{cases} h_l & x < x_0 \\ h_r & x \geq x_0, \end{cases} \quad (2.14a)$$

$$u(x, 0) = 0, \quad (2.14b)$$

$$G(x, 0) = 0, \quad (2.14c)$$

$$b(x) = 0 \quad (2.14d)$$

where h_l and h_r are the water depth to the left and right of x_0 respectively.

Currently, there are no known analytic solutions of the Serre equations to the dam-break problem (2.14). Although, some insight into the behaviour of the evolution of these dam-break problems has been gained from asymptotic [32] and linear [37] analyses.

There have been a number of numerical solutions to dam-break problems presented in the literature [27, 32, 34, 35, 36] which have been produced by a variety of numerical methods. Some of these numerical methods could not handle discontinuous initial conditions [32, 34, 35, 36] and so smooth approximations to the initial conditions (2.14) were employed. The variety of numerical approaches has lead to different conclusions about the behaviour of the evolution of dam-break problems for the Serre equations in the literature. To resolve the differences in the literature, a comprehensive review of a particular dam-break problem using a variety of numerical methods was performed [18].

The convergence and conservation properties of many numerical solutions of the first-, second- and third-order finite difference volume methods as well as two second-order finite difference methods were investigated [18]. All the numerical solutions affirmed the existence of a previously unreported behaviour of the Serre equations in the presence of steep gradients over short time spans. This new

behaviour was justified using the convergence of the numerical solutions to one another as the resolution of the method increased. Since this convergence was observed for all high-order accurate methods; the likelihood that this new behaviour was a result of a particular quirk of the method was small. Therefore, it was concluded that this behaviour was representative of the true solution of the Serre equations in the presence of steep gradients.

The analytic solution of the Serre equations to the dam-break problem is not currently known. The numerical solutions of Pitt et al. [18] currently provide the best approximation to the solution of the Serre equations to the dam-break problem over short time spans. Furthermore, the impact of varying the resolution of the numerical methods and the smoothing applied to the initial conditions on the observed behaviour was determined. These results allow others to appropriately choose the smoothness of the initial conditions and the resolution of their numerical method.

The relevant results garnered from the asymptotic [32] and linear [37] analyses for the evolution of the dam-break problem are presented here, followed by a summary of the numerical results [18], which constituted a significant portion of the research underpinning this thesis. Since the rest of this thesis is focused around the development and validation of the second-order finite element volume method and the paper [18] is focused on other methods, only a brief summary of this work is provided in this thesis.

Asymptotic and Linear Results

The asymptotic analysis of El et al. [32] used Whitham modulation to study the evolution of dispersive shock waves of the Serre equations as $t \rightarrow \infty$. Dispersive shock waves are generated during the evolution of the dam-break problem in the Serre equations and thus the Whitham modulation results are very useful for understanding the dam-break problem. The Whitham modulation results provide a relationship between the initial heights of the dam-break problem h_l and h_r and the amplitude A^+ and speed S^+ of the leading wave in the resulting dispersive wave train. These quantities are obtained by solving

$$\frac{\Delta}{(A^+ + 1)^{1/4}} - \left(\frac{3}{4 - \sqrt{A^+ + 1}} \right)^{21/10} \left(\frac{2}{1 + \sqrt{A^+ + 1}} \right)^{2/5} = 0, \quad (2.15a)$$

$$S^+ = \sqrt{g(A^+ + 1)} \quad (2.15b)$$

where

$$\Delta = \frac{1}{4h_r} \left(\sqrt{\frac{h_l}{h_r}} + 1 \right)^2.$$

These expressions were found to agree well with numerical simulations provided that $\Delta < 1.43$ [32].

For the linearised Serre equations the dispersion relation is (2.9) from which the phase and group velocities can be determined (2.10). The negative and positive branches of the phase and group velocities are separated, implying a separation of the upwind and downwind parts of the dispersive wave-train [37]. Therefore, the structure of dispersive shock waves of the Serre equations should be two separate dispersive wave trains.

Numerical Solutions for the Smoothed Dam-break Problem

To resolve the differences present in the literature a variety of numerical methods were used to solve the smoothed dam-break problem, the most common class of smoothed versions of the dam-break problem initial conditions (2.14). The smoothed dam-break problem is given by

$$h(x, 0) = h_r + \frac{h_l - h_r}{2} \left(1 + \tanh \left(\frac{x_0 - x}{\alpha} \right) \right), \quad (2.16a)$$

$$u(x, 0) = 0, \quad (2.16b)$$

$$G(x, 0) = 0, \quad (2.16c)$$

$$b(x) = 0 \quad (2.16d)$$

where α controls the width of the transition from h_l to h_r and thus the steepness of the initial gradient in the water surface. Most of the presented numerical simulations were focused on solutions to smooth dam break problems where $h_l = 1.8m$, $h_r = 1m$ and $x_0 = 500m$ with a final time of $t = 30s$. The smoothing parameter α and the resolution of the methods were varied to investigate their influence on the observed behaviour of the numerical solution. Four structures were observed in the numerical solutions; the non-oscillatory structure, the flat structure, the node structure and the growth structure. Numerical solutions at $t = 30s$ for the mentioned h_l , h_r and x_0 values demonstrating examples of the observed structures are shown in Figure 2.2.

The growth structure was previously not reported in the literature. The growth structure was consistently observed in numerical solutions of the smoothed dam-break problem for high-order accurate methods as α and the grid resolution

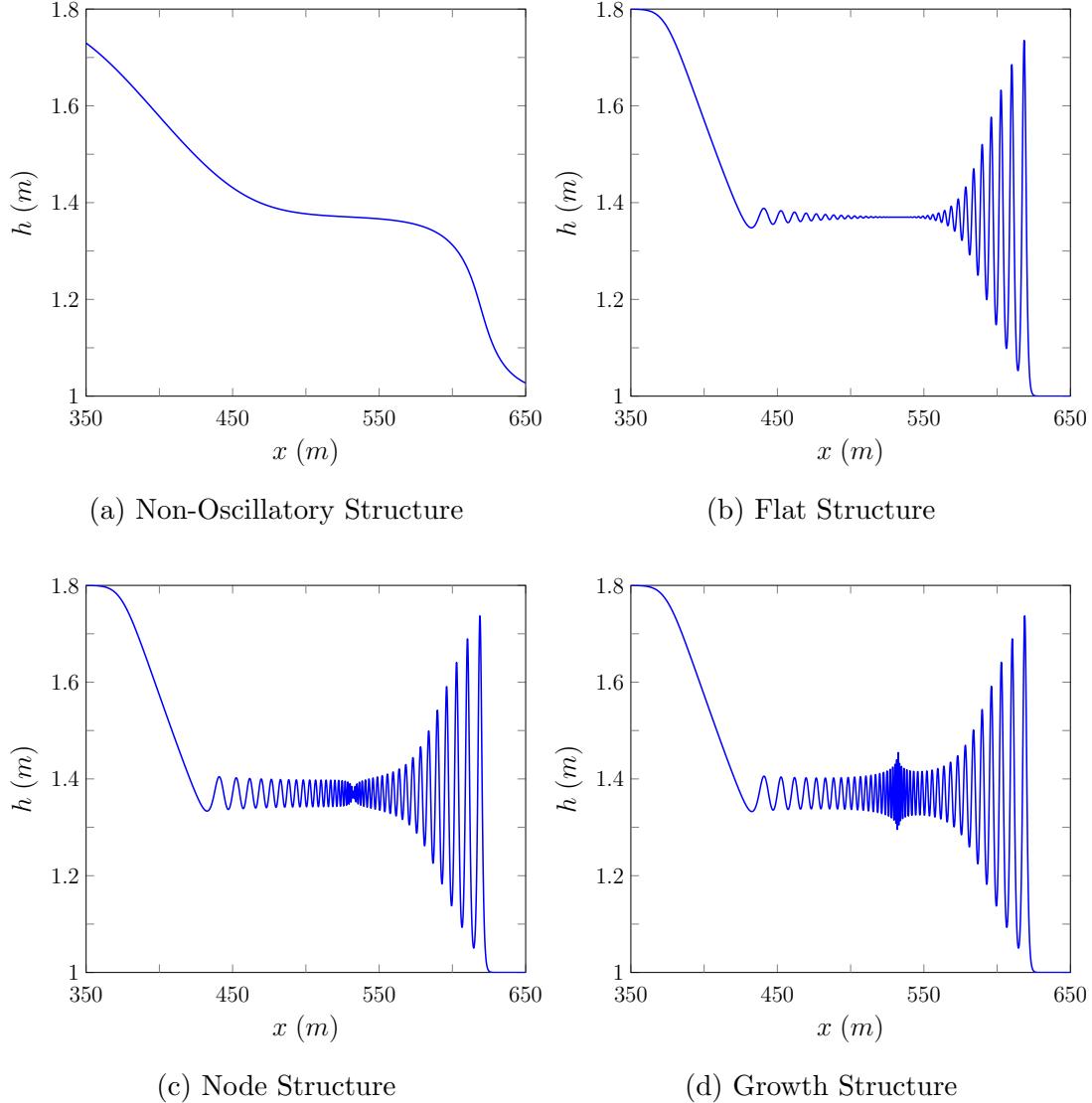


Figure 2.2: Examples of the different structures observed in numerical solutions pf the smoothed dam-break problem (2.16) displayed by Pitt et al. [18].

were increased. Therefore, the growth structure represents the true structure of the solution of the Serre equations for the dam-break problem with the corresponding h_l and h_r values at $t = 30s$. This conclusion was supported by increasing the resolution of the numerical methods and observing the convergence of the numerical solutions to one another and the conservation of h , uh and \mathcal{H} .

The finite difference volume methods introduced diffusive errors due to their flux approximation causing the amplitude and number of oscillations in the dispersive wave train to increase as the resolution increased. The finite difference methods introduce dispersive errors, causing a decrease in the amplitude and number of oscillations in the dispersive wave train as the resolution was increased. Therefore, the numerical solutions of both classes of methods serve as upper and lower bounds on the amplitude and number of oscillations in the dispersive wave train. These bounds affirmed that the observed growth structure best represented the true solution of the Serre equations to this dam-break problem.

The observation of other behaviours at $t = 30s$ was caused by; small α values which overly smooth the initial conditions, low-order numerical methods which introduce large diffusive errors and low numerical resolutions which cannot resolve the high frequency waves observed in the growth structure. The structures exist on a spectrum where the severity of these effects determine the observed behaviour. Hence, the most severe damping effects produced the non-oscillatory structure and the least severe effects produced the growth structure. These effects explained the observations of different structures previously present in the literature [27, 32, 34, 35].

The differences in the observed structures are primarily driven by the different internal structures of the dispersive shock wave while the front of the waves are indistinguishable, as demonstrated in Figure 2.2 for the flat, node and growth structures. Therefore, the results of numerical solutions that have not resolved all the internal structure present in the growth structure still agree well with the Whitham modulation results (2.15) of El et al. [32].

The amplitude of waves at the centre of the growth structure decays over time, resulting in the observation of the flat structure when t is large. These results agree with the linear argument put forth by Dougalis et al. [37]. Therefore, for smaller times the non-linear terms of the Serre equations play a significant role in the evolution of steep gradients while over longer times the linear terms dominate and thus a separation of the dispersive wave trains is observed.

In this chapter the Serre equations and their relevant properties were given.

The forced Serre equations were introduced and a summary of the main results for the evolution of steep gradients in the free surface was provided. In the following chapter the finite element volume method, whose development was a primary objective of this thesis, is described.

Chapter 3

Finite Element Volume Method

In this chapter we introduce the notation for the numerical grids and then describe the second-order Finite Element Volume Method (FEVM) in detail.

A variety of numerical methods have been used to solve the Serre equations; from complete finite difference methods [32, 38] and finite element methods [28, 34, 35] to combinations of finite difference and finite volume methods [15, 27]. Splitting techniques have also been employed, most commonly to split the Serre equations into their non-linear and dispersive parts; resulting in an elliptic operator for the dispersive part and the SWWE for the non-linear part [33, 39, 40].

Numerical methods that make use of the conservation law form of the Serre equations (2.6) [15, 27, 28] are the most promising for the two dimensional Serre equations with variable bathymetry. The primary reason for this is that these methods are robust and extend well to unstructured meshes with complex geometries which are the meshes most commonly used for modelling physical scenarios. Additionally, to properly handle the elliptic operator produced by the non-linear and dispersive splitting requires overly restrictive assumptions about the smoothness of the physical quantities, particularly the water depth.

An extension of the Finite Difference Volume Methods (FDVM) [15, 27] that uses a finite element method in place of the finite difference method was developed during the research underpinning this thesis. This second-order FEVM which will be referred to as FEVM₂ was a main objective of this thesis; it consists of two main parts a Finite Element Method (FEM) to solve (2.7) and a Finite Volume Method (FVM) to solve (2.6), hence its name. Making use of a FEM and a FVM results in a numerical method with a number of desirable properties: it is accurate in the presence of steep gradients in the free surface [18], it accurately models the wetting and drying of beds and the FEM provides a complete profile

of the horizontal velocity inside a cell. This last point indicates that this method is the ideal variant of the finite volume based methods [15] for solving the two-dimensional Serre equations on unstructured meshes with parallelised code.

In addition to the FEVM₂, the first- and second-order FDVM of Le Métayer et al. [27] and Zoppou et al. [15] were reproduced. These methods will be referred to as FDVM₁ and FDVM₂ respectively. Furthermore, the third-order FDVM₃ was implemented during the research underpinning this thesis. Finally, the second-order naive finite difference method [41] and the finite difference method of El et al. [32] which are referred to as \mathcal{D} and \mathcal{W} respectively were reproduced. Descriptions of these methods have already been published [15, 18] and therefore, are omitted from this thesis.

3.1 Notation for Numerical Grids

In the FEVM₂ time is discretised into time levels separated by a constant duration Δt while space is discretised into cells of constant width Δx . The FEVM can be extended to allow for varying Δt and Δx values, with this description restricted to the constant case for simplicity. The notation for time is quite simple; from an initial time t^0 the n^{th} time level where $n \in \mathbb{N}$ is

$$t^n = t^0 + n\Delta t.$$

The goal of FEVM₂ is to update the quantities at the current time level t^n to the next time level t^{n+1} by solving the Serre equations.

The notation for space is more complicated as multiple locations inside a cell require definition. The cells are defined by their midpoints; which are given from a starting location x_0 , thus the midpoint of the j^{th} cell where $j \in \mathbb{N}$ is

$$x_j = x_0 + j\Delta x.$$

Other points inside the j^{th} cell can be defined in relation to the midpoint so that

$$x_{j+s} = x_j + s\Delta x$$

where $s \in [-\frac{1}{2}, \frac{1}{2}] \subset \mathbb{R}$, although for the remainder of the thesis only rational values of s are required. With this notation the j^{th} cell spans $[x_{j-1/2}, x_{j+1/2}]$. These discretisations in space and time result in the grids displayed in Figure 3.1.

The temporal and spatial grid notation naturally extends to the quantities of interest, for example, for a general quantity q

$$q_j^n = q(x_j, t^n)$$

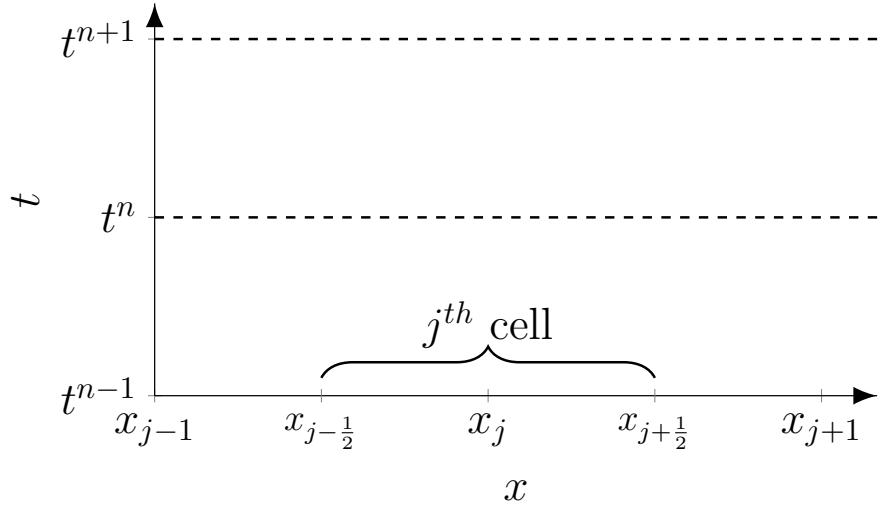


Figure 3.1: Diagram of the time levels t^{n-1} , t^n and t^{n+1} at which the numerical solution of the Serre equations will be calculated. Also shown is the j^{th} cell with midpoint x_j spanning $x_{j-1/2}$ to $x_{j+1/2}$ which is a volume of the FVM and an element of the FEM.

which are the nodal values of q . Since the FEVM uses a FVM, a definition of the cell averages of quantities are also required. The average of a quantity q over the j^{th} cell at time level t^n is

$$\bar{q}_j^n = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} q(x, t^n) dx.$$

In the FEVM the quantities are reconstructed at various points inside the cell using the adjacent cell average values. At the cell edges $x_{j\pm 1/2}$, two reconstructions are possible from each of the neighbouring cells, these two possible reconstructions are distinguished using superscripts. For example, for the cell edge $x_{j+1/2}$ and a general quantity q , there is the reconstructed value $q_{j+1/2}^-$ from the j^{th} cell and the reconstructed value $q_{j+1/2}^+$ from the $(j + 1)^{th}$ cell.

3.2 Structure Overview

The description of the FEVM begins with an overview of the evolution step, followed by the details of each of its components.

The evolution step from t^n to t^{n+1} begins with the cell averages for the height of the water above the bed h , the absolute location of the free-surface w and the conserved quantity G at time t^n . Additionally, all the nodal values of the bed

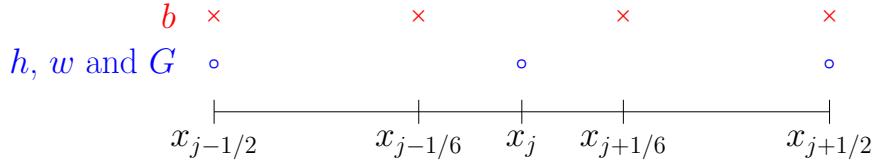


Figure 3.2: The locations of the reconstructions for h , w , G (○) and b (✖) inside the j^{th} cell.

profile b are known. These quantities are written as vectors capturing their values from the 0^{th} to the m^{th} cell in the following way

$$\bar{\mathbf{q}}^n = \begin{bmatrix} \bar{q}_0^n \\ \bar{q}_1^n \\ \vdots \\ \bar{q}_m^n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix}$$

where q is a generic quantity representing the vectors for h , G and w . The evolution step proceeds by (i) reconstructing the quantities over the cell, (ii) calculating the fluid velocity, (iii) approximating the flux, (iv) approximating the source term, (v) updating the cell averages and then (vi) applying second-order Strong Stability Preserving (SSP) Runge-Kutta time stepping.

- (i) Reconstruction: The locations for the reconstruction of all the quantities in the j^{th} cell are displayed in Figure 3.2. For the j^{th} cell the quantities h , w and G are reconstructed at $x_{j-1/2}$, x_j and $x_{j+1/2}$ from their cell average values using the second-order reconstruction operators $\mathcal{R}_{j-1/2}^+$, \mathcal{R}_j and $\mathcal{R}_{j+1/2}^-$ respectively. The bed profile b in the j^{th} cell is reconstructed at $x_{j-1/2}$, $x_{j-1/6}$, $x_{j+1/6}$ and $x_{j+1/2}$ from its nodal values using the fourth-order reconstruction operators $\mathcal{B}_{j-1/2}$, $\mathcal{B}_{j-1/6}$, $\mathcal{B}_{j+1/6}$ and $\mathcal{B}_{j+1/2}$ respectively.

Therefore, the following reconstructions are performed

$$\begin{aligned} q_{j\pm 1/2}^\pm &= \mathcal{R}_{j\pm 1/2}^\pm(\bar{\mathbf{q}}^n), & b_{j\pm 1/2} &= \mathcal{B}_{j\pm 1/2}(\mathbf{b}), \\ q_j &= \mathcal{R}_j(\bar{\mathbf{q}}^n), & b_{j\pm 1/6} &= \mathcal{B}_{j\pm 1/6}(\mathbf{b}) \end{aligned}$$

where q is a generic quantity representing h , w and G . Note that the time superscript is omitted from the reconstructed quantities to simplify the

notation. Performing the reconstruction operations on all the cells produces the following vectors; $\hat{\mathbf{h}}$, $\hat{\mathbf{w}}$, $\hat{\mathbf{G}}$ and $\hat{\mathbf{b}}$ at time t^n which are given by

$$\hat{\mathbf{q}} = \begin{bmatrix} q_{-1/2}^+ \\ q_0 \\ q_{1/2}^- \\ \vdots \\ q_{m+1/2}^- \end{bmatrix}, \quad \hat{\mathbf{b}} = \begin{bmatrix} b_{-1/2} \\ b_{-1/6} \\ b_{1/6} \\ b_{1/2} \\ \vdots \\ b_{m+1/2} \end{bmatrix}$$

where q is a generic quantity demonstrating the vectors for h , w and G . Note that since the domain contains all the cells from the 0^{th} cell to the m^{th} cell that $q_{-1/2}^+$ and $q_{m+1/2}^-$ are the reconstructed values of q at the left and right boundaries of the domain, respectively.

- (ii) Fluid Velocity: The remaining unknown quantity, the horizontal velocity of the fluid column, u is calculated at $x_{j-1/2}$, x_j and $x_{j+1/2}$ in each cell by solving (2.7) with a second-order FEM. The solution of the FEM for u is denoted by the function \mathcal{G} , which takes $\hat{\mathbf{h}}$, $\hat{\mathbf{G}}$ and $\hat{\mathbf{b}}$ as inputs and thus,

$$\hat{\mathbf{u}} = \begin{bmatrix} u_{-1/2} \\ u_0 \\ u_{1/2} \\ \vdots \\ u_{m+1/2} \end{bmatrix} = \mathcal{G}(\hat{\mathbf{h}}, \hat{\mathbf{G}}, \hat{\mathbf{b}}).$$

- (iii) Flux Across Cell Interfaces: The temporally averaged fluxes $F_{j-1/2}^n$ and $F_{j+1/2}^n$ across the cell boundaries $x_{j-1/2}$ and $x_{j+1/2}$ are calculated using $\mathcal{F}_{j-1/2}$ and $\mathcal{F}_{j+1/2}$ hence,

$$F_{j\pm 1/2}^n = \mathcal{F}_{j\pm 1/2}(\hat{\mathbf{h}}, \hat{\mathbf{G}}, \hat{\mathbf{b}}, \hat{\mathbf{u}}).$$

- (iv) Source Terms: The contribution of the source term to the cell average of a quantity over a time step S_j^n is calculated using the operator \mathcal{S}_j

$$S_j^n = \mathcal{S}_j(\hat{\mathbf{h}}, \hat{\mathbf{w}}, \hat{\mathbf{b}}, \hat{\mathbf{u}}).$$

- (v) Update Cell Averages: The cell average values are updated from time t^n to t^{n+1} with a forward Euler approximation, resulting in a method that is second-order accurate in space and first-order in time.

- (vi) Second-Order SSP Runge-Kutta Method: Steps (i)-(v) are repeated and an SSP Runge-Kutta time stepping method is employed to obtain $\bar{\mathbf{h}}$ and $\bar{\mathbf{G}}$ at t^{n+1} with second-order accuracy in space and time.

3.2.1 Reconstruction

The details for the reconstruction of h , w , G and b in the j^{th} cell at the locations shown in Figure 3.2 are now provided. For h , w and G the reconstructions are performed from the cell averages, while b is reconstructed from the nodal values. For simplicity it is assumed that the mesh is structured. The reconstruction methods described below can be extended to unstructured meshes through generalisations of the employed interpolation techniques [42, 43].

Reconstruction of the h , w and G

The quantities h , w and G are reconstructed with piecewise linear functions over a cell from neighbouring cell averages. Since h , w and G use the same reconstruction operators, a general quantity q will be used to demonstrate the operators. For the j^{th} cell the values of q are reconstructed at $x_{j-1/2}$, x_j and $x_{j+1/2}$ in the following way

$$q_{j-1/2}^+ = \mathcal{R}_{j-1/2}^+(\bar{\mathbf{q}}) = \bar{q}_j - \frac{\Delta x}{2} d_j, \quad (3.1a)$$

$$q_j = \mathcal{R}_j(\bar{\mathbf{q}}) = \bar{q}_j, \quad (3.1b)$$

$$q_{j+1/2}^- = \mathcal{R}_{j+1/2}^-(\bar{\mathbf{q}}) = \bar{q}_j + \frac{\Delta x}{2} d_j \quad (3.1c)$$

where

$$d_j = \text{minmod} \left(\theta \frac{\bar{q}_j - \bar{q}_{j-1}}{\Delta x}, \frac{\bar{q}_{j+1} - \bar{q}_{j-1}}{2\Delta x}, \theta \frac{\bar{q}_{j+1} - \bar{q}_j}{\Delta x} \right) \quad (3.2)$$

with $\theta \in [1, 2]$. The choice of the θ parameter alters the diffusion introduced by the reconstruction. When $\theta = 1$ the reconstruction introduces the most diffusion and is equivalent to the minmod reconstruction [44]. When $\theta = 2$ the reconstruction introduces the least diffusion and is equivalent to the monotized central reconstruction [45].

Definition 3.1. The minmod function

$$\text{minmod}(a_0, a_1, \dots) := \begin{cases} \min \{a_i\} & a_i > 0 \text{ for all } i \\ \max \{a_i\} & a_i < 0 \text{ for all } i \\ 0 & \text{otherwise} \end{cases}$$

takes a list of $a_i \in \mathbb{R}$. If all the inputs have the same sign then minmod returns the input with smallest absolute value, otherwise it returns zero.

The non-linear limiting used to calculate d_j ensures that the reconstruction of h , w and G inside the cell is Total Variation Diminishing (TVD) [46], hence it does not introduce non-physical oscillations and is therefore stable. The TVD property of this reconstruction is achieved by constraining the slope d_j to zero near local extrema, resulting in a piecewise constant reconstruction which is TVD. Away from local extrema d_j will be the gradient with the smallest absolute value, producing a second-order accurate reconstruction. The flexibility offered by the free parameter θ and its successful use when solving the SWWE with ANUGA [10] makes the generalised minmod limiter an attractive option for the FEVM₂.

The reconstruction operator \mathcal{R}_j is second-order accurate, due to the second-order accuracy of the midpoint quadrature rule [47]

$$\bar{q}_j^n = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} q(x, t^n) dx = q_j^n + \mathcal{O}(\Delta x^2). \quad (3.3)$$

Reconstruction of the Bed Profile

The FEVM₂ requires a reconstruction of the bed profile that is at least second-order accurate for b , $\partial b / \partial x$ and $\partial^2 b / \partial x^2$. To accomplish this b is reconstructed with a cubic polynomial $C_j(x)$ centred around x_j

$$C_j(x) = c_0 (x - x_j)^3 + c_1 (x - x_j)^2 + c_2 (x - x_j) + c_3.$$

The cubic polynomial $C_j(x)$ passes through the nodal values b_{j-2} , b_{j-1} , b_{j+1} and b_{j+2} therefore,

$$\begin{bmatrix} -8\Delta x^3 & 4\Delta x^2 & -2\Delta x & 1 \\ -\Delta x^3 & \Delta x^2 & -\Delta x & 1 \\ \Delta x^3 & \Delta x^2 & \Delta x & 1 \\ 8\Delta x^3 & 4\Delta x^2 & 2\Delta x & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} b_{j-2} \\ b_{j-1} \\ b_{j+1} \\ b_{j+2} \end{bmatrix}.$$

Solving this matrix equation the polynomial coefficients for $C_j(x)$ are obtained

$$c_0 = \frac{-b_{j-2} + 2b_{j-1} - 2b_{j+1} + b_{j+2}}{12\Delta x^3},$$

$$c_1 = \frac{b_{j-2} - b_{j-1} - b_{j+1} + b_{j+2}}{6\Delta x^2},$$

$$c_2 = \frac{b_{j-2} - 8b_{j-1} + 8b_{j+1} - b_{j+2}}{12\Delta x},$$

$$c_3 = \frac{-b_{j-2} + 4b_{j-1} + 4b_{j+1} - b_{j+2}}{6}.$$

The method requires a continuous bed profile and so the two reconstructions at the cell edge from the adjacent cells are averaged. Therefore, the reconstructing cubic for the bed profile in the j^{th} cell is the cubic which takes these values

$$b_{j-1/2} = \mathcal{B}_{j-1/2}(\mathbf{b}) = \frac{1}{2} (C_j(x_{j-1/2}) + C_{j-1}(x_{j-1/2})), \quad (3.4a)$$

$$b_{j-1/6} = \mathcal{B}_{j-1/6}(\mathbf{b}) = C_j(x_{j-1/6}), \quad (3.4b)$$

$$b_{j+1/6} = \mathcal{B}_{j+1/6}(\mathbf{b}) = C_j(x_{j+1/6}), \quad (3.4c)$$

$$b_{j+1/2} = \mathcal{B}_{j+1/2}(\mathbf{b}) = \frac{1}{2} (C_j(x_{j+1/2}) + C_{j+1}(x_{j+1/2})). \quad (3.4d)$$

3.2.2 Fluid Velocity

To calculate the remaining unknown primitive variable u from the known quantities h , G and b a FEM is used to solve (2.7). The FEM uses the weak form of (2.7) with a test function v over the spatial domain Ω which is

$$\int_{\Omega} Gv \, dx = \int_{\Omega} uh \left(1 + \frac{\partial h}{\partial x} \frac{\partial b}{\partial x} + \frac{1}{2} h \frac{\partial^2 b}{\partial x^2} + \left[\frac{\partial b}{\partial x} \right]^2 \right) v - \frac{\partial}{\partial x} \left(\frac{1}{3} h^3 \frac{\partial u}{\partial x} \right) v \, dx.$$

Integrating this equation by parts with zero Dirichlet boundary conditions produces

$$\begin{aligned} \int_{\Omega} Gv \, dx &= \int_{\Omega} uh \left(1 + \left[\frac{\partial b}{\partial x} \right]^2 \right) v \, dx + \int_{\Omega} \frac{1}{3} h^3 \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \, dx \\ &\quad - \int_{\Omega} \frac{1}{2} uh^2 \frac{\partial b}{\partial x} \frac{\partial v}{\partial x} \, dx - \int_{\Omega} \frac{1}{2} h^2 \frac{\partial b}{\partial x} \frac{\partial u}{\partial x} v \, dx. \end{aligned} \quad (3.5)$$

Assuming that time is fixed so that all the functions only vary in space, this formulation implies that by ensuring that G , h , b and $\partial b/\partial x$ have finite integrals over Ω , then u and $\partial u/\partial x$ must have finite integrals as well. To approximate the flux and source terms (2.6) requires u and $\partial u/\partial x$ to be well defined, and thus integrable. Therefore, it will be assumed that for each time t that $h, G \in \mathbb{L}^2(\Omega)$, the space of square integrable functions and $b \in \mathbb{W}^{1,2}(\Omega)$ the space of square integrable functions whose first weak derivatives are also square integrable. These assumptions imply $u \in \mathbb{W}^{1,2}(\Omega)$ when solving (3.5) and thus u and $\partial u/\partial x$ will be well defined, as desired. For a precise definition of $\mathbb{L}^2(\Omega)$ and $\mathbb{W}^{1,2}(\Omega)$, see Appendix B.

To apply the FEM to (3.5) the integration is performed over the cells and then summed to get the equation for the entire domain

$$\sum_{j=0}^m \left(\int_{x_{j-1/2}}^{x_{j+1/2}} \left[\left(uh \left(1 + \left[\frac{\partial b}{\partial x} \right]^2 \right) - \frac{1}{2} h^2 \frac{\partial b}{\partial x} \frac{\partial u}{\partial x} - G \right) v + \left(\frac{1}{3} h^3 \frac{\partial u}{\partial x} - \frac{1}{2} uh^2 \frac{\partial b}{\partial x} \right) \frac{\partial v}{\partial x} \right] dx \right) = 0 \quad (3.6)$$

which holds for all test functions v . The next step is to replace the functions for h , G , b , v and u with their corresponding basis function approximations. Expressions for these basis functions are provided in Appendix B.

Basis Function Approximations

For h and G the basis functions ψ (B.1) are used, the functions are linear inside a cell and zero elsewhere and so are not continuous as shown in Figure 3.3. These basis functions are consistent with the reconstructions of h and G which are linear inside a cell and discontinuous across the cell edges. Since these basis functions are in $\mathbb{L}^2(\Omega)$ the basis function approximations to h and G are in the appropriate function space.

The basis functions ψ produce the following representation of h and G in the FEM written using the generic quantity q

$$q = \sum_{j=0}^m \left(q_{j-1/2}^+ \psi_{j-1/2}^+ + q_{j+1/2}^- \psi_{j+1/2}^- \right). \quad (3.7)$$

To calculate the flux and source terms in (2.6b) a locally calculated second-order accurate approximation to the first derivative of u is required. To produce the appropriate approximation a quadratic representation of u in each cell is

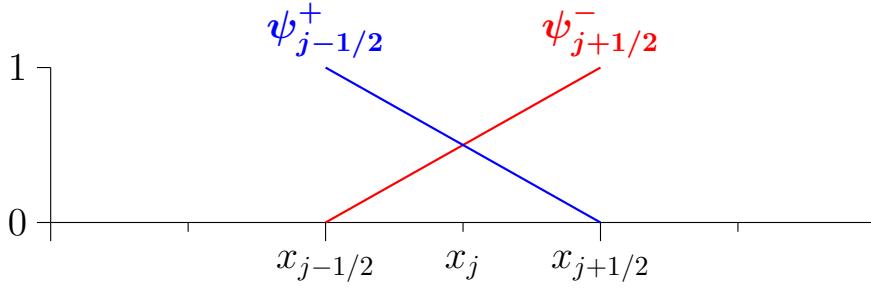


Figure 3.3: Support of the discontinuous linear basis functions ψ which are non-zero over the j^{th} cell.

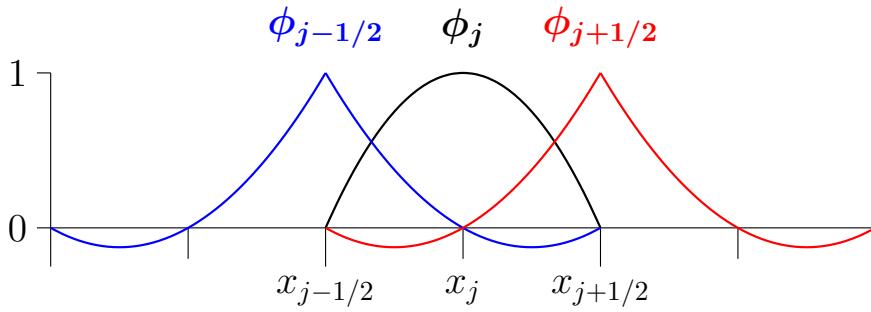


Figure 3.4: Support of the continuous piecewise quadratic basis functions ϕ which are non-zero over the j^{th} cell.

used. Furthermore, since $u \in \mathbb{W}^{1,2}(\Omega)$ the quadratic representation of u must be continuous across the cell edges $x_{j\pm 1/2}$. Therefore, the continuous quadratic basis functions $\phi_{j\pm 1/2}$ and ϕ_j (B.2) depicted in Figure 3.4 are used.

From the basis functions ϕ the basis function approximation to u is

$$u = u_{-1/2}\phi_{-1/2} + \sum_{j=0}^m (u_j\phi_j + u_{j+1/2}\phi_{j+1/2}). \quad (3.8)$$

Approximating the source term of the evolution of G equation (2.6b) requires a local approximation to the second derivative of the bed that is also second-order accurate. To allow for an appropriate second derivative of the bed profile, b must be a member of $\mathbb{W}^{2,2}(\Omega)$ which is smoother than required by (3.5). Therefore, the cubic basis functions γ (B.3) which are continuous across the cell edges are used. These basis functions are shown in Figure 3.5 and from them the basis function approximation to b is

$$b = b_{-1/2}\gamma_{-1/2} + \sum_{j=0}^m (b_{j-1/6}\gamma_{j-1/6} + b_{j+1/6}\gamma_{j+1/6} + b_{j+1/2}\gamma_{j+1/2}). \quad (3.9)$$

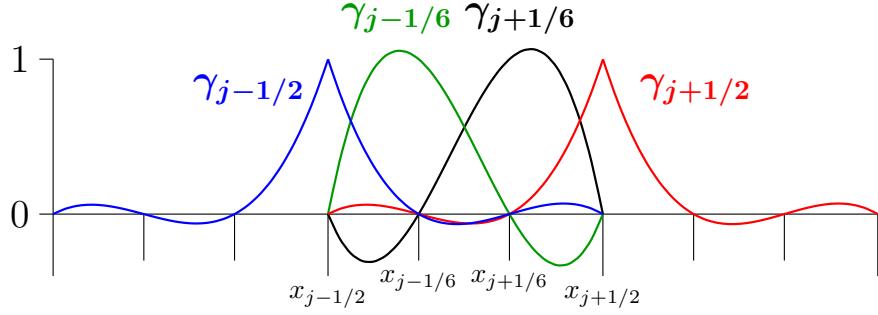


Figure 3.5: Support of the continuous piecewise cubic basis functions γ which are non-zero over the j^{th} cell.

Calculation of Elementwise Matrices

The integral equation (3.6) holds for all v . However, since the solution space has the basis functions ϕ it is sufficient to satisfy (3.6) for all ϕ to generate the solution. Since only the basis functions $\phi_{j-1/2}$, ϕ_j and $\phi_{j+1/2}$ are non-zero over the j^{th} cell the j^{th} term in the sum (3.6) can be calculated as follows

$$\begin{aligned} & \int_{x_{j-1/2}}^{x_{j+1/2}} \left(\left[uh \left(1 + \left[\frac{\partial b}{\partial x} \right]^2 \right) - \frac{1}{2} h^2 \frac{\partial b}{\partial x} \frac{\partial u}{\partial x} - G \right] \begin{bmatrix} \phi_{j-1/2} \\ \phi_j \\ \phi_{j+1/2} \end{bmatrix} \right. \\ & \quad \left. + \left[\frac{1}{3} h^3 \frac{\partial u}{\partial x} - \frac{1}{2} h^2 \frac{\partial b}{\partial x} u \right] \frac{\partial}{\partial x} \begin{bmatrix} \phi_{j-1/2} \\ \phi_j \\ \phi_{j+1/2} \end{bmatrix} \right) dx \quad (3.10) \end{aligned}$$

where the finite element approximations for h (3.7), G (3.7), u (3.8) and b (3.9) are used. This integral can be generalised for all cells by moving to the natural reference ξ -space, as the basis functions which are non-zero in one element are just translations of the non-zero basis functions in another element. The mapping from the x -space to the ξ -space is

$$x = x_j + \xi \frac{\Delta x}{2}.$$

Therefore, the j^{th} cell $[x_{j-1/2}, x_{j+1/2}]$ gets mapped to $[-1, 1]$ in the ξ -space. Performing the change of variables from x to ξ in (3.10) produces

$$\begin{aligned} \frac{\Delta x}{2} \int_{-1}^1 & \left(\left[uh \left(1 + \frac{4}{\Delta x^2} \left[\frac{\partial b}{\partial \xi} \right]^2 \right) - \frac{2}{\Delta x^2} h^2 \frac{\partial b}{\partial \xi} \frac{\partial u}{\partial \xi} - G \right] \begin{bmatrix} \phi_{j-1/2} \\ \phi_j \\ \phi_{j+1/2} \end{bmatrix} \right. \\ & \left. + \frac{4}{\Delta x^2} \left[\frac{1}{3} h^3 \frac{\partial u}{\partial \xi} - \frac{1}{2} h^2 \frac{\partial b}{\partial \xi} u \right] \frac{\partial}{\partial \xi} \begin{bmatrix} \phi_{j-1/2} \\ \phi_j \\ \phi_{j+1/2} \end{bmatrix} \right) d\xi. \end{aligned}$$

The remainder of the process will be demonstrated for the uh term as an example with the remaining integrals provided [online](https://sites.google.com/view/jordanpitt/phd-thesis-resources/finite-element-integrals) (<https://sites.google.com/view/jordanpitt/phd-thesis-resources/finite-element-integrals>). The uh term is

$$\frac{\Delta x}{2} \int_{-1}^1 uh \begin{bmatrix} \phi_{j-1/2} \\ \phi_j \\ \phi_{j+1/2} \end{bmatrix} d\xi.$$

Since the integral is computed over $[-1, 1]$, there are only a few non-zero contributions from the finite element approximations to h and u , so the uh term becomes

$$\begin{aligned} \frac{\Delta x}{2} \int_{-1}^1 & \left((u_{j-1/2} \phi_{j-1/2} + u_j \phi_j + u_{j+1/2} \phi_{j+1/2}) \right. \\ & \times \left. \left(h_{j-1/2}^+ \psi_{j-1/2}^+ + h_{j+1/2}^- \psi_{j+1/2}^- \right) \begin{bmatrix} \phi_{j-1/2} \\ \phi_j \\ \phi_{j+1/2} \end{bmatrix} \right) d\xi \\ = \frac{\Delta x}{2} & \left(h_{j-1/2}^+ \int_{-1}^1 \psi_{j-1/2}^+ \begin{bmatrix} \phi_{j-1/2} \phi_{j-1/2} & \phi_j \phi_{j-1/2} & \phi_{j+1/2} \phi_{j-1/2} \\ \phi_{j-1/2} \phi_j & \phi_j \phi_j & \phi_{j+1/2} \phi_j \\ \phi_{j+1/2} \phi_{j-1/2} & \phi_{j+1/2} \phi_j & \phi_{j+1/2} \phi_{j+1/2} \end{bmatrix} d\xi \right. \\ & \left. + h_{j+1/2}^- \int_{-1}^1 \psi_{j+1/2}^- \begin{bmatrix} \phi_{j-1/2} \phi_{j-1/2} & \phi_j \phi_{j-1/2} & \phi_{j+1/2} \phi_{j-1/2} \\ \phi_{j-1/2} \phi_j & \phi_j \phi_j & \phi_{j+1/2} \phi_j \\ \phi_{j+1/2} \phi_{j-1/2} & \phi_{j+1/2} \phi_j & \phi_{j+1/2} \phi_{j+1/2} \end{bmatrix} d\xi \right) \begin{bmatrix} u_{j-1/2} \\ u_j \\ u_{j+1/2} \end{bmatrix}. \end{aligned}$$

Calculating the integrals of all the basis function combinations yields

$$\begin{aligned} \frac{\Delta x}{2} \int_{-1}^1 u h \begin{bmatrix} \phi_{j-1/2} \\ \phi_j \\ \phi_{j+1/2} \end{bmatrix} d\xi = \\ \frac{\Delta x}{60} \begin{bmatrix} 7h_{j-1/2}^+ + h_{j+1/2}^- & 4h_{j-1/2}^+ & -h_{j-1/2}^+ - h_{j+1/2}^- \\ 4h_{j-1/2}^+ & 16h_{j-1/2}^+ + 16h_{j+1/2}^- & 4h_{j+1/2}^- \\ -h_{j-1/2}^+ - h_{j+1/2}^- & 4h_{j+1/2}^- & h_{j-1/2}^+ + 7h_{j+1/2}^- \end{bmatrix} \begin{bmatrix} u_{j-1/2} \\ u_j \\ u_{j+1/2} \end{bmatrix}. \end{aligned}$$

Assembly of the Global Matrix

By combining all the matrices generated by the integral of each of the u terms, the contribution of the j^{th} cell to the stiffness matrix \mathbf{A}_j is obtained. Likewise all the integrals of the remaining term Gv in (3.6) generate the elementwise vector \mathbf{g}_j . These elementwise matrices and vectors are then assembled into the global stiffness matrix \mathbf{A} and the global right hand-side term \mathbf{g} thus (3.6) is rewritten as

$$\mathbf{A}\hat{\mathbf{u}} = \mathbf{g}. \quad (3.11)$$

This is a penta-diagonal matrix equation which can be solved by direct banded matrix solution techniques such as those of Press et al. [48] to obtain

$$\hat{\mathbf{u}} = \mathcal{G}(\hat{\mathbf{h}}, \hat{\mathbf{G}}, \hat{\mathbf{b}}) = \mathbf{A}^{-1}\mathbf{g} \quad (3.12)$$

as desired.

The matrix \mathbf{A} is well conditioned as long as $h \gg 0$. However, as h vanishes the condition number of the matrix increases. To avoid an inaccurate solution for u when h is small the techniques described in section 3.5 are employed.

3.2.3 Flux Across the Cell Interfaces

The method of Kurganov et al. [49] is used to calculate the flux across a cell interface. This method was employed because it can handle discontinuities across the cell boundary and only requires an estimate of the maximum and minimum wave speeds. This is precisely the situation for the Serre equations which do not have a known expression for the characteristics but do possess estimates on the maximum and minimum wave speeds (2.10).

Only the calculation of the flux term $F_{j+1/2}$ is demonstrated as the process to calculate the flux term $F_{j-1/2}$ is identical but with different cells. For a general

quantity q the approximation of the flux term given by Kurganov et al. [49] is

$$F_{j+\frac{1}{2}} = \frac{a_{j+\frac{1}{2}}^+ f(q_{j+\frac{1}{2}}^-) - a_{j+\frac{1}{2}}^- f(q_{j+\frac{1}{2}}^+)}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} + \frac{a_{j+\frac{1}{2}}^+ a_{j+\frac{1}{2}}^-}{a_{j+\frac{1}{2}}^+ - a_{j+\frac{1}{2}}^-} (q_{j+\frac{1}{2}}^+ - q_{j+\frac{1}{2}}^-) \quad (3.13)$$

where $a_{j+\frac{1}{2}}^+$ and $a_{j+\frac{1}{2}}^-$ are given by bounds on the wave speed. Applying the wave speed bounds (2.10)

$$a_{j+\frac{1}{2}}^- = \min \left\{ 0, u_{j+1/2}^- - \sqrt{gh_{j+1/2}^-}, u_{j+1/2}^+ - \sqrt{gh_{j+1/2}^+} \right\}, \quad (3.14)$$

$$a_{j+\frac{1}{2}}^+ = \max \left\{ 0, u_{j+1/2}^- + \sqrt{gh_{j+1/2}^-}, u_{j+1/2}^+ + \sqrt{gh_{j+1/2}^+} \right\}. \quad (3.15)$$

The flux functions $f(q_{j+\frac{1}{2}}^-)$ and $f(q_{j+\frac{1}{2}}^+)$ across the cell edge $x_{j+1/2}$ are evaluated using the reconstructed values $q_{j+\frac{1}{2}}^-$ from the j^{th} cell and $q_{j+\frac{1}{2}}^+$ from the $(j+1)^{th}$ cell. From the continuity equation (2.6a) it is obtained that

$$f(h_{j+\frac{1}{2}}^\pm) = u_{j+1/2}^\pm h_{j+1/2}^\pm.$$

For the evolution of G equation (2.6b) it is obtained that

$$\begin{aligned} f(G_{j+\frac{1}{2}}^\pm) &= u_{j+1/2}^\pm G_{j+1/2}^\pm + \frac{g}{2} \left(h_{j+1/2}^\pm \right)^2 - \frac{2}{3} \left(h_{j+1/2}^\pm \right)^3 \left[\left(\frac{\partial u}{\partial x} \right)_{j+1/2}^\pm \right]^2 \\ &\quad + \left(h_{j+1/2}^\pm \right)^2 u_{j+1/2}^\pm \left(\frac{\partial u}{\partial x} \right)_{j+1/2}^\pm \left(\frac{\partial b}{\partial x} \right)_{j+1/2}^\pm. \end{aligned} \quad (3.16)$$

The quantities $h_{j-1/2}^+$, $h_{j+1/2}^-$, $G_{j-1/2}^+$ and $G_{j+1/2}^-$ were calculated during the reconstruction and the FEM provided $u_{j+1/2}^\pm = u_{j+1/2}$ as u is continuous across the cell boundaries.

Calculation of Derivatives

Approximations to $\left(\frac{\partial b}{\partial x} \right)_{j+1/2}^\pm$ and $\left(\frac{\partial u}{\partial x} \right)_{j+1/2}^\pm$ are required to calculate the flux (3.16). To calculate these derivatives in u and b the basis function approximation to these quantities in the FEM are used. For u the reconstructing quadratic polynomial is

$$P_j^u(x) = p_0^u (x - x_j)^2 + p_1^u (x - x_j) + p_2^u \quad (3.17)$$

which passes through $u_{j-1/2}$, u_j and $u_{j+1/2}$ (3.12). While for b the reconstructing cubic polynomial is

$$P_j^b(x) = p_0^b (x - x_j)^3 + p_1^b (x - x_j)^2 + p_2^b (x - x_j) + p_3^b \quad (3.18)$$

which passes through $b_{j-1/2}$, $b_{j-1/6}$, $b_{j+1/6}$ and $b_{j+1/2}$ (3.4). Because the cell edge values were averaged during the reconstruction of the bed, $P_j^b(x)$ will be different from $C_j(x)$.

For $P_j^u(x)$ the coefficients are

$$\begin{aligned} p_0^u &= \frac{u_{j-1/2} - 2u_j + u_{j+1/2}}{2\Delta x^2}, \\ p_1^u &= \frac{-u_{j-1/2} + u_{j+1/2}}{\Delta x}, \\ p_2^u &= u_j. \end{aligned}$$

While for $P_j^b(x)$ the coefficients are

$$\begin{aligned} p_0^b &= \frac{-9b_{j-1/2} + 27b_{j-1/6} - 27b_{j+1/6} + 9b_{j+1/2}}{2\Delta x^3}, \\ p_1^b &= \frac{9b_{j-1/2} - 9b_{j-1/6} - 9b_{j+1/6} + 9b_{j+1/2}}{4\Delta x^2}, \\ p_2^b &= \frac{b_{j-1/2} - 27b_{j-1/6} + 27b_{j+1/6} - b_{j+1/2}}{8\Delta x}, \\ p_3^b &= \frac{-b_{j-1/2} + 9b_{j-1/6} + 9b_{j+1/6} - b_{j+1/2}}{16}. \end{aligned}$$

Taking the derivative of the polynomials (3.17) and (3.18) produces

$$\begin{aligned} \frac{\partial}{\partial x} P_j^u(x) &= 2p_0^u(x - x_j) + p_1^u, \\ \frac{\partial}{\partial x} P_j^b(x) &= 3p_0^b(x - x_j)^2 + 2p_1^b(x - x_j) + p_2^b. \end{aligned}$$

This is a second-order approximation to the derivative of u and b at $x_{j+1/2}$ for the j^{th} cell. The process for the $(j+1)^{th}$ cell is the same and therefore,

$$\begin{aligned} \left(\frac{\partial u}{\partial x} \right)_{j+1/2}^- &= \frac{\partial}{\partial x} P_j^u(x_{j+1/2}), \\ \left(\frac{\partial u}{\partial x} \right)_{j+1/2}^+ &= \frac{\partial}{\partial x} P_{j+1}^u(x_{j+1/2}), \\ \left(\frac{\partial b}{\partial x} \right)_{j+1/2}^- &= \frac{\partial}{\partial x} P_j^b(x_{j+1/2}), \\ \left(\frac{\partial b}{\partial x} \right)_{j+1/2}^+ &= \frac{\partial}{\partial x} P_{j+1}^b(x_{j+1/2}). \end{aligned}$$

Hence, all the terms needed to calculate the approximation to the flux (3.13) for h and G are possessed. However, to ensure that the FEVM is well-balanced and recovers the lake at rest steady state solution, the approximation to the intercell fluxes must be modified.

Well-Balancing Modification to Flux Approximation

To recover the lake at rest steady state solution the method of Audusse et al. [50] is employed. This method was initially designed for the SWWE and has been extended to the Serre equations [41]. To enforce well-balancing the reconstruction of h is modified at the cell edges in the following way.

First calculate

$$\dot{b}_{j+1/2}^- = w_{j+1/2}^- - h_{j+1/2}^-, \quad \dot{b}_{j+1/2}^+ = w_{j+1/2}^+ - h_{j+1/2}^+. \quad (3.19)$$

Find the maximum

$$\ddot{b}_{j+1/2} = \max \left\{ \dot{b}_{j+1/2}^-, \dot{b}_{j+1/2}^+ \right\}$$

then define

$$\ddot{h}_{j+1/2}^- = \max \left\{ 0, w_{j+1/2}^- - \ddot{b}_{j+1/2} \right\}, \quad (3.20a)$$

$$\ddot{h}_{j+1/2}^+ = \max \left\{ 0, w_{j+1/2}^+ - \ddot{b}_{j+1/2} \right\}. \quad (3.20b)$$

This generates the vector $\ddot{\mathbf{h}}$

$$\ddot{\mathbf{h}} = \begin{bmatrix} \ddot{h}_{-1/2}^+ \\ h_0 \\ \ddot{h}_{1/2}^- \\ \vdots \\ \ddot{h}_{m+1/2}^- \end{bmatrix}$$

which is used to calculate the flux term $F_{j+1/2}$ in (3.13) for h and G instead of $\hat{\mathbf{h}}$. Applying the same process but with different cells $F_{j-1/2}$ is obtained and thus

$$F_{j\pm 1/2}^n = \mathcal{F}_{j\pm 1/2} \left(\ddot{\mathbf{h}}, \hat{\mathbf{G}}, \hat{\mathbf{b}}, \hat{\mathbf{u}} \right)$$

for the evolution of h and G equations as desired. The modification to the intercell flux as well as the process to approximate the source terms described below ensures that the method recovers the lake at rest steady state, as shown in Chapter 5.

3.2.4 Source Terms

To evolve the Serre equations (2.6), an approximation to the source term at the cell centre x_j from time t^n to t^{n+1} is required, it will be denoted as S_j^n . Equation (2.6a) has no source term, therefore only the calculation of the source term for equation (2.6b) is presented.

Following the work of Audusse et al. [50] to produce a well-balanced method, the approximation to S_j^n is split into the naive source term approximation S_{ci} and the corrective interface source terms $S_{j+\frac{1}{2}}^-$ and $S_{j+\frac{1}{2}}^+$

$$S_j^n = S_{j+\frac{1}{2}}^- + \Delta x S_{ci} + S_{j-\frac{1}{2}}^+.$$

The corrective interface terms $S_{j+\frac{1}{2}}^-$ and $S_{j+\frac{1}{2}}^+$ together with the modifications to the intercell flux ensure that the flux and source terms cancel for the lake at rest solution.

The centred source term is calculated like so

$$S_{ci} = -\frac{1}{2} (h_j)^2 u_j \left(\frac{\partial u}{\partial x} \right)_j \left(\frac{\partial^2 b}{\partial x^2} \right)_j + h_j (u_j)^2 \left(\frac{\partial b}{\partial x} \right)_j \left(\frac{\partial^2 b}{\partial x^2} \right)_j - g h_j \left(\frac{\partial b}{\partial x} \right)_j.$$

Where h_j from the reconstruction process (3.1) and u_j from the solution of (3.12) are used. To calculate the derivatives of u and b their polynomial representations (3.17) and (3.18) inside a cell are used. However, to ensure that the terms cancel properly for a lake at rest the approximation to $\frac{\partial b}{\partial x}$ is modified to use $\dot{b}_{j+1/2}^-$ and $\dot{b}_{j+1/2}^+$ from (3.19). Therefore, the following approximations are used to calculate S_{ci}

$$\begin{aligned} \left(\frac{\partial u}{\partial x} \right)_j &= \frac{\partial}{\partial x} P_j^u(x_j), \\ \left(\frac{\partial b}{\partial x} \right)_j &= \frac{\dot{b}_{j+1/2}^- - \dot{b}_{j-1/2}^+}{\Delta x}, \\ \left(\frac{\partial^2 b}{\partial x^2} \right)_j &= \frac{\partial^2}{\partial x^2} P_j^b(x_j). \end{aligned}$$

To ensure well-balancing the corrective interface source terms

$$S_{j+\frac{1}{2}}^- = \frac{g}{2} \left(\ddot{h}_{j+\frac{1}{2}}^- \right)^2 - \frac{g}{2} \left(h_{j+\frac{1}{2}}^- \right)^2,$$

$$S_{j-\frac{1}{2}}^+ = \frac{g}{2} \left(h_{j-\frac{1}{2}}^+ \right)^2 - \frac{g}{2} \left(\ddot{h}_{j-\frac{1}{2}}^+ \right)^2$$

are also added. These corrective terms make use of $h_{j+\frac{1}{2}}^-$ and $h_{j+\frac{1}{2}}^+$ obtained from the reconstruction (3.1) and the modified values $\ddot{h}_{j+\frac{1}{2}}^-$ and $\ddot{h}_{j+\frac{1}{2}}^+$ from (3.20). Combining the centred and interface source terms the approximation to the source term for G is

$$S_j^n = \mathcal{S}_j(\hat{\mathbf{h}}, \ddot{\mathbf{h}}, \hat{\mathbf{w}}, \hat{\mathbf{b}}, \hat{\mathbf{u}}) = S_{j+\frac{1}{2}}^- + \Delta x S_{ci} + S_{j-\frac{1}{2}}^+.$$

This operator \mathcal{S}_j is slightly different to the example given in the overview of the evolution step as it takes $\hat{\mathbf{h}}$ and $\ddot{\mathbf{h}}$ as inputs. This change was made to make clear that $\ddot{\mathbf{h}}$ is required to calculate S_j^n . This would be obscured by a more consistent notation that is possible since $\ddot{\mathbf{h}}$ only depends on $\hat{\mathbf{h}}$, $\hat{\mathbf{w}}$ and $\hat{\mathbf{b}}$.

3.2.5 Update Cell Averages

Applying a forward Euler approximation with the approximation to the flux and source terms it is obtained that

$$\bar{q}_j^{n+1} = \bar{q}_j^n + \frac{\Delta t}{\Delta x} \left(F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n + S_j^n \right) \quad (3.21)$$

where $F_{j+\frac{1}{2}}^n$, $F_{j-\frac{1}{2}}^n$ and S_j^n are all calculated using the quantities at time t^n . This update formula is first-order in time.

3.2.6 Second-Order SSP Runge-Kutta Method

To increase the order of accuracy in time a Strong Stability Preserving (SSP) Runge-Kutta method [51] is employed. The second-order SSP Runge-Kutta method is a convex combination of the first-order time steps (3.21) in the following way

$$\bar{q}_j^{(1)} = \bar{q}_j^n + \frac{\Delta t}{\Delta x} \left(F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n + S_j^n \right), \quad (3.22a)$$

$$\bar{q}_j^{(2)} = \bar{q}_j^{(1)} + \frac{\Delta t}{\Delta x} \left(F_{j+\frac{1}{2}}^{(1)} - F_{j-\frac{1}{2}}^{(1)} + S_j^{(1)} \right), \quad (3.22b)$$

$$\bar{q}_j^{n+1} = \frac{1}{2} \left(\bar{q}_j^n + \bar{q}_j^{(2)} \right). \quad (3.22c)$$

The second-order SSP Runge-Kutta method is a time stepping method that preserves the stability of the first-order method (3.21) and is second-order accurate in time. Since all the spatial approximations are second-order accurate, the steps (i)-(vi) should result in a second-order accurate FEVM for the Serre equations, as desired.

3.3 CFL condition

To ensure the stability of the FEVM the Courant-Friedrichs-Lowy (CFL) condition [52] is used, as it is a necessary condition for stability. The CFL condition ensures that time steps are small enough so that information is only transferred between neighbouring cells. For the Serre equations the CFL condition is

$$\Delta t \leq \frac{Cr}{\max_j \{a_{j+1/2}^\pm\}} \Delta x \quad (3.23)$$

where $a_{j+1/2}^\pm$ are the wave-speed bounds used in the flux approximation (3.15) and $0 \leq Cr \leq 1$ is the Courant number. Typically, the conservative $Cr = 0.5$ is used in the numerical experiments in this thesis.

3.4 Boundary Conditions

To numerically model the Serre equations over finite spatial domains boundary conditions must be enforced at the left and right edge of the domain; $x_{-1/2}$ and $x_{m+1/2}$ respectively. For simplicity, only Dirichlet boundary conditions for the FEVM were considered in this thesis. These Dirichlet boundary conditions are enforced using ghost cells located outside the domain boundaries. These ghost cells contain the complete representation of their respective quantities over the cell. Additionally, since a finite volume based method is employed, only a first-order accurate approximation at the boundaries is required to maintain global second-order accuracy [53]. For h , w , G and u one ghost cell at each boundary is required, while for b two ghost cells at each boundary are required. The ghost cells for h , w and G written for a generic quantity q are

$$\hat{\mathbf{q}}_{-1} = \begin{bmatrix} q_{-3/2}^+ \\ q_{-1}^- \\ q_{-1/2}^- \end{bmatrix}, \quad \hat{\mathbf{q}}_{m+1} = \begin{bmatrix} q_{m+1/2}^+ \\ q_{m+1}^- \\ q_{m+3/2}^- \end{bmatrix}.$$

For u and b the ghost cells are

$$\hat{\mathbf{u}}_{-1} = \begin{bmatrix} u_{-3/2} \\ u_{-1} \\ u_{-1/2} \end{bmatrix}, \quad \hat{\mathbf{u}}_{m+1} = \begin{bmatrix} u_{m+1/2} \\ u_{m+1} \\ u_{m+3/2} \end{bmatrix},$$

$$\hat{\mathbf{b}}_{-2} = \begin{bmatrix} b_{-5/2} \\ b_{-13/6} \\ b_{-11/6} \\ b_{-3/2} \end{bmatrix}, \quad \hat{\mathbf{b}}_{-1} = \begin{bmatrix} b_{-3/2} \\ b_{-7/6} \\ b_{-5/6} \\ b_{-1/2} \end{bmatrix}, \quad \hat{\mathbf{b}}_{m+1} = \begin{bmatrix} b_{m+1/2} \\ b_{m+5/6} \\ b_{m+7/6} \\ b_{m+3/2} \end{bmatrix}, \quad \hat{\mathbf{b}}_{m+2} = \begin{bmatrix} b_{m+3/2} \\ b_{m+11/6} \\ b_{m+13/6} \\ b_{m+5/2} \end{bmatrix}.$$

To ensure that the solution of u using (3.12) agrees with the boundary conditions $\hat{\mathbf{u}}_{-1}$ and $\hat{\mathbf{u}}_m$ the elementwise stiffness matrices \mathbf{A}_0 and \mathbf{A}_m and vectors \mathbf{g}_0 and \mathbf{g}_m must be modified in the following way

$$\mathbf{A}_0 = \begin{bmatrix} 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \mathbf{g}_0 = \begin{bmatrix} u_{-1/2} \\ g_1 \\ g_2 \end{bmatrix},$$

$$\mathbf{A}_m = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{g}_m = \begin{bmatrix} g_0 \\ g_1 \\ u_{m+1/2} \end{bmatrix}.$$

These are then assembled with the other element contributions in the global stiffness matrix \mathbf{A} and right hand side vector \mathbf{g} in (3.11).

3.5 Dry Beds

Dry beds are handled adequately by all steps of the FEVM in their current form, except the FEM for u . A dry bed presents two issues; when h and G are small then small errors in h and G can produce large errors in u leading to instabilities and when $h = 0$ the stiffness matrix \mathbf{A} (3.12) becomes singular.

The issue of large errors in u when h is small also arises when solving the SWWE; due to $u = (uh)/h$ being undefined as uh and h go to zero. For the Serre equations with horizontal beds when $h \ll 1$ (2.8c) becomes

$$G = uh + \mathcal{O}(h^3). \quad (3.24)$$

Since $h \ll 1$ the $\mathcal{O}(h^3)$ terms are neglected, and thus when h is small G is equal to the momentum uh , and the challenges posed by $h \rightarrow 0$ for the SWWE and the Serre equations are equivalent. Therefore, the dry bed handling techniques for the SWWE can be applied to the Serre equations; in particular a desingularisation transformation [54].

These desingularisation transforms act by modifying the calculation of u given h and uh to avoid the singularity as the numerator and denominator go to zero, hence their name. The simplest such transformation is

$$u = \frac{(uh)h}{h(h + h_{base})} \quad (3.25)$$

where h_{base} is a small chosen parameter. The analytical error introduced by this transformation decreases as h_{base} decreases. However, as noted by Kurganov and Petrova [54] small values of h_{base} lead to large numerical errors in the calculation of u . To avoid such errors h_{base} can be made larger or following Kurganov and Petrova [54] different desingularisation transformations can be employed. For the main purpose of this thesis; the validation tests reported in Chapter 5 the simpler transformation with small values of h_{base} were more useful, although large numerical errors in u were possible for small values of h .

To adapt the calculation of u in (3.25) to (2.7) it is viewed as a transformation of the quantity h which is equivalent to

$$h \rightarrow h \left(\frac{h + h_{base}}{h} \right). \quad (3.26)$$

This transformation is ill-defined when $h = 0$ so a small term h_{tol} is added to the denominator. The term h_{tol} serves as the cut-off value with any cells with $h < h_{tol}$ being considered dry. Therefore, the transformation for the reconstructed values of h in the finite element method is

$$h_{j-1/2}^+ = h_{j-1/2}^+ \left(\frac{h_{j-1/2}^+ + h_{base}}{h_{j-1/2}^+ + h_{tol}} \right), \quad (3.27a)$$

$$h_{j+1/2}^- = h_{j+1/2}^- \left(\frac{h_{j+1/2}^- + h_{base}}{h_{j+1/2}^- + h_{tol}} \right) \quad (3.27b)$$

where on the right hand side are the reconstructed values of h from (3.1) and the left hand side are the values of h used to defined the basis functions of the FEM (3.7). This transformation is applied to all terms in the FEM avoiding the singularity as $h \rightarrow 0$; and in the case where $G = uh$ the transformation is equivalent to (3.25) for the SWWE.

Even with the transform (3.27), the matrix \mathbf{A} can become singular. The methods of Zoppou et al. [15] made use of direct banded matrix solvers such as the Thomas algorithm [55] to solve (3.12) which rely on non-singular matrices

making them unsuitable when $h = 0$. This was resolved by employing an LU decomposition algorithm described by Press et al. [48]. This algorithm solves banded matrix problems using an LU decomposition with partial pivoting. When the value of a pivot is below some tolerance value p_{tol} the algorithm replaces its value with p_{tol} , avoiding the errors associated with small pivot values during an LU decomposition. Furthermore, this LU decomposition retains the banded matrix structure, and so is not as memory intensive as a standard LU decomposition. Typically, the value $p_{tol} = 10^{-20}$ was used, allowing the matrix solver to accurately invert \mathbf{A} and thus solve (3.12) when $h = 0$.

Solving (3.12) using the LU decomposition algorithm of Press et al. [48] where the transformation (3.27) has been applied to the reconstructed values of h an approximation to u is obtained that is valid in the presence of dry beds. Additionally, to avoid numerical errors becoming dominant when h is very small a cut-off is placed on h beyond which $h = G = u = 0$ and the cells are dry; this is given by h_{tol} . The drying of the cells is performed for the whole cell based on the cell average value of h so that if $\bar{h}_j \leq h_{tol}$ then

$$\begin{array}{lll} h_{j-1/2}^+ = 0 & G_{j-1/2}^+ = 0 & w_{j-1/2}^+ = b_{j-1/2} \\ h_j = 0 & G_j = 0 & w_j = b_j, \\ h_{j+1/2}^- = 0 & G_{j+1/2}^- = 0 & w_{j+1/2}^- = b_{j+1/2} \end{array}$$

and

$$\begin{array}{lll} u_{j-1/2} = 0 & \text{if} & \bar{h}_{j-1} \leq h_{tol} \\ u_j = 0 & & \\ u_{j+1/2} = 0 & \text{if} & \bar{h}_{j+1} \leq h_{tol}. \end{array}$$

The drying procedure occurs after the solution of (3.12). In the numerical experiments the typical values used were $h_{tol} = 10^{-12}$ and $h_{base} = 10^{-8}$.

In this chapter FEVM₂ was described, including the details for the well-balancing and dry bed handling procedures. A linear analysis of the convergence and dispersion properties of FEVM₂ will now be performed.

Chapter 4

Linear Analysis

In this chapter a linear analysis is used to study the convergence and dispersion properties of:

- FEVM₂ which is a second-order Finite Element Volume Method (FEVM) described in Chapter 3.
- The Finite Difference Volume Methods (FDVM) of first- (FDVM₁), second- (FDVM₂) and third-order (FDVM₃) described by Zoppou et al. [15].
- The second-order finite difference methods \mathcal{D} and \mathcal{W} described by Pitt et al. [18].

A full description of the linear analysis is provided for FEVM₂ as an example, while only the results are presented for FDVM₁, FDVM₂, FDVM₃, \mathcal{D} and \mathcal{W} . Supplementary details for the linear analysis of FDVM₁, FDVM₂, FDVM₃, \mathcal{D} and \mathcal{W} can be found in Appendix C.

An important property of a numerical method is convergence. Convergence guarantees that as the spatial and temporal resolution of a numerical method is increased, then the numerical solution approaches the solution of the partial differential equations it approximates. For linear partial differential equations the Lax-equivalence theorem states that a numerical method is convergent if and only if it is stable and consistent [56]. A numerical scheme is consistent if the error introduced by the numerical method over a time step approaches zero as the spatial and temporal resolution increases. A numerical method is stable if the errors from previous time steps are not amplified over subsequent time steps.

Another important attribute of a numerical method modelling dispersive wave equations, such as the Serre equations is its dispersion properties. The dispersion relation of a system determines the phase and group velocity of travelling

waves in that system. The Serre equations possess a dispersion relation that well approximates the dispersion relation given by linear theory for water waves [24]. Therefore, how well the dispersion relation of a numerical method approximates the dispersion relation of the Serre equations is of particular interest.

The convergence and dispersion properties of the whole numerical method applied to the solution of the linearised Serre equations with a horizontal bed were analysed. The whole scheme is considered as the spatial and temporal approximations were analysed simultaneously. The effect of variations in the bed and non-linear terms are important when studying the convergence properties of numerical methods solving the full Serre equations. However, these effects greatly increase the complexity of the convergence analysis. In this thesis the convergence properties of the non-linear Serre equations with varying bathymetry are estimated by investigating the linearised Serre equations with a horizontal bed. In general, it is expected that a numerical method with poor convergence properties for the linearised Serre equations with a horizontal bed will also possess poor convergence properties when the bed and non-linear terms are included.

The dispersion properties of the Serre equations are derived from the linearised Serre equations with a horizontal bed [15]. Because the dispersion analysis includes the spatial and temporal approximations simultaneously the presented analysis of the dispersion properties of the numerical method is a complete analysis, which extends the work of Filippini et al. [40].

The linear analyses of convergence and dispersion properties for the finite volume based methods rely on establishing a relationship of the form

$$\left[\begin{array}{c} \bar{h} \\ \bar{G} \end{array} \right]_j^{n+1} = \mathbf{E} \left[\begin{array}{c} \bar{h} \\ \bar{G} \end{array} \right]_j^n \quad (4.1)$$

where \mathbf{E} is the 2×2 evolution matrix relating the cell average conserved quantities h and G at time level t^n with the cell average conserved quantities at time level t^{n+1} . For a linear analysis the evolution matrix \mathbf{E} will be independent of n and j . The evolution matrix \mathbf{E} is obtained in the analyses by representing the numerical solution in terms of Fourier modes. By analysing the properties of \mathbf{E} and comparing it with the exact evolution matrix the convergence and dispersion properties of its associated numerical method can be determined.

The evolution matrix \mathbf{E} (4.1) is derived for FEVM₂ and then used to perform a convergence and dispersion analysis. Additionally, the results of the analyses for FDVM₁, FDVM₂, FDVM₃, \mathcal{D} and \mathcal{W} are also provided. Appendix C provides a description of the process to produce the evolution matrix \mathbf{E} for FDVM₁, FDVM₂,

FDVM₃ and provides the evolution matrices for \mathcal{D} and \mathcal{W} . The results in this thesis extend those of Zoppou et al. [15] by including more methods, analysing the convergence properties and allowing non-zero background mean velocities.

4.1 Linearised Serre Equations

The Serre equations with a horizontal bed (2.5) are linearised by considering waves as small perturbations $\delta \times \eta(x, t)$ and $\delta \times \mu(x, t)$ on a flow with a mean height H and a mean velocity U respectively, where $\delta \ll 1$. Therefore,

$$h(x, t) = H + \delta\eta(x, t) + \mathcal{O}(\delta^2), \quad (4.2a)$$

$$u(x, t) = U + \delta\mu(x, t) + \mathcal{O}(\delta^2). \quad (4.2b)$$

These waves are relatively small so terms of order δ^2 are negligible. Substituting (4.2) into the Serre equations and neglecting terms of order δ^2 one obtains

$$\frac{\partial(\delta\eta)}{\partial t} + H \frac{\partial(\delta\mu)}{\partial x} + U \frac{\partial(\delta\eta)}{\partial x} = 0, \quad (4.3a)$$

$$H \frac{\partial(\delta\mu)}{\partial t} + gH \frac{\partial(\delta\eta)}{\partial x} + UH \frac{\partial(\delta\mu)}{\partial x} - \frac{H^3}{3} \left(U \frac{\partial^3(\delta\mu)}{\partial x^3} + \frac{\partial^3(\delta\mu)}{\partial x^2 \partial t} \right) = 0 \quad (4.3b)$$

and for G

$$G = UH + U\delta\eta + H\delta\mu - \frac{H^3}{3} \frac{\partial^2(\delta\mu)}{\partial x^2}. \quad (4.3c)$$

Absorbing the δ factor into the corresponding η and μ terms and rewriting these equations in conservation law form for η and G produces

$$\frac{\partial\eta}{\partial t} + \frac{\partial}{\partial x} (H\mu + U\eta) = 0, \quad (4.4a)$$

$$\frac{\partial G}{\partial t} + \frac{\partial}{\partial x} (UG + UH\mu + gH\eta) = 0 \quad (4.4b)$$

where

$$G = UH + U\eta + H\mu - \frac{H^3}{3} \frac{\partial^2\mu}{\partial x^2}. \quad (4.4c)$$

4.2 Evolution Matrix

To derive the evolution matrix, \mathbf{E} the behaviour of (4.4) when η and μ are Fourier modes is studied. A Fourier mode $q(x, t)$ is

$$q(x, t) = q(0, 0) e^{i(\omega t + kx)} \quad (4.5)$$

where k is the wavenumber, ω^\pm is the frequency (2.9) and i is the imaginary number. The Fourier modes are the eigenfunctions of these linearised Serre equations (4.4). Since the eigenfunctions form a basis of the solution space, their dispersion and convergence properties are inherited by all solutions of (4.4). Therefore, it is sufficient to study only the convergence and dispersion properties for Fourier mode solutions captured by the evolution matrix \mathbf{E} .

A consequence of a quantity q being a Fourier mode represented on a uniform temporal and spatial grid is that for any real numbers m and l ,

$$q_{j+l}^{n+m} = q_j^n e^{i(m\omega^\pm \Delta t + lk\Delta x)}. \quad (4.6)$$

Since η and μ are Fourier modes then so is G . Furthermore, the cell averages of these quantities $\bar{\eta}$, $\bar{\mu}$ and \bar{G} are Fourier modes as well.

For Fourier modes the operators $\mathcal{R}_{j-1/2}^+$, \mathcal{R}_j , $\mathcal{R}_{j+1/2}^-$, \mathcal{G} , $\mathcal{F}_{j-1/2}$ and $\mathcal{F}_{j+1/2}$ from Chapter 3 act as factors on their inputs. These factors vary with H , U , k , ω^\pm , Δx and Δt and are independent of j and n . By combining the factors associated with these operators the evolution matrix \mathbf{E} can be derived for FEVM₂ for the linearised Serre equations with a horizontal bed. Since all the constituent factors that produce \mathbf{E} are independent of j and n then \mathbf{E} will also be independent of j and n , as desired. The expressions for all these factors are now derived, following the structure of the method laid out in Section 3.2. Since the linearised Serre equations with a horizontal bed have no source terms step (iv), which approximates the source terms is not necessary.

4.2.1 Reconstruction

Given the cell averages $\bar{\eta}$ and \bar{G} at t^n , the first step of the numerical method is to reconstruct η and G inside the j^{th} cell at $x_{j-1/2}$, x_j and $x_{j+1/2}$ using $\mathcal{R}_{j-1/2}^+$, \mathcal{R}_j and $\mathcal{R}_{j+1/2}^-$ from (3.1). Since η and G are Fourier modes and therefore smooth non-linear limiters are not required to ensure the scheme is TVD and thus the slope $d_j = (-\bar{q}_{j-1} + \bar{q}_{j+1}) / (2\Delta x)$ is used in the reconstruction. Applying (4.6) to the reconstructions (3.1) with the centred slope approximation produces

$$q_{j-\frac{1}{2}}^+ = \bar{q}_j - \frac{-\bar{q}_j e^{-ik\Delta x} + \bar{q}_j e^{ik\Delta x}}{4} = \left(1 - \frac{i \sin(k\Delta x)}{2}\right) \bar{q}_j = \mathcal{R}_{j-1/2}^+ \bar{q}_j, \quad (4.7a)$$

$$q_j = \bar{q}_j = \mathcal{R}_j \bar{q}_j, \quad (4.7b)$$

$$q_{j+\frac{1}{2}}^- = \bar{q}_j + \frac{-\bar{q}_j e^{-ik\Delta x} + \bar{q}_j e^{ik\Delta x}}{4} = \left(1 + \frac{i \sin(k\Delta x)}{2}\right) \bar{q}_j = \mathcal{R}_{j+1/2}^- \bar{q}_j. \quad (4.7c)$$

Note that the factors given by the reconstructions operators $\mathcal{R}_{j-1/2}^+$, \mathcal{R}_j and $\mathcal{R}_{j+1/2}^-$ are independent of j . Furthermore, from (4.6) implies that the reconstructions of η and G at other locations are $\mathcal{R}_{j-1/2}^+$, \mathcal{R}_j and $\mathcal{R}_{j+1/2}^-$ multiplied by a shift factor. In particular, the reconstruction operator $\mathcal{R}_{j+1/2}^+$ for $q_{j+\frac{1}{2}}^+$ is given by

$$q_{j+\frac{1}{2}}^+ = e^{ik\Delta x} q_{j-\frac{1}{2}}^+ = e^{ik\Delta x} \mathcal{R}_{j-1/2}^+ \bar{q}_j = \mathcal{R}_{j+1/2}^+ \bar{q}_j. \quad (4.7d)$$

4.2.2 Fluid Velocity

To calculate the velocity perturbation $\mu_{j+1/2}$ a second-order FEM is used. The FEM commences with the weak formulation of (4.4c), obtained by multiplying (4.4c) by a test function v and integrating over the spatial domain Ω

$$\int_{\Omega} Gv \, dx = UH \int_{\Omega} v \, dx + U \int_{\Omega} \eta v \, dx + H \int_{\Omega} \mu v \, dx + \frac{H^3}{3} \int_{\Omega} \frac{\partial \mu}{\partial x} \frac{\partial v}{\partial x} \, dx.$$

The FEM then proceeds for (4.4c) as in Chapter 3 for the non-linear Serre equation using the basis functions defined in Appendix B. Therefore, G has the basis functions $\psi_{j-1/2}^+$ and $\psi_{j+1/2}^-$ (B.1), which means the approximation to G is linear inside a cell with discontinuous jumps at the cell edges. For v and μ the basis functions $\phi_{j-1/2}$, ϕ_j and $\phi_{j+1/2}$ (B.2) are used thus, v and the approximation to μ are quadratic polynomials inside a cell and are continuous across the cell edges.

Given the detailed description of the method in Chapter 3, the full derivation is omitted. The elementwise matrix \mathbf{A}_j and vector \mathbf{g}_j for the finite element approximation to (4.4) are

$$\begin{aligned} \mathbf{A}_j &= H \frac{\Delta x}{30} \begin{bmatrix} 4 & 2 & -1 \\ 2 & 16 & 2 \\ -1 & 2 & 4 \end{bmatrix} + \frac{H^3}{3} \frac{1}{3\Delta x} \begin{bmatrix} 7 & -8 & 1 \\ -8 & 16 & -8 \\ 1 & -8 & 7 \end{bmatrix}, \\ \mathbf{g}_j &= \frac{\Delta x}{6} \left(\begin{bmatrix} G_{j-1/2}^+ \\ 2G_{j-1/2}^+ + 2G_{j+1/2}^- \\ G_{j+1/2}^- \end{bmatrix} - UH \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix} - U \begin{bmatrix} \eta_{j-1/2}^+ \\ 2\eta_{j-1/2}^+ + 2\eta_{j+1/2}^- \\ \eta_{j+1/2}^- \end{bmatrix} \right). \end{aligned}$$

To calculate the intercell flux requires μ at $x_{j+1/2}$. From the elementwise matrices and vectors for the j^{th} and $(j+1)^{th}$ cells the equation that relates all the

quantities at $x_{j+1/2}$ is

$$\begin{aligned} \frac{\Delta x}{6} & \left(G_{j+1/2}^- + G_{j+1/2}^+ \right) \\ & = 2UH \frac{\Delta x}{6} + U \frac{\Delta x}{6} \left(\eta_{j+1/2}^- + \eta_{j+1/2}^+ \right) \\ & + \left(H \frac{\Delta x}{30} \left[-\mu_{j-1/2} + 2\mu_j + 8\mu_{j+1/2} + 2\mu_{j+1} - \mu_{j+3/2} \right] \right. \\ & \quad \left. + \frac{H^3}{3} \frac{1}{3\Delta x} \left[\mu_{j-1/2} - 8\mu_j + 14\mu_{j+1/2} - 8\mu_{j+1} + \mu_{j+3/2} \right] \right). \end{aligned}$$

Using (4.7) and (4.6),

$$\begin{aligned} \frac{\Delta x}{6} & \left(\mathcal{R}_{j+1/2}^- + \mathcal{R}_{j+1/2}^+ \right) \bar{G}_j \\ & = 2UH \frac{\Delta x}{6} + U \frac{\Delta x}{6} \left(\mathcal{R}_{j+1/2}^- + \mathcal{R}_{j+1/2}^+ \right) \bar{\eta}_j \\ & + \left(H \frac{\Delta x}{30} \left[-e^{-ik\Delta x} + 2e^{-ik\frac{\Delta x}{2}} + 8 + 2e^{ik\frac{\Delta x}{2}} - e^{ik\Delta x} \right] \right. \\ & \quad \left. + \frac{H^3}{3} \frac{1}{3\Delta x} \left[e^{-ik\Delta x} - 8e^{-ik\frac{\Delta x}{2}} + 14 - 8e^{ik\frac{\Delta x}{2}} + e^{ik\Delta x} \right] \right) \mu_{j+1/2}. \end{aligned}$$

Rearranging this equation it is obtained that

$$\mu_{j+1/2} = \mathcal{G}^\eta \bar{\eta}_j + \mathcal{G}^G \bar{G}_j + \mathcal{G}^c \quad (4.8)$$

where

$$\mathcal{G}^\eta = -U\mathcal{G}^G,$$

$$\mathcal{G}^G = \frac{\Delta x}{6\mathcal{G}_D} \left(\mathcal{R}_{j+1/2}^- + \mathcal{R}_{j+1/2}^+ \right),$$

$$\mathcal{G}^c = -2UH \frac{\Delta x}{6\mathcal{G}_D}$$

and the common divisor \mathcal{G}_D is

$$\begin{aligned} \mathcal{G}_D & = H \frac{\Delta x}{30} \left(4 \cos \left(\frac{k\Delta x}{2} \right) - 2 \cos(k\Delta x) + 8 \right) \\ & \quad + \frac{H^3}{3} \frac{1}{3\Delta x} \left(-16 \cos \left(\frac{k\Delta x}{2} \right) + 2 \cos(k\Delta x) + 14 \right). \end{aligned}$$

Hence, the factors \mathcal{G}^η , \mathcal{G}^G , \mathcal{G}^c that make up the operator \mathcal{G} do not depend on n or j as desired.

4.2.3 Flux Across the Cell Interfaces

The average intercell flux $F_{j+1/2}$ is approximated using (3.13). The linearised Serre equations possess the wave speed bounds (2.10), thus

$$a_{j+1/2}^- = \min \left\{ 0, U - \sqrt{gH} \right\}, \quad a_{j+1/2}^+ = \max \left\{ 0, U + \sqrt{gH} \right\}. \quad (4.9)$$

The method has three different approximations to $F_{j+1/2}$ depending on the Froude number $Fr = U/\sqrt{gH}$; (i) supercritical flow to the left where $Fr < -1$, (ii) critical and subcritical flow in both directions where $-1 \leq Fr \leq 1$ and (iii) supercritical flow to the right where $Fr > 1$. The flux operators are derived for each of these cases separately.

Left Supercritical Flow $Fr < -1$:

For left supercritical flow; $Fr < -1$ and thus $U + \sqrt{gH} < 0$, since (4.9) it must be that $a_{j+1/2}^- = U - \sqrt{gH}$ and $a_{j+1/2}^+ = 0$. For these values the flux approximation reduces to

$$F_{j+\frac{1}{2}} = f \left(q_{j+\frac{1}{2}}^+ \right) \quad (4.10)$$

for a generic quantity q .

Substituting the flux function from the continuity equation (4.4a) into the flux approximation produces

$$F_{j+\frac{1}{2}}^\eta = H\mu_{j+1/2} + U\eta_{j+1/2}^+$$

since μ is continuous $\mu_{j+1/2} = \mu_{j+1/2}^+ = \mu_{j+1/2}^-$.

Using the FEM for $\mu_{j+1/2}$ (4.8) and the reconstruction (4.7) it is obtained that

$$\begin{aligned} F_{j+\frac{1}{2}}^\eta &= H \left(\mathcal{G}^\eta \bar{\eta}_j + \mathcal{G}^G \bar{G}_j + \mathcal{G}^c \right) + U\eta_{j+1/2}^+ \\ &= \left(H\mathcal{G}^\eta + U\mathcal{R}_{j+1/2}^+ \right) \bar{\eta}_j + H\mathcal{G}^G \bar{G}_j + H\mathcal{G}^c. \end{aligned}$$

This can be written as factors for $\bar{\eta}_j$ and \bar{G}_j as follows

$$F_{j+\frac{1}{2}}^\eta = \mathcal{F}_{j+\frac{1}{2}}^{\eta,\eta} \bar{\eta}_j + \mathcal{F}_{j+\frac{1}{2}}^{\eta,G} \bar{G}_j + \mathcal{F}_{j+\frac{1}{2}}^{\eta,c}$$

where

$$\begin{aligned} \mathcal{F}_{j+\frac{1}{2}}^{\eta,\eta} &= H\mathcal{G}^\eta + U\mathcal{R}_{j+1/2}^+, \\ \mathcal{F}_{j+\frac{1}{2}}^{\eta,G} &= H\mathcal{G}^G, \\ \mathcal{F}_{j+\frac{1}{2}}^{\eta,c} &= H\mathcal{G}^c. \end{aligned}$$

Substituting the flux function for the G equation (4.4b) into the flux approximation (4.10) gives rise to

$$F_{j+\frac{1}{2}}^G = UG_{j+1/2}^+ + UH\mu_{j+1/2} + gH\eta_{j+1/2}^+.$$

Using the FEM (4.8) to calculate $\mu_{j+1/2}$ and the interface reconstruction (4.7) produces

$$F_{j+\frac{1}{2}}^G = UG_{j+1/2}^+ + UH(\mathcal{G}^\eta \bar{\eta}_j + \mathcal{G}^G \bar{G}_j + \mathcal{G}^c) + gH\eta_{j+1/2}^+$$

which can be rewritten as

$$F_{j+\frac{1}{2}}^G = \mathcal{F}_{j+\frac{1}{2}}^{G,\eta} \bar{\eta}_j + \mathcal{F}_{j+\frac{1}{2}}^{G,G} \bar{G}_j + \mathcal{F}_{j+\frac{1}{2}}^{G,c}$$

where

$$\begin{aligned}\mathcal{F}_{j+\frac{1}{2}}^{G,\eta} &= UH\mathcal{G}^\eta + gH\mathcal{R}_{j+1/2}^+, \\ \mathcal{F}_{j+\frac{1}{2}}^{G,G} &= U\mathcal{R}_{j+1/2}^+ + UH\mathcal{G}^G, \\ \mathcal{F}_{j+\frac{1}{2}}^{G,c} &= UH\mathcal{G}^c.\end{aligned}$$

Subcritical Flow $-1 \leq Fr \leq 1$:

When the flow is subcritical $-1 \leq Fr \leq 1$, implying that $a_{j+1/2}^- = U - \sqrt{gH}$ and $a_{j+1/2}^+ = U + \sqrt{gH}$. Therefore, the flux approximation (3.13) becomes

$$\begin{aligned}F_{j+\frac{1}{2}} &= \frac{U}{2\sqrt{gH}} \left[f(q_{j+\frac{1}{2}}^-) - f(q_{j+\frac{1}{2}}^+) \right] + \frac{1}{2} \left[f(q_{j+\frac{1}{2}}^-) + f(q_{j+\frac{1}{2}}^+) \right] \\ &\quad + \frac{U^2 - gH}{2\sqrt{gH}} \left[q_{j+\frac{1}{2}}^+ - q_{j+\frac{1}{2}}^- \right].\end{aligned}\tag{4.11}$$

Substituting in the flux function for η given by (4.4a) we get

$$\begin{aligned}F_{j+\frac{1}{2}}^\eta &= \frac{U}{2\sqrt{gH}} \left(H\mu_{j+1/2} + U\eta_{j+\frac{1}{2}}^- - H\mu_{j+1/2} - U\eta_{j+\frac{1}{2}}^+ \right) \\ &\quad + \frac{1}{2} \left(H\mu_{j+1/2} + U\eta_{j+\frac{1}{2}}^- + H\mu_{j+1/2} + U\eta_{j+\frac{1}{2}}^+ \right) \\ &\quad + \frac{U^2 - gH}{2\sqrt{gH}} \left(\eta_{j+\frac{1}{2}}^+ - \eta_{j+\frac{1}{2}}^- \right).\end{aligned}$$

Using the reconstruction factors (4.7) and (4.8) and rearranging produces

$$F_{j+\frac{1}{2}}^\eta = \mathcal{F}_{j+\frac{1}{2}}^{\eta,\eta} \bar{\eta}_j + \mathcal{F}_{j+\frac{1}{2}}^{\eta,G} \bar{G}_j + \mathcal{F}_{j+\frac{1}{2}}^{\eta,c}$$

where

$$\begin{aligned}\mathcal{F}_{j+\frac{1}{2}}^{\eta,\eta} &= H\mathcal{G}^\eta + \frac{U}{2} \left[\mathcal{R}_{j+1/2}^- + \mathcal{R}_{j+1/2}^+ \right] - \frac{\sqrt{gH}}{2} \left[\mathcal{R}_{j+1/2}^+ - \mathcal{R}_{j+1/2}^- \right] \\ \mathcal{F}_{j+\frac{1}{2}}^{\eta,G} &= H\mathcal{G}^G \\ \mathcal{F}_{j+\frac{1}{2}}^{\eta,c} &= H\mathcal{G}^c\end{aligned}$$

For the flux function of G (4.4b) the flux approximation (4.11) becomes

$$\begin{aligned}F_{j+\frac{1}{2}}^G &= \frac{U}{2\sqrt{gH}} \left(UG_{j+\frac{1}{2}}^- + UH\mu_{j+1/2} + gH\eta_{j+\frac{1}{2}}^- - UG_{j+\frac{1}{2}}^+ - UH\mu_{j+1/2} - gH\eta_{j+\frac{1}{2}}^+ \right) \\ &\quad + \frac{1}{2} \left(UG_{j+\frac{1}{2}}^- + UH\mu_{j+1/2} + gH\eta_{j+\frac{1}{2}}^- + UG_{j+\frac{1}{2}}^+ + UH\mu_{j+1/2} + gH\eta_{j+\frac{1}{2}}^+ \right) \\ &\quad + \frac{U^2 - gH}{2\sqrt{gH}} \left(G_{j+\frac{1}{2}}^+ - G_{j+\frac{1}{2}}^- \right).\end{aligned}$$

Using the reconstruction factors (4.7) and (4.8) it follows that

$$F_{j+\frac{1}{2}}^G = \mathcal{F}_{j+\frac{1}{2}}^{G,\eta} \bar{\eta}_j + \mathcal{F}_{j+\frac{1}{2}}^{G,G} \bar{G}_j + \mathcal{F}_{j+\frac{1}{2}}^{G,c}$$

where

$$\begin{aligned}\mathcal{F}_{j+\frac{1}{2}}^{G,\eta} &= \frac{U\sqrt{gH}}{2} \left[\mathcal{R}_{j+1/2}^- - \mathcal{R}_{j+1/2}^+ \right] + UH\mathcal{G}^\eta + \frac{gH}{2} \left[\mathcal{R}_{j+1/2}^- + \mathcal{R}_{j+1/2}^+ \right], \\ \mathcal{F}_{j+\frac{1}{2}}^{G,G} &= UH\mathcal{G}^G + \frac{U}{2} \left[\mathcal{R}_{j+1/2}^- + \mathcal{R}_{j+1/2}^+ \right] - \frac{\sqrt{gH}}{2} \left[\mathcal{R}_{j+1/2}^+ - \mathcal{R}_{j+1/2}^- \right], \\ \mathcal{F}_{j+\frac{1}{2}}^{G,c} &= UH\mathcal{G}^c.\end{aligned}$$

Right Supercritical Flow $Fr > 1$:

For a rightward supercritical flow $Fr > 1$, implying that $a_{j+1/2}^- = 0$ and $a_{j+1/2}^+ = U + \sqrt{gH}$. Thus the situation for rightward supercritical flow is very similar to the left supercritical case, except $\mathcal{R}_{j+1/2}^+$ is replaced with $\mathcal{R}_{j+1/2}^-$ in the flux approximation for a general quantity (3.13) which reduces to

$$F_{j+\frac{1}{2}} = f \left(q_{j+\frac{1}{2}}^- \right).$$

Substituting in the flux function into (4.4a) and (4.4b) produces

$$F_{j+\frac{1}{2}}^\eta = \mathcal{F}_{j+\frac{1}{2}}^{\eta,\eta} \bar{\eta}_j + \mathcal{F}_{j+\frac{1}{2}}^{\eta,G} \bar{G}_j + \mathcal{F}_{j+\frac{1}{2}}^{\eta,c}$$

where

$$\begin{aligned}\mathcal{F}_{j+\frac{1}{2}}^{\eta,\eta} &= H\mathcal{G}^\eta + U\mathcal{R}_{j+1/2}^-, \\ \mathcal{F}_{j+\frac{1}{2}}^{\eta,G} &= H\mathcal{G}^G, \\ \mathcal{F}_{j+\frac{1}{2}}^{\eta,c} &= H\mathcal{G}^c\end{aligned}$$

and

$$F_{j+\frac{1}{2}}^G = \mathcal{F}_{j+\frac{1}{2}}^{G,\eta} \bar{\eta}_j + \mathcal{F}_{j+\frac{1}{2}}^{G,G} \bar{G}_j + \mathcal{F}_{j+\frac{1}{2}}^{G,c}$$

where

$$\begin{aligned}\mathcal{F}_{j+\frac{1}{2}}^{G,\eta} &= UH\mathcal{G}^\eta + gH\mathcal{R}_{j+1/2}^-, \\ \mathcal{F}_{j+\frac{1}{2}}^{G,G} &= U\mathcal{R}_{j+1/2}^+ + UH\mathcal{G}^G, \\ \mathcal{F}_{j+\frac{1}{2}}^{G,c} &= UH\mathcal{G}^c\end{aligned}$$

respectively.

4.2.4 Update Cell Averages

The operators for the flux functions for supercritical, critical and subcritical flow have been obtained. Substituting the appropriate flux approximation into the forward Euler step, (3.21) it follows that

$$\begin{aligned}\bar{\eta}_j^{n+1} &= \bar{\eta}_j^n - \frac{\Delta t}{\Delta x} \left(\left[\mathcal{F}_{j+\frac{1}{2}}^{\eta,\eta} \bar{\eta}_j + \mathcal{F}_{j+\frac{1}{2}}^{\eta,G} \bar{G}_j + \mathcal{F}_{j+\frac{1}{2}}^{\eta,c} \right] - \left[\mathcal{F}_{j-\frac{1}{2}}^{\eta,\eta} \bar{\eta}_j + \mathcal{F}_{j-\frac{1}{2}}^{\eta,G} \bar{G}_j + \mathcal{F}_{j-\frac{1}{2}}^{\eta,c} \right] \right), \\ \bar{G}_j^{n+1} &= \bar{G}_j^n - \frac{\Delta t}{\Delta x} \left(\left[\mathcal{F}_{j+\frac{1}{2}}^{G,\eta} \bar{\eta}_j + \mathcal{F}_{j+\frac{1}{2}}^{G,G} \bar{G}_j + \mathcal{F}_{j+\frac{1}{2}}^{G,c} \right] - \left[\mathcal{F}_{j-\frac{1}{2}}^{G,\eta} \bar{\eta}_j + \mathcal{F}_{j-\frac{1}{2}}^{G,G} \bar{G}_j + \mathcal{F}_{j-\frac{1}{2}}^{G,c} \right] \right).\end{aligned}$$

Since $\mathcal{F}_{j-\frac{1}{2}}^{\eta,\eta} = e^{-ik\Delta x} \mathcal{F}_{j+\frac{1}{2}}^{\eta,\eta}$, $\mathcal{F}_{j-\frac{1}{2}}^{\eta,G} = e^{-ik\Delta x} \mathcal{F}_{j+\frac{1}{2}}^{\eta,G}$, $\mathcal{F}_{j-\frac{1}{2}}^{G,\eta} = e^{-ik\Delta x} \mathcal{F}_{j+\frac{1}{2}}^{G,\eta}$ and $\mathcal{F}_{j-\frac{1}{2}}^{G,G} = e^{-ik\Delta x} \mathcal{F}_{j+\frac{1}{2}}^{G,G}$ it must be that

$$\begin{aligned}\bar{\eta}_j^{n+1} &= \bar{\eta}_j^n - \frac{\Delta t}{\Delta x} \left([1 - e^{-ik\Delta x}] \left[\mathcal{F}_{j+\frac{1}{2}}^{\eta,\eta} \bar{\eta}_j + \mathcal{F}_{j+\frac{1}{2}}^{\eta,G} \bar{G}_j \right] \right), \\ \bar{G}_j^{n+1} &= \bar{G}_j^n - \frac{\Delta t}{\Delta x} \left([1 - e^{-ik\Delta x}] \left[\mathcal{F}_{j+\frac{1}{2}}^{G,\eta} \bar{\eta}_j + \mathcal{F}_{j+\frac{1}{2}}^{G,G} \bar{G}_j \right] \right).\end{aligned}$$

This can be written in matrix form using the identity matrix \mathbf{I} as

$$\begin{aligned}\begin{bmatrix} \bar{\eta} \\ \bar{G} \end{bmatrix}_j^{n+1} &= \begin{bmatrix} \bar{\eta} \\ \bar{G} \end{bmatrix}_j^n - (1 - e^{-ik\Delta x}) \frac{\Delta t}{\Delta x} \begin{bmatrix} \mathcal{F}^{\eta,\eta} & \mathcal{F}^{\eta,G} \\ \mathcal{F}^{G,\eta} & \mathcal{F}^{G,G} \end{bmatrix} \begin{bmatrix} \bar{\eta} \\ \bar{G} \end{bmatrix}_j^n \\ &= (\mathbf{I} - \Delta t \mathbf{F}) \begin{bmatrix} \bar{\eta} \\ \bar{G} \end{bmatrix}_j^n\end{aligned}\tag{4.12}$$

for a single Euler step which is first-order in time.

4.2.5 Second-Order SSP Runge-Kutta Method

To achieve second-order accurate time stepping, the second-order SSP Runge-Kutta scheme (3.22) is used. This scheme uses the following convex combination of the Euler steps (4.12)

$$\left[\frac{\bar{\eta}}{G} \right]_j^{(1)} = (\mathbf{I} - \Delta t \mathbf{F}) \left[\frac{\bar{\eta}}{G} \right]_j^n, \quad (4.13a)$$

$$\left[\frac{\bar{\eta}}{G} \right]_j^{(2)} = (\mathbf{I} - \Delta t \mathbf{F}) \left[\frac{\bar{\eta}}{G} \right]_j^{(1)}, \quad (4.13b)$$

$$\left[\frac{\bar{\eta}}{G} \right]_j^{n+1} = \frac{1}{2} \left(\left[\frac{\bar{\eta}}{G} \right]_j^n + \left[\frac{\bar{\eta}}{G} \right]_j^{(2)} \right). \quad (4.13c)$$

Substituting (4.13a) and (4.13b) into (4.13c) it can be written in terms of the flux matrix \mathbf{F} and the cell averages at t^n as

$$\left[\frac{\bar{\eta}}{G} \right]_j^{n+1} = \frac{1}{2} \left(\left[\frac{\bar{\eta}}{G} \right]_j^n + (\mathbf{I} - \Delta t \mathbf{F})^2 \left[\frac{\bar{\eta}}{G} \right]_j^n \right).$$

Expanding $(\mathbf{I} - \Delta t \mathbf{F})^2$ produces

$$\begin{aligned} \left[\frac{\bar{\eta}}{G} \right]_j^{n+1} &= \left(\mathbf{I} - \Delta t \mathbf{F} + \frac{1}{2} \Delta t^2 \mathbf{F}^2 \right) \left[\frac{\bar{\eta}}{G} \right]_j^n \\ &= \mathbf{E} \left[\frac{\bar{\eta}}{G} \right]_j^n \end{aligned} \quad (4.14)$$

which is in the desired form (4.1).

This is the evolution matrix \mathbf{E} for FEVM₂. The matrix \mathbf{E} is dependent on the flux matrix \mathbf{F} and therefore will depend on the Froude number. The Froude number is constant over time in this analysis therefore, supercritical, subcritical and critical flow can be investigated individually.

The convergence and dispersion analysis proceed by studying the properties of the evolution matrix \mathbf{E} for FEVM₂. As a comparison the results for the finite difference volume methods FDVM₁, FDVM₂ and FDVM₃ described by Zoppou et al. [15] and the finite difference methods \mathcal{D} and \mathcal{W} described by Pitt et al. [18] are also provided.

The evolution matrices for FDVM₁, FDVM₂, FDVM₃ can be derived following the derivation of the evolution matrix of FEVM₂ using the expressions for its constituent operators provided in Appendix C. For \mathcal{D} and \mathcal{W} the evolution matrices are (C.10) and (C.11) respectively. The results of the convergence analysis using these evolution matrices will now be presented.

4.3 Convergence Analysis

The Lax-equivalence theorem is employed to demonstrate the convergence of the numerical methods by establishing their stability and consistency. Stability is established for the numerical methods using a Von Neumann stability analysis. While consistency is established for the numerical methods by representing the numerical solution as Fourier modes (4.5) and demonstrating the consistency of the numerical method for these Fourier mode solutions. Since the Fourier modes form a basis of the solution space of the linearised Serre equations, this demonstrates consistency for all numerical solutions. The presented stability and consistency analyses imply convergence of the numerical method under the L_2 norm.

4.3.1 Stability

For a numerical method to be stable the errors from previous time steps must not be amplified over the current time step. To ensure that this is the case it is necessary that

$$\rho(\mathbf{E}) \leq 1 \quad (4.15)$$

where $\rho(\mathbf{E})$ is the spectral radius of \mathbf{E} . Since \mathbf{E} was derived for the methods using Fourier modes, this condition implies Von Neumann stability.

The spectral radius $\rho(\mathbf{E})$ was calculated numerically for various values of Δx , Δt , k , H and U to investigate the assertion that (4.15). The results of this investigation are summarised in Figure 4.1 which is a plot of $\rho(\mathbf{E})$ against $\Delta x/\lambda$ for representative values of k , H and U ; where $\lambda = 2\pi/k$ is the wavelength. The standard gravitational acceleration $g = 9.81m/s^2$ was used and the time step $\Delta t = 0.5/(U + \sqrt{gH})\Delta x$ was chosen to satisfy the CFL condition (3.23). This was the common choice of Δt in the numerical experiments discussed in Chapters 5 and 6.

The behaviour of $\rho(\mathbf{E})$ for $H = 1m$, $k = \pi/10m^{-1}$ with $U = 0m/s$ and $U = 1m/s$ is shown in Figure 4.1 and is representative of the behaviour for all

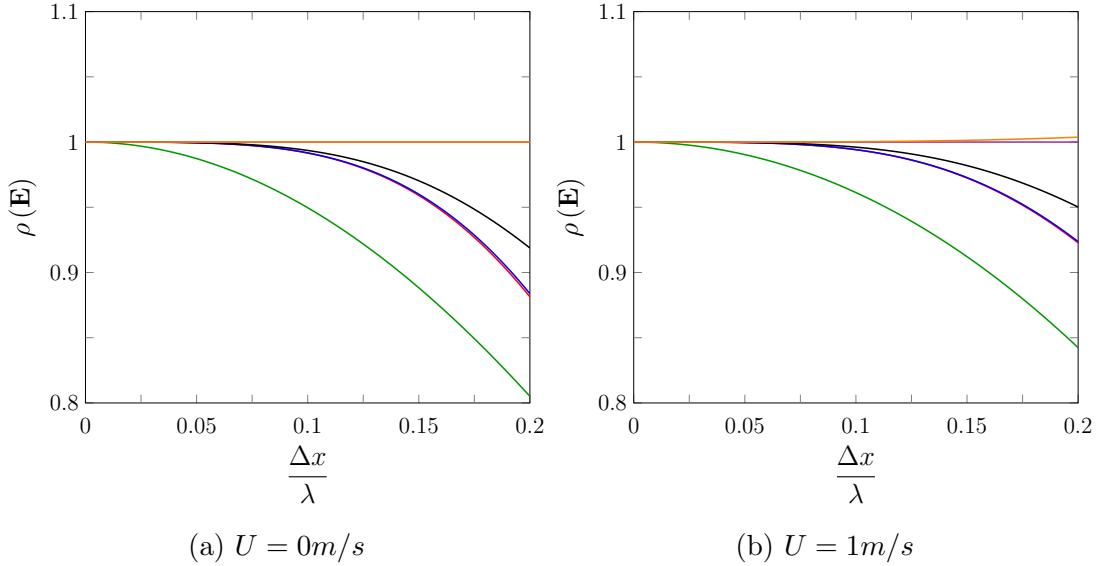


Figure 4.1: Spectral radius of \mathbf{E} against $\Delta x/\lambda$ for FDVM₁ (green line), FDVM₂ (red line), FEVM₂ (blue line), FDVM₃ (black line), \mathcal{D} (purple line) and \mathcal{W} (orange line). With $H = 1m$ and $k = \pi/10m^{-1}$.

other values of H , k and U . For these k and H values the shallowness parameter $\sigma = 1/20$ and so the Serre equations are applicable [24].

In Figure 4.1 it can be seen that all methods have $\rho(\mathbf{E}) \leq 1$ for $U = 0\text{m/s}$ and are therefore stable. The two finite difference methods overlap and have $\rho(\mathbf{E}) = 1$ for all Δx values, while FDVM₂ and FEVM₂ also overlap with $\rho(\mathbf{E}) < 1$. However, when $U \neq 0\text{m/s}$ the method \mathcal{W} has $\rho(\mathbf{E}) > 1$ for all Δx values and is therefore unstable. All other methods have $\rho(\mathbf{E}) \leq 1$, retaining their stability when $U \neq 0\text{m/s}$.

Similar results were observed for a wide range of k , H and U values and Froude numbers. All methods except \mathcal{W} were found to be stable for any combination of these variables. While \mathcal{W} was only stable when $U = 0\text{m/s}$. This differs from the stability result for \mathcal{D} and \mathcal{W} reported by Pitt et al. [18] as that analysis assumed that $U = 0\text{m/s}$.

4.3.2 Consistency

For a numerical method to be consistent the error introduced by the method for a single time step must approach zero as the spatial and temporal resolution is increased. It is sufficient to demonstrate consistency for the eigenfunctions of the linearised Serre equations, which are the Fourier modes as the eigenfunctions form a basis of the solution space. Therefore, the consistency of the methods are

demonstrated using the evolution matrix \mathbf{E} .

The error introduced for a single time step from t^n to t^{n+1} , \mathcal{T}_j^n is

$$\mathcal{T}_j^n = \mathbf{E} \left[\frac{\bar{\eta}}{G} \right]_j^n - \left[\frac{\bar{\eta}}{G} \right]_j^{n+1}. \quad (4.16)$$

To ensure consistency it must be that $\lim_{\Delta x, \Delta t \rightarrow 0} \|\mathcal{T}_j^n\|_2 = 0$ for all n and j , where $\|\cdot\|_2$ is the L_2 vector norm. Note that the location x_j and time t^n will be fixed as Δx and Δt vary, so that the measurement of the error \mathcal{T}_j^n is consistent. Taking the L_2 norm of both sides of (4.16) and using (4.6) it follows that

$$\|\mathcal{T}_j^n\|_2 = \left\| \mathbf{E} \left[\frac{\bar{\eta}}{G} \right]_j^n - e^{i\omega^\pm \Delta t} \left[\frac{\bar{\eta}}{G} \right]_j^n \right\|_2.$$

Using the matrix norm induced by L_2 , the Frobenius norm $\|\cdot\|_F$ produces

$$\|\mathcal{T}_j^n\|_2 \leq \left\| \mathbf{E} - e^{i\omega^\pm \Delta t} \mathbf{I} \right\|_F \left\| \left[\frac{\bar{\eta}}{G} \right]_j^n \right\|_2.$$

Both $\bar{\eta}_j^n$ and \bar{G}_j^n approach their respective midpoint values η_j^n and G_j^n as Δx goes to zero due to the shrinking of the j^{th} cell. Additionally, $\bar{\eta}_j^n$ and \bar{G}_j^n will vary as Δx varies however, they remain finite. Therefore, if $\lim_{\Delta x, \Delta t \rightarrow 0} \left\| \mathbf{E} - e^{i\omega^\pm \Delta t} \mathbf{I} \right\|_F = 0$ then $\lim_{\Delta x, \Delta t \rightarrow 0} \|\mathcal{T}_j^n\|_2 = 0$ as desired.

The Taylor series of $\mathbf{E} - e^{i\omega^\pm \Delta t} \mathbf{I}$ was calculated for all the numerical methods for all flow scenarios; subcritical, critical and supercritical flows. Since the results are the same for ω^+ and ω^- only the results for ω^+ are reported. For FEVM₂ the lowest order Δx and Δt terms of the Taylor series of $\mathbf{E} - e^{i\omega^+ \Delta t} \mathbf{I}$ can be found in Table 4.1. From Table 4.1 it can be seen that the Taylor series of all the elements of $\mathbf{E} - e^{i\omega^+ \Delta t} \mathbf{I}$ have a factor of Δt . So that

$$\begin{aligned} \left\| \mathbf{E} - e^{i\omega^+ \Delta t} \mathbf{I} \right\|_F &= \left\| \Delta t (\mathbf{M}_0 + \mathcal{O}(\Delta t)) \right\|_F \\ &= |\Delta t| \left\| \mathbf{M}_0 + \mathcal{O}(\Delta t) \right\|_F \\ &\leq |\Delta t| (\|\mathbf{M}_0\|_F + \|\mathcal{O}(\Delta t)\|_F) \end{aligned}$$

where \mathbf{M}_0 is some matrix.

From Table 4.1 it is evident that \mathbf{M}_0 is independent of Δt and finite so that as $\Delta t \rightarrow 0$ then $|\Delta t| (\|\mathbf{M}_0\|_F + \|\mathcal{O}(\Delta t)\|_F) \rightarrow 0$ and therefore $\left\| \mathbf{E} - e^{i\omega^+ \Delta t} \mathbf{I} \right\|_F \rightarrow 0$. Hence, for FEVM₂ it must be that $\lim_{\Delta x, \Delta t \rightarrow 0} \|\mathcal{T}_j^n\|_2 = 0$ for all n and j

Element	Lowest Order Terms of $\mathbf{E} - e^{i\omega^+\Delta t}\mathbf{I}$ for FEVM ₂	
	Δx	Δt
$E_{0,0} - e^{i\omega^+\Delta t}$	$-\frac{i(54 + 45H^2k^2 + 10H^4k^4)}{120\beta^2}Uk^3\Delta t\Delta x^2$	$\frac{\sqrt{3gH\beta} + 3U}{\beta}ik\Delta t$
$E_{0,1}$	$\frac{\beta - 3}{\beta^2}\frac{ik^3}{40}\Delta t\Delta x^2$	$-\frac{3}{\beta}ik\Delta t$
$E_{1,0}$	$-\left(gH - \frac{15U^2}{\beta} + \frac{9U^2}{\beta}\right)\frac{k^3}{120}\Delta t\Delta x^2$	$\left(-gH + \frac{3U^2}{\beta}\right)ik\Delta t$
$E_{1,1} - e^{i\omega^+\Delta t}$	$\frac{126 + 75H^2k^2 + 10H^4k^4}{\beta^2}\frac{k^3}{120}iU\Delta t\Delta x^2$	$\frac{\sqrt{3gH\beta} - 3U}{\beta}ik\Delta t$

Table 4.1: Lowest order terms of the Taylor series for the elements of $\mathbf{E} - e^{i\omega^+\Delta t}\mathbf{I}$ for FEVM₂ for all values of Fr . Here $\beta = 3 + k^2H^2$.

thus, FEVM₂ is consistent for Fourier mode solutions implying consistency for all solutions as desired.

All methods were found to be consistent. The associated Taylor series Tables for all other methods can be found in Appendix C. More specifically, Tables C.9 and C.10 for FDVM₁, Table C.11 for FDVM₂, Tables C.12 and C.13 for FDVM₃, Table C.14 for \mathcal{D} and Table C.15 for \mathcal{W} .

4.4 Dispersion Analysis

To study the dispersion properties of the numerical method, the dispersion relation of the numerical method that relates the frequency $\tilde{\omega}^\pm$ of the numerical solution to the wavenumber k must be calculated. Making use of (4.6) in (4.14) it follows that

$$\mathbf{E} \left[\frac{\bar{\eta}}{G} \right]_j^n = e^{i\omega^\pm \Delta t} \left[\frac{\bar{\eta}}{G} \right]_j^n.$$

Therefore, the evolution matrix \mathbf{E} of an exact method has the eigenvalues $e^{i\omega^+ \Delta t}$ and $e^{i\omega^- \Delta t}$ where ω^\pm are the positive and negative branches of the dispersion relation of the linearised Serre equations (2.9). For approximate numerical methods the dispersion relation for $\tilde{\omega}^\pm$ can be calculated by taking the eigenvalues of its evolution matrix τ^\pm like so

$$\tilde{\omega}^\pm = \frac{1}{i\Delta t} \log [\tau^\pm].$$

By comparing $\tilde{\omega}^\pm$ with the analytic ω^\pm given by the linearised Serre equations (2.9) the error in the dispersion relation for the numerical method can be determined. The real part of the frequency indicates the speed of a wave, while the imaginary part determines the change in amplitude. For the linearised Serre equations the imaginary part of ω^\pm is zero and so the amplitude of waves are constant in time. Only the results for the positive branch of the dispersion relation comparing $\tilde{\omega}^+$ and ω^+ is presented as the behaviour of the negative branch is similar.

The relative error in the dispersion relation was plotted against $\Delta x/\lambda$ for representative values of H , U and k . Where $\lambda = 2\pi/k$ is the wavelength of the wave with wave number k . The gravitational acceleration $g = 9.81 m/s^2$ was used and the time step value $\Delta t = 0.5 / (U + \sqrt{gH}) \Delta x$ was chosen to satisfy the CFL condition (3.23).

Figures 4.2 and 4.3 presents the dispersion error for $kH = \pi/10$ where $\sigma = kH/2\pi = 1/20$ and so the water is shallow and thus the Serre equations are appropriate. The real and imaginary errors are displayed separately to isolate the errors in the speed and amplitude of the wave for the numerical method. The total error is also reported as a measure of the overall error in the dispersion relation of the numerical method.

From Figures 4.2 and 4.3 it is evident that all methods approximate the dispersion relation of the Serre equations well with the approximation improving as $\Delta x \rightarrow 0$, as expected.

For the real part of the dispersion error all the FEVM and the FDVM outperform the two finite difference methods. Therefore, the FEVM and the FDVM will better approximate the speed of waves in the linearised Serre equations than \mathcal{D} and \mathcal{W} . However, for the amplitude of waves the roles are reversed with the two finite difference methods either scaling the waves very little or not at all. When taking both effects into account, FDVM₁ has the largest dispersion error followed by \mathcal{W} , \mathcal{D} , FEVM₂, FDVM₂ and finally FDVM₃ has the lowest dispersion error. Hence, the size of the total dispersion error is mainly determined by the order of accuracy of the numerical scheme. Overall the finite volume based methods perform better than the finite difference methods of the same order.

Figures 4.2 and 4.3 demonstrate that FDVM₂ is superior to FEVM₂ for the complete dispersion error and the real and imaginary parts separately as well. Therefore, FDVM₂ should more accurately model the speed and amplitude of waves.

Similar results were observed across a wide array of k , H and U values. However, as kH increases the distinction between FDVM₂ and FEVM₂ becomes less pronounced, as can be observed in Figure 4.4.

The real part of the dispersion error for $kH = 2.5$ plotted in Figure 4.4 was also produced by Filippini et al. [40] for their methods. The results for the FDVM and the FEVM compare favourably with the methods described and analysed by Filippini et al. [40]. Furthermore, their work is extended here by allowing for non-zero values of U , combining the spatial and temporal approximations and examining the imaginary and total error in the dispersion relation. This work also extended the dispersion analysis presented by Zoppou et al. [15] for FDVM₁, FDVM₂ and FDVM₃ by including non-zero values of U .

Figure 4.5 demonstrates that the results of the real part of the dispersion error is slightly different when U is not zero. For example the non-zero value of U significantly changes the real part of the dispersion error for FDVM₁ when

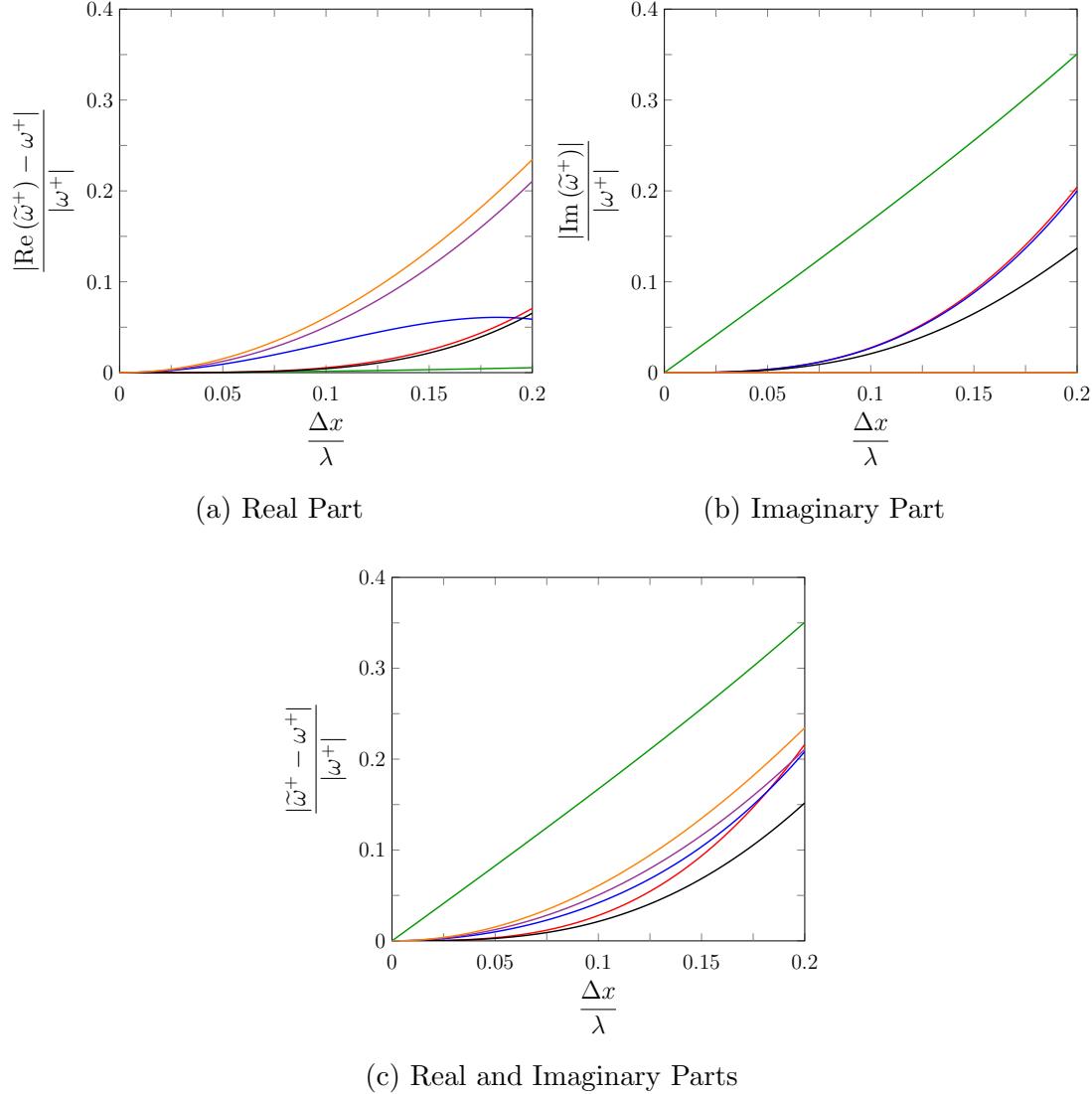


Figure 4.2: Relative dispersion error against $\Delta x/\lambda$ when $H = 1m$, $k = \pi/10m^{-1}$ and $U = 0m/s$ for FDVM₁ (—), FDVM₂ (—), FEVM₂ (—), FDVM₃ (—), \mathcal{D} (—) and \mathcal{W} (—).

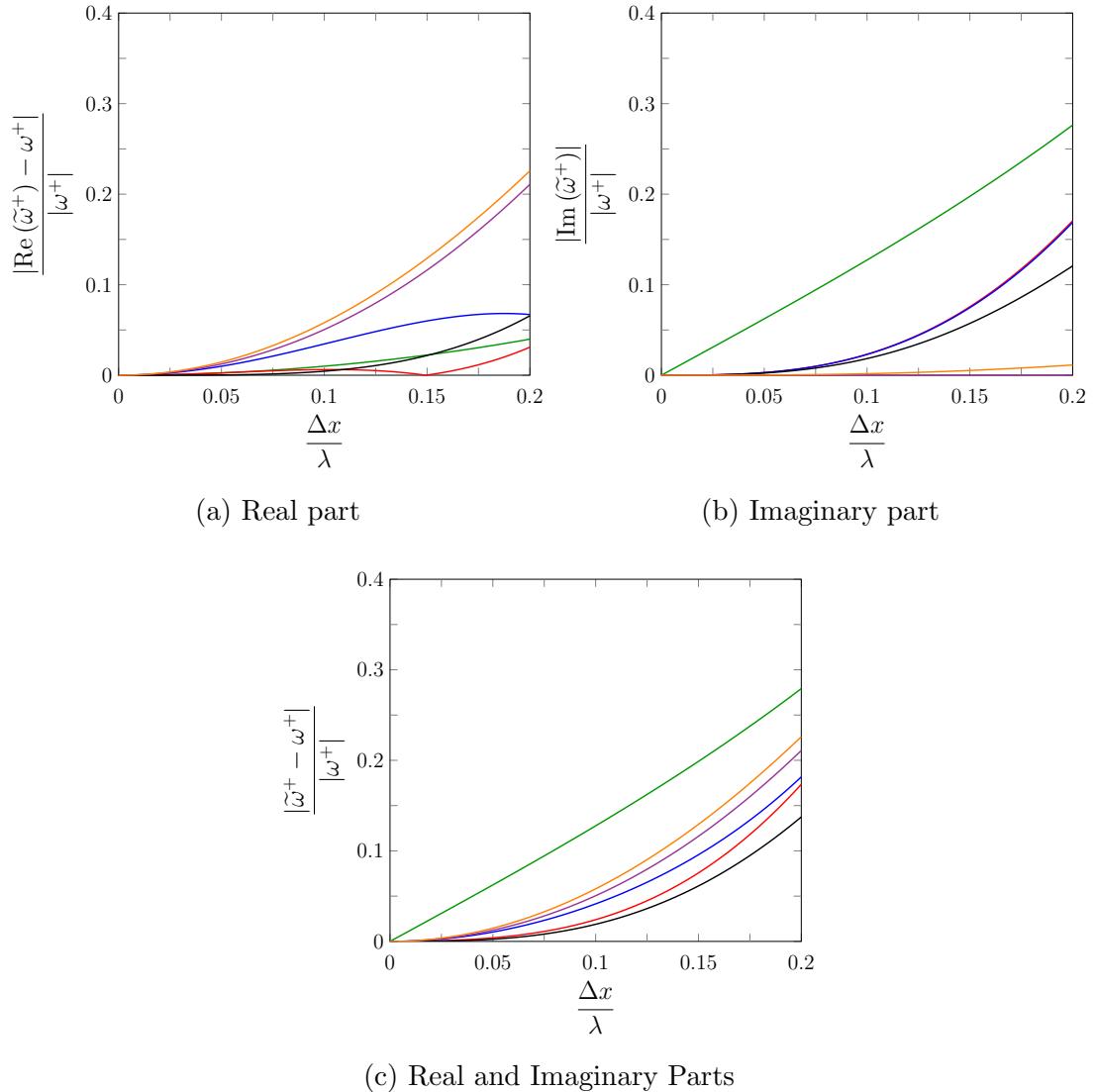


Figure 4.3: Relative dispersion error against $\Delta x/\lambda$ when $H = 1m$, $k = \pi/10m^{-1}$ and $U = 1m/s$ for FDVM₁ (—), FDVM₂ (—), FEVM₂ (—), FDVM₃ (—), \mathcal{D} (—) and \mathcal{W} (—).

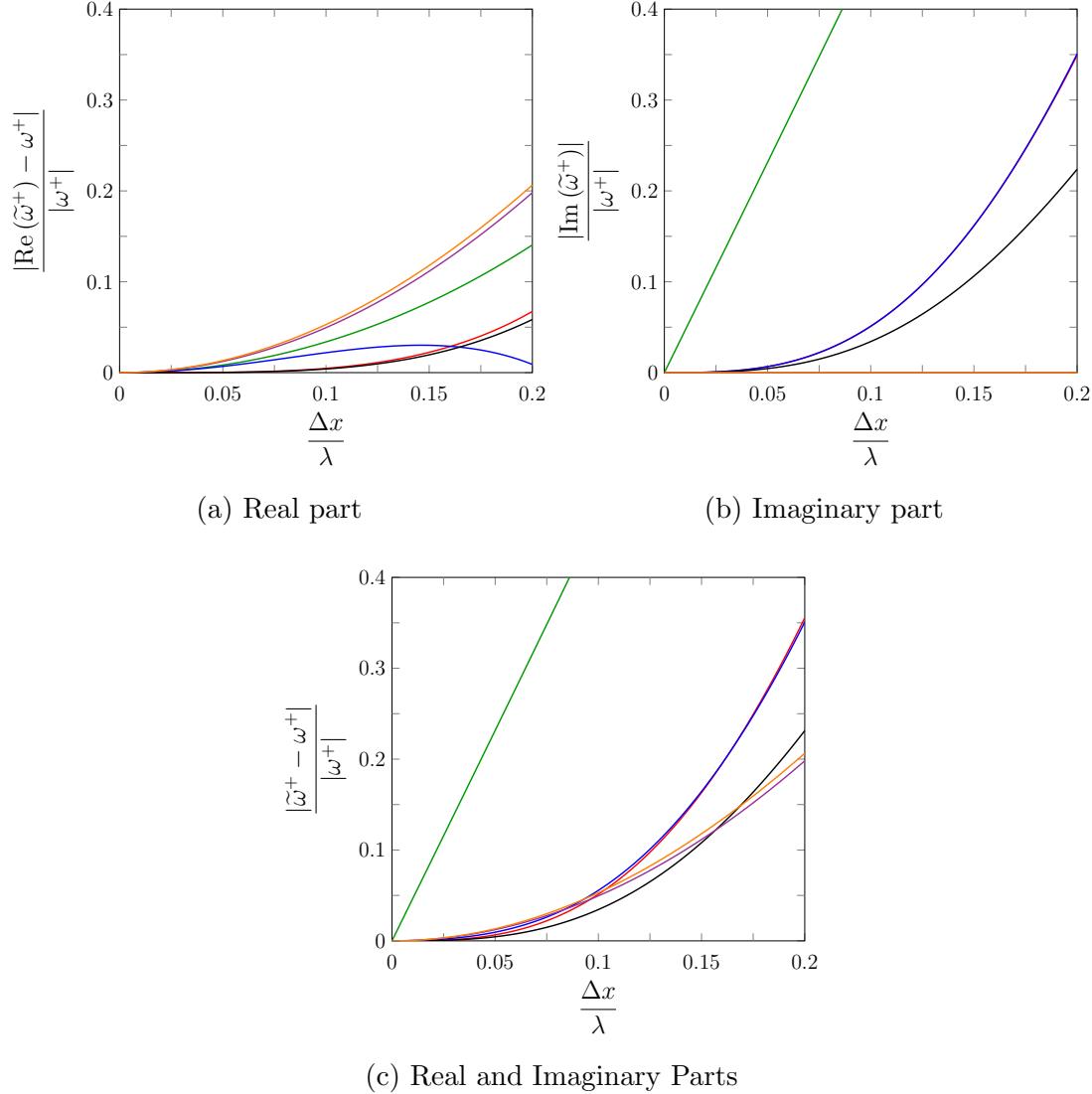


Figure 4.4: Relative dispersion error against $\Delta x/\lambda$ when $H = 1m$, $k = 2.5m^{-1}$ and $U = 0m/s$ for FDVM₁ (green), FDVM₂ (red), FEVM₂ (blue), FDVM₃ (black), \mathcal{D} (purple) and \mathcal{W} (orange).

$kH = 2.5$. Therefore, for some methods allowing for non-zero values of U can have a significant impact on the conclusions drawn from the dispersion analysis. Furthermore, taking the imaginary part of the dispersion error into account is important as ω^\pm determines not only the speed of waves but also their amplitude. For instance the FDVM₁ performs very well for the real part of the dispersion error and poorly for the imaginary part, and so false conclusions about the accuracy of the method could be drawn from only considering the real part of the dispersion error.

The Taylor series expansion of $\tilde{\omega}^\pm$ was also derived for all the numerical methods. The lowest order terms of the Taylor series for $\tilde{\omega}^+ - \omega^+$ are compiled in Table 4.2 when $-1 \leq Fr \leq 1$ for the FDVM and FEVM. In Table 4.2 it is clear that these schemes estimated ω^+ with the expected order of accuracy in both space and time. This was also the case for ω^- .

The lowest order terms of the Taylor series for $\tilde{\omega}^+ - \omega^+$ for both $Fr < -1$ and $Fr > 1$ are presented in Table 4.3. Only the errors that are different from those reported in Table 4.2 are presented, which was only the case for the spatial error of the first- and third-order numerical methods. From Tables 4.3 and 4.2 it is can be seen that the FDVM and the FEVM retain their order of accuracy when approximating ω^+ when the flow is supercritical, this was also the case for ω^- .

Finally, the lowest order terms of the Taylor series for $\tilde{\omega}^+ - \omega^+$ for the finite difference methods are displayed in Table 4.4. These methods do not change depending on the value of the physical quantities. The two finite difference methods retain their order of accuracy in space and time when approximating ω^+ .

Because all methods were demonstrated to have the expected order of accuracy in approximating ω^\pm for the linearised Serre equations this implies that for small Δx values the order of accuracy will be the primary driver of the dispersion error, as was observed.

In this chapter the convergence and dispersion properties of the numerical methods were studied using a linear analysis. The results of this analysis demonstrated the superiority of the high-order accurate FDVM and FEVM over the finite difference methods. The numerical methods will now be validated against analytic and forced solutions.

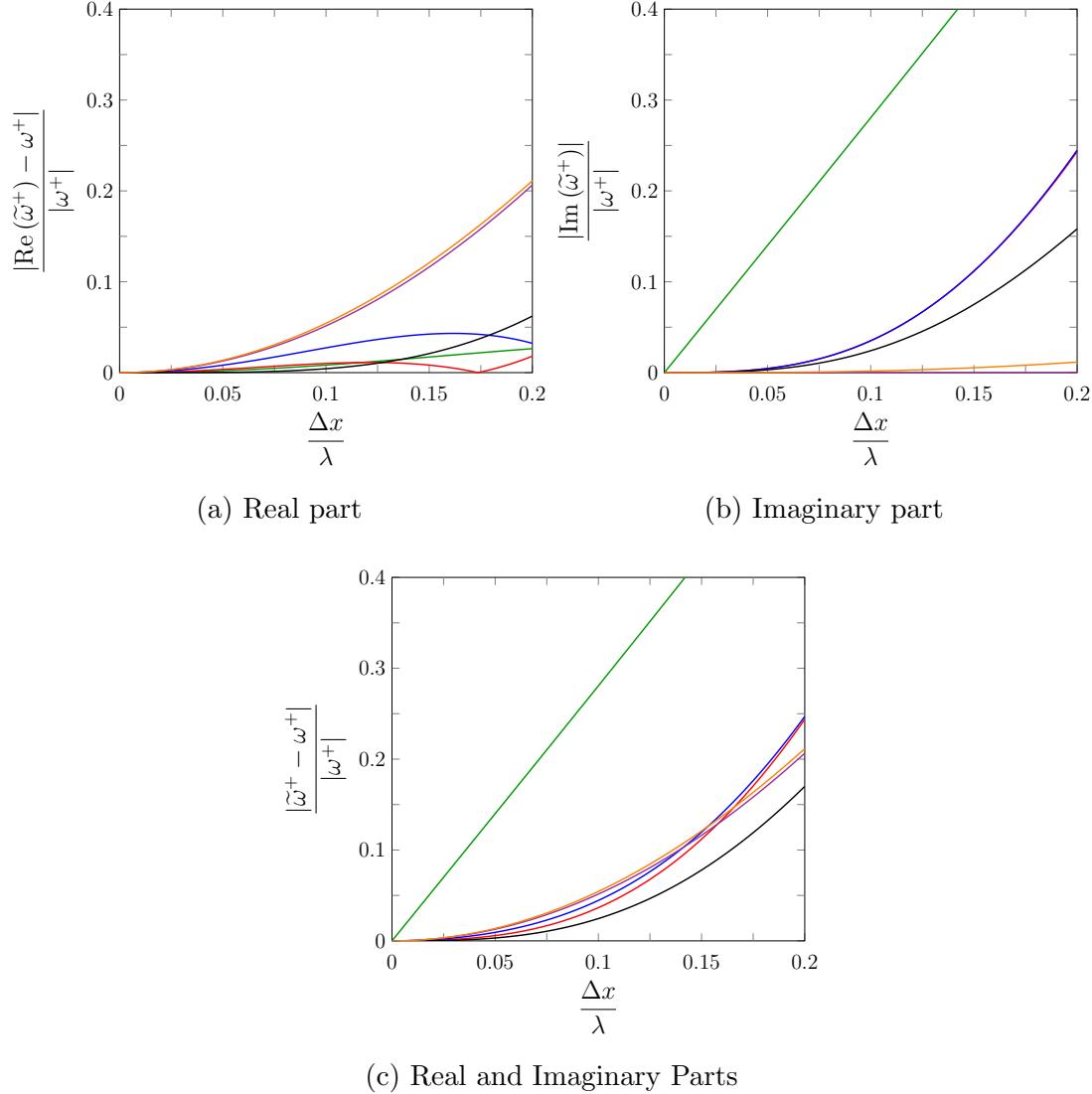


Figure 4.5: Relative dispersion error against $\Delta x/\lambda$ when $H = 1m$, $k = 2.5m^{-1}$ and $U = 1m/s$ for FDVM₁ (—), FDVM₂ (—), FEVM₂ (—), FEVM₃ (—), \mathcal{D} (—) and \mathcal{W} (—).

Scheme	Lowest Order Terms of $\tilde{\omega}^+ - \omega^+$	
	Δx	Δt
FDVM ₁	$-\left(2\sqrt{gH} - \sqrt{\frac{3U}{\beta}}\right) \frac{ik^2}{4} \Delta x$	$\frac{i(\omega^+)^2}{2} \Delta t$
FDVM ₂	$\frac{2\beta U - 3\sqrt{3gH\beta}}{\beta^2} \frac{k^3}{24} \Delta x^2$	$-\frac{(\omega^+)^3}{6} \Delta t^2$
FEVM ₂	$\left(U + \frac{(42 + 15k^2H^2)\sqrt{3gH\beta}}{20\beta^2}\right) \frac{k^3}{12} \Delta x^2$	$-\frac{(\omega^+)^3}{6} \Delta t^2$
FDVM ₃	$-\left(2\sqrt{gH} - \sqrt{3\beta}U\right) \frac{ik^4}{24} \Delta x^3$	$-\frac{i(\omega^+)^4}{24} \Delta t^3$

Table 4.2: Lowest order terms for Taylor series of $\tilde{\omega}^+ - \omega^+$ for all FDVM and the FEVM. With $-1 \leq Fr \leq 1$ and $\beta = 3 + H^2k^2$.

Scheme	Lowest Order Δx Term of $\tilde{\omega}^+ - \omega^+$	
	$Fr < -1$	$Fr > 1$
FDVM ₁	$-\left(2U + \sqrt{\frac{3gH}{\beta}}\right) \frac{ik^2}{4} \Delta x$	$\left(2U + \sqrt{\frac{3gH}{\beta}}\right) \frac{ik^2}{4} \Delta x$
FDVM ₃	$-\left(2U + \sqrt{\frac{3gH}{\beta}}\right) \frac{ik^4}{24} \Delta x^3$	$\left(2U + \sqrt{\frac{3gH}{\beta}}\right) \frac{ik^4}{24} \Delta x^3$

Table 4.3: Lowest order Δx term for Taylor series of $\tilde{\omega}^+ - \omega^+$ for all FDVM for supercritical Froude numbers where different from Table 4.2. With $\beta = 3 + H^2k^2$.

Scheme	Lowest Order Terms of $\tilde{\omega}^+ - \omega^+$	
	Δx	Δt
\mathcal{D}	$-\chi\Delta x^2$	$-\frac{(\omega^+)^3}{3}\Delta t^2$
\mathcal{W}	$\chi\Delta x^2$	$\begin{aligned} & \frac{1}{\beta^2} \left(\beta U^2 [9\sqrt{3gH\beta} + 4\beta U] \right. \\ & \left. + 3gH^2 [\sqrt{3gH\beta} + 6\beta U] \right) \frac{k^3}{18} \Delta t^2 \end{aligned}$

Table 4.4: Lowest order terms of the Taylor series of $\tilde{\omega}^+ - \omega^+$ for \mathcal{D} and \mathcal{W} .
With $\beta = 3 + H^2 k^2$ and $\chi = \left(U + \frac{(4 + H^2 k^2) \sqrt{3gH\beta}}{4\beta^2} \right)$.

Chapter 5

Numerical Validation

In this chapter analytic and forced solutions are used to validate the numerical methods. These numerical methods are:

- The second-order Finite Element Volume Method (FEVM) termed FEVM₂ which is described in Chapter 3.
- The first-, second- and third-order Finite Difference Volume Methods (FDVM) termed FDVM₁, FDVM₂ and FDVM₃ respectively. Descriptions of these methods were published by Zoppou et al. [15]. Additionally, FDVM₂ was extended to handle dry beds using the technique outlined in Chapter 3.
- The second-order finite difference methods \mathcal{D} and \mathcal{W} whose descriptions were published by Pitt et al. [18].

For all these numerical methods a constant spatial and temporal resolution is used to keep the analysis simple.

The analytic and forced solutions described in Chapter 2 are used to verify that the numerical methods have the expected convergence and conservation properties. Firstly, the measures of convergence and conservation used to verify the methods are introduced. These measures are then used to compare all the numerical methods for the solitary travelling wave solution. The convergence and conservation properties of FDVM₂ and FEVM₂ are then compared using the lake at rest solution. Currently, the FDVM₂ and FEVM₂ are the only methods in this thesis that incorporate varying bathymetry.

Finally FDVM₂ and FEVM₂ are validated using forced solutions which test the accuracy of their approximations to all terms in the Serre equations. The forced solutions are obtained by adding terms to the Serre equations (2.13) to

force any desired solution. Since these forced solutions are arbitrary they allow the method to be validated against more flow scenarios than possible given the limited number of currently known analytic solutions. However, since these forced solutions are arbitrary they are no longer necessarily conservative. Therefore, the forced solutions are only used to assess the convergence properties of these numerical methods.

5.1 Measuring Convergence and Conservation

The convergence of the numerical methods is studied by comparing their numerical solutions to the analytic solutions or the forced solutions of the Serre equations. While conservation is investigated by comparing the total amount of a conserved quantity in a numerical solution at some time with the total amount of that quantity present in the initial conditions. The notation for these measures is now introduced and their calculation described, beginning with convergence.

5.1.1 Measure of Convergence

By measuring the relative difference between the numerical and analytic solutions as Δx varies, the convergence of the numerical methods can be investigated. To measure the relative difference the L_2 vector norm is used to compare the numerical and analytic solutions at the cell midpoints x_j at the end of the simulations. For a quantity q , the vector of its values \mathbf{q} at the cell midpoints x_j and the corresponding numerical solution at those locations \mathbf{q}^* ; the L_2 norm is

$$L_2(\mathbf{q}, \mathbf{q}^*) = \begin{cases} \frac{\|\mathbf{q}^* - \mathbf{q}\|_2}{\|\mathbf{q}\|_2} & \|\mathbf{q}\|_2 > 0 \\ \|\mathbf{q}^*\|_2 & \|\mathbf{q}\|_2 = 0. \end{cases}$$

5.1.2 Measures of Conservation

The conservation properties of the methods are established by calculating the total amount of a conserved quantity in the numerical solution $\mathcal{C}^*(\mathbf{q}^*)$ at the end of the simulation and comparing it to the total amount of that quantity present

in the initial conditions $\mathcal{C}(q(x, 0))$ derived analytically. The relative measure

$$C(q, \mathbf{q}^*) = \begin{cases} \frac{|\mathcal{C}^*(\mathbf{q}^*) - \mathcal{C}(q(x, 0))|}{|\mathcal{C}(q(x, 0))|} & |\mathcal{C}(q(x, 0))| > 0 \\ |\mathcal{C}^*(\mathbf{q}^*)| & |\mathcal{C}(q(x, 0))| = 0 \end{cases} \quad (5.1)$$

is used. The total amount in a numerical solution $\mathcal{C}^*(\mathbf{q}^*)$ was calculated using 3 point Gaussian quadrature over the j^{th} cell and summing these cell integrals for all j . The value of q at the three points needed to perform the Gaussian quadrature were calculated by interpolating the j^{th} cell using a quartic polynomial that fits the nodal values $q_{j-2}, q_{j-1}, q_j, q_{j+1}$ and q_{j+2} at the surrounding cell midpoints. Gaussian quadrature using three points is 5^{th} order accurate and interpolation by quartics is 5^{th} order accurate for the quantity q and 4^{th} order accurate for its spatial derivative $\partial q / \partial x$. Since all methods are third-order accurate or less, the error introduced by the calculation of $\mathcal{C}^*(\mathbf{q}^*)$ for h , uh , G and \mathcal{H} will be dominated by the error introduced by the numerical solvers.

In some cases $\mathcal{C}(q(x, 0))$ may be difficult to derive analytically, when this occurs, it is approximated with $\mathcal{C}^*(\mathbf{q}^0)$ in (5.1); where \mathbf{q}^0 is the vector of the quantity at the cell midpoints used as the initial conditions of the numerical method. The numerical approximation to the conservation error (5.1) is denoted by C^* . Errors generated by the discretisation will be similar for $\mathcal{C}^*(\mathbf{q}^0)$ and $\mathcal{C}^*(\mathbf{q}^*)$ and will therefore cancel when studying C^* . Hence, the completely numerically calculated conservation error will also be useful in cases where the errors introduced by the discretisation dominate.

5.2 Solitary Travelling Wave Solution

To assess the ability of the numerical methods to solve the Serre equations with a horizontal bed the solitary travelling wave solution (2.11) described in Chapter 2 was used. The solitary travelling wave solution is a member of the family of periodic travelling wave solutions [32]. Every member of this family of solutions except the trivial stationary one have the same non-zero terms and thus provide similar tests for the numerical methods. Hence, it is sufficient to only study the solitary travelling wave solution.

For the solitary wave solution all the terms in (2.8) must be adequately approximated by the numerical method to properly reproduce the analytic solution. Therefore, this analytic solution serves as a good benchmark for the ability of a

numerical method to accurately solve the Serre equations with a horizontal bed for smooth solutions.

For the numerical tests the solitary travelling wave solution (2.11) with $a_0 = 1m$, $a_1 = 0.7m$ and $g = 9.81m/s^2$ at $t = 0s$ was used as the initial conditions. The spatial domain was $[-250m, 250m]$ and the problem was solved until $t = 50s$. Various numerical solutions were produced with a range of Δx values that had the following form; $\Delta x = 100/2^k m$ with $k \in [6, \dots, 19]$. The CFL condition was satisfied with CFL number $Cr = 0.5$ by setting $\Delta t = Cr\Delta x / \sqrt{g(a_0 + a_1)}$. For FDVM₂ and FEVM₂ the limiting parameter $\theta = 1.2$ was used in the generalised minmod limiter (3.2) employed by both methods during the reconstruction step. While FDVM₃ used the Koren limiter [57] in its reconstruction [15], which has no limiting parameter.

For the parameters $a_0 = 1m$ and $a_1 = 0.7m$ the non-linearity is $\epsilon = a_1/a_0 = 0.7$; this is large but beneath most of the well known breaking thresholds for water waves $\epsilon \leq 0.8$ [58]. Because ϵ is large the non-linear effects are large and therefore, so are the balancing dispersive effects making this particular analytic solution a rigorous test of the numerical methods. With the chosen spatial domain and the final time $t = 50s$ there is no interaction of the wave with the boundary, therefore the Dirichlet boundary conditions were appropriate.

The results of this analytic solution validation were published by Zoppou et al. [15] for FDVM₁, FDVM₂ and FDVM₃. These results have been expanded here to include an investigation of the convergence of G and the conservation of h , uh and G .

An example numerical solution with $\Delta x = 100/2^{11}m \approx 0.049m$ from all methods was plotted in Figure 5.1 against the analytic solution at $t = 50s$. Only an illustrative amount of points in the numerical solution have been plotted. From these plots it is evident that FDVM₁ performs significantly worse than the high-order methods at reproducing the analytic solution, even for this relatively fine grid where the wave is captured by more than 200 cells. The poor performance of FDVM₁ is due to the numerical diffusion introduced by the method, which has caused the wave in the numerical solution to decrease in amplitude and widen significantly. The high-order numerical methods all accurately replicate the analytic solution, with insignificant visual differences in these plots due to the high resolution of the grid.

The L_2 error was calculated for h , u and G for all numerical solutions and was plotted against Δx for all numerical methods in Figure 5.2. From these plots it can be seen that all numerical methods are convergent. The rate at which

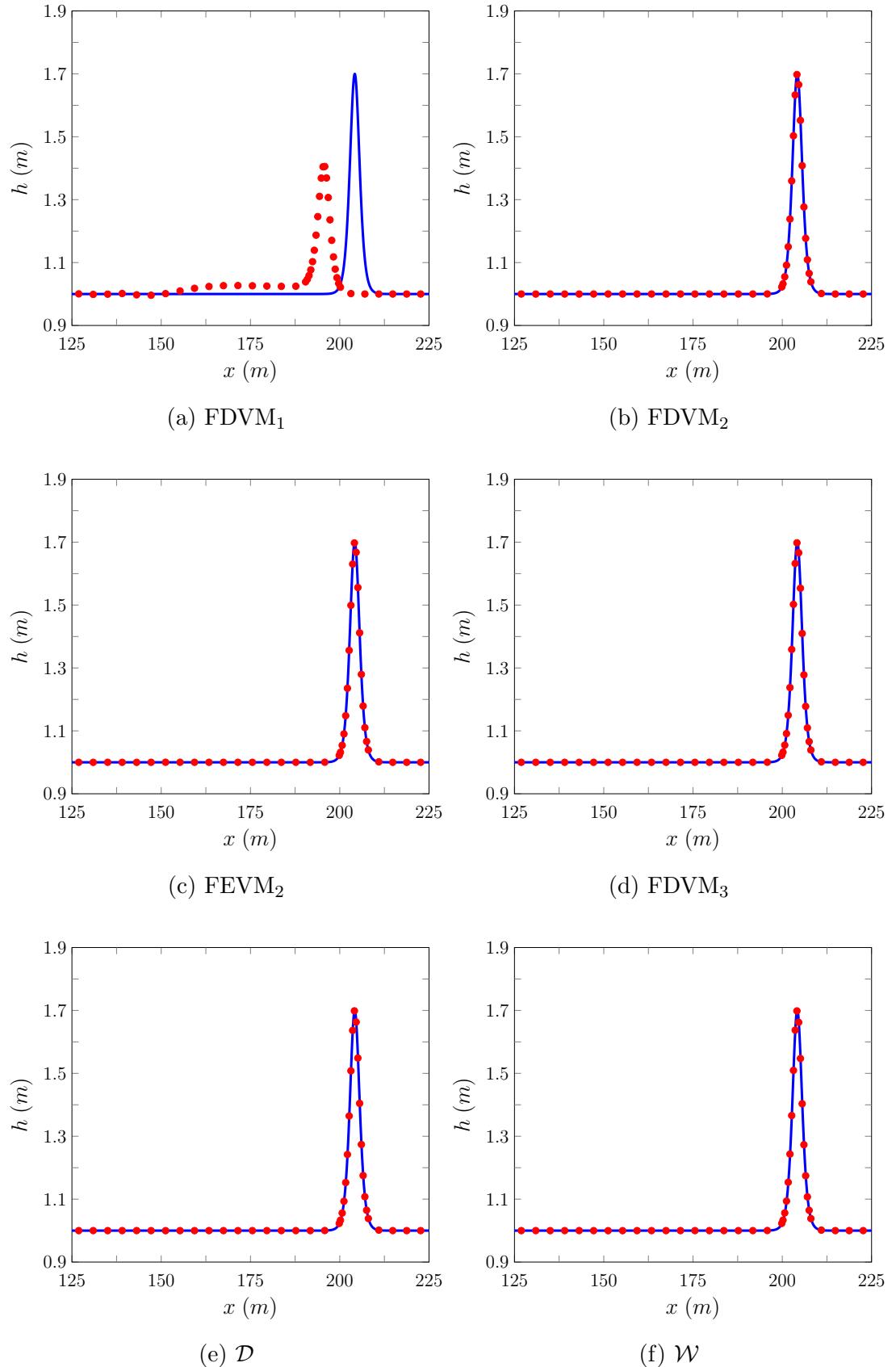


Figure 5.1: Comparison of the analytic solution (—) and numerical solution with $\Delta x = 100/2^{11}m$ (●) for the solitary travelling wave solution at $t = 50s$ for all methods.

the numerical solutions converge to the analytic solution as a function of Δx is determined by the order of accuracy of the numerical scheme. All methods demonstrate the expected order of accuracy given the order of accuracy of the approximations used in the method. The convergence results of all methods agree with the results of the linear analysis in Chapter 4 and Appendix C.

All methods more accurately reproduced the analytic solution for h than either G or u across all Δx values. This is due to the simplicity of the evolution equation for h (2.6a) compared to the evolution equations of G (2.6b) and uh (2.4b). While the error in G is transmitted to u when solving (2.7) in the finite volume based methods.

Increasing the order of accuracy of the numerical methods leads to smaller errors when comparing two numerical solutions for the same Δx value, as Figure 5.2 demonstrates. This behaviour is consistent with the example numerical solution in Figure 5.1, where the lowest order accuracy scheme, FDVM₁ had the poorest reproduction of the analytic solution. However, there is only a slight benefit from moving from the second-order FEVM₂ and FDVM₂ to the third-order FDVM₃.

For the second-order methods it is observed that FDVM₂ consistently produces the smallest L_2 error followed by FEVM₂, \mathcal{W} and \mathcal{D} . The difference between the FDVM₂ and FEVM₂ is significant with the errors of FEVM₂ being 2 to 4 times larger than those of FDVM₂. Therefore, FDVM₂ is reproducing the solitary wave solution more accurately than FEVM₂.

The finite difference methods produce very similar errors which are twice as large as the errors of FEVM₂. Additionally, the round-off effects dominate the L_2 error of the finite difference methods at larger Δx values than for the finite volume based methods.

The error in conservation C was calculated for all methods using the analytic expressions for the total amounts of the conserved quantities in the initial conditions (A.1) provided in Appendix A. The error in conservation was plotted against the spatial resolution in Figure 5.3. These results demonstrate that due to the use of the finite volume methods for h and G , both are conserved at machine precision for all the finite volume based methods as expected. While the finite difference methods only conserved h at machine precision because the employed finite difference method for the continuity equation (2.6a) is a conservative method.

No methods conserve \mathcal{H} or uh within machine precision. Since none of the methods were designed to conserve these quantities this is not surprising, although the error in conservation of all methods for these quantities does exhibit the order

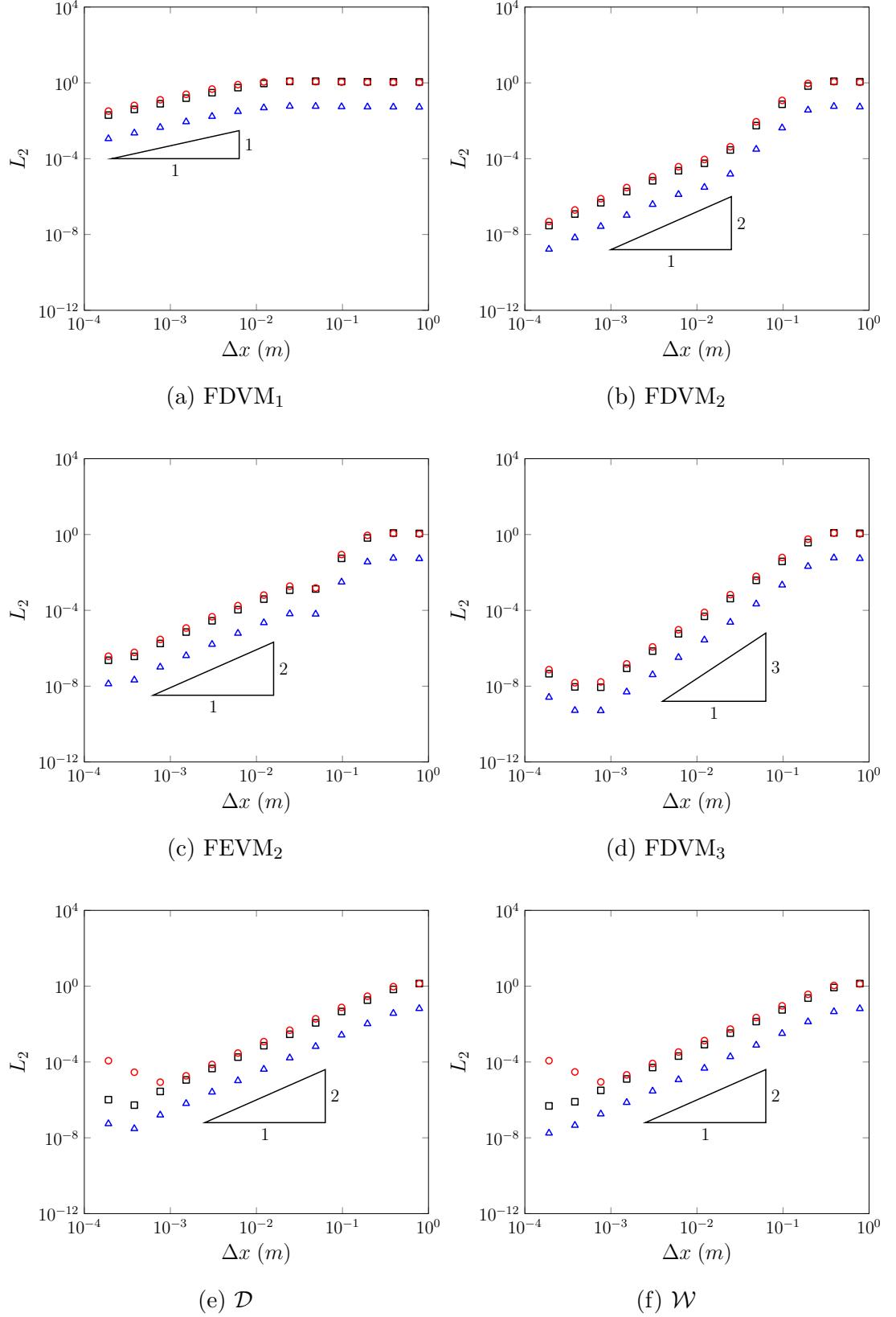


Figure 5.2: Convergence as measured by the L_2 norm against Δx for h (Δ), u (\square) and G (\circ) for the soliton problem for all methods.

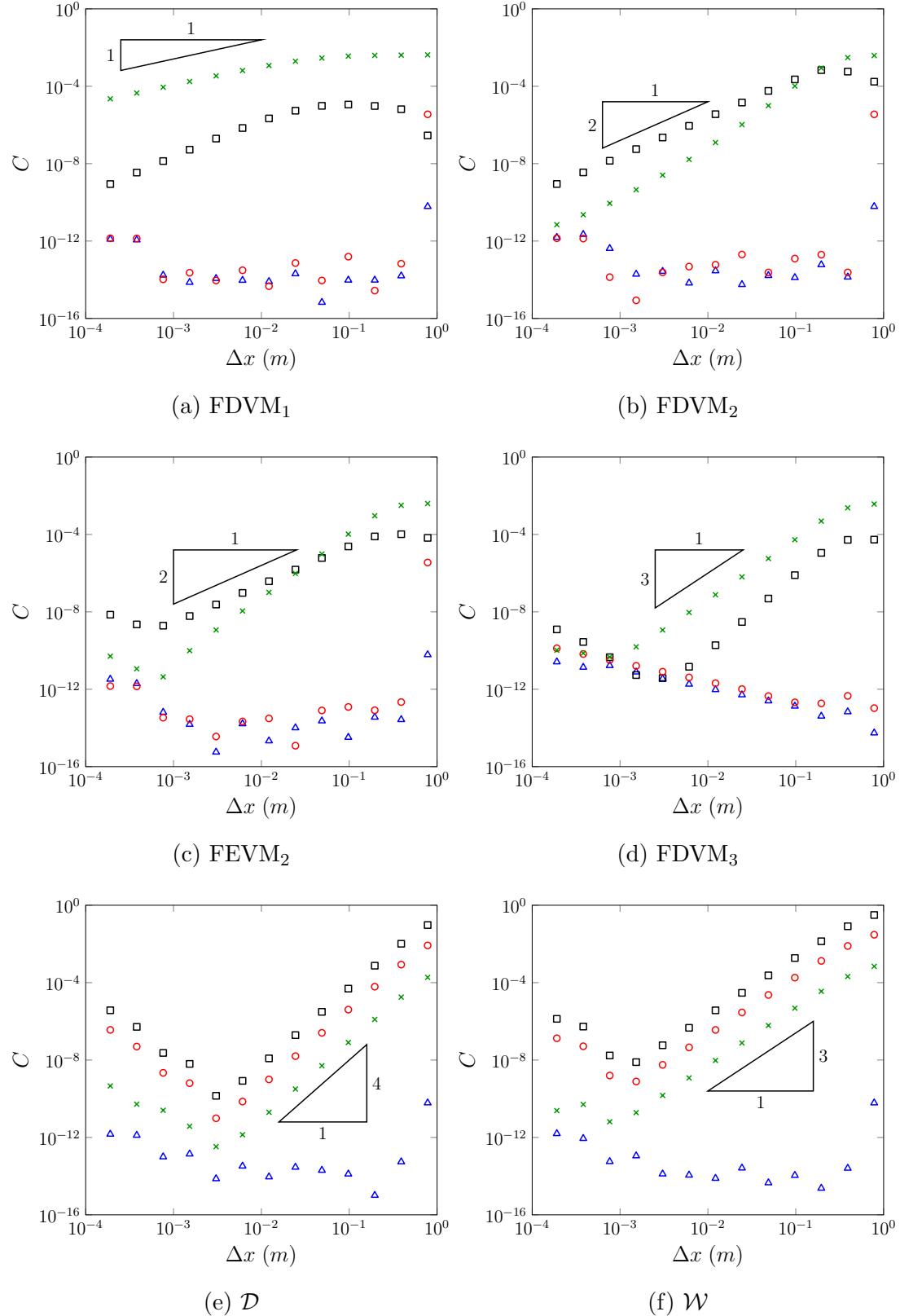


Figure 5.3: Conservation error C against Δx for h (Δ), uh (\square), G (\circ) and \mathcal{H} (\times) for the solitary travelling wave solution for all methods.

of accuracy of the convergence of the numerical method or better, as expected.

For small Δx values the round-off errors dominate the conservation error, particularly for the finite difference methods. Interestingly, FDVM₃ has an accumulation of round-off error increasing the conservation error for h and G as Δx decreases. This was found to be caused by the Runge-Kutta coefficients of the third-order time stepping method [15] not being exactly represented as floating point numbers. For the third-order SSP Runge-Kutta time stepping method the coefficients in the last step are $1/3$ and $2/3$. Since these numbers are not exactly represented in floating point they are approximated with a small error that when summed does not maintain the conservation properties of h and G . Thus, every time step accumulates a small conservation error of machine precision size leading to the observed increase as Δx becomes small and the number of time steps increases. Remedies for this error in conservation were attempted such as using the coefficients $1/3$ and $(1 - 1/3)$ and bringing the common divisor out but this problem persisted. Some other numerical techniques are required to resolve this issue such as those of Higham [59]. Ultimately, since the convergence of FDVM₃ was only slightly better than FEVM₂ and FDVM₂ [15] it was not developed further and this issue was not resolved.

These results demonstrate the need for high-order accurate schemes to accurately approximate the Serre equations. Furthermore, the results suggest that second-order accuracy is sufficient, with third-order accurate schemes showing only a slight improvement. This was also the conclusion of the analytic validation by Zoppou et al. [15]. Finally, these results demonstrate the ability of FEVM and FDVM to conserve h and G up to machine precision, as desired. Given these results, only FEVM₂ and FDVM₂ have been extended to allow for variable bathymetry and dry beds. Consequently, the rest of the results in this chapter and Chapter 6 will only consider FEVM₂ and FDVM₂.

5.3 Lake at Rest Solution

To verify the validity of the numerical methods for the Serre equations with variable bathymetry and assess the well-balancing modifications various numerical solutions to the lake at rest solution (2.12) were compared.

The particular lake at rest solution (2.12) associated with the bed profile

$$b(x) = a_1 \sin(a_2 x) \quad (5.2)$$

was chosen for this validation to ensure that all terms with derivatives of the bed

were tested. To demonstrate the capability of the methods in the presence of dry and wet beds the parameter values $a_0 = 0m$, $a_1 = 1m$ and $a_2 = 2\pi/50m^{-1}$ were chosen. These parameter values result in wet regions with a horizontal free surface where the stage $w(x, t) = h(x, t) + b(x) = a_0 = 0$ (2.12). Therefore, the bed is periodic and water submerges the troughs of the bed while the peaks of the bed are dry.

For the numerical solutions, the spatial domain was $x \in [-112.5m, 87.5m]$ and the final time was $t = 10s$, with the standard gravitational acceleration $g = 9.81m/s^2$. The spatial resolution of the method was varied so that $\Delta x = 100/2^k m$ with $k \in [8, \dots, 17]$ and the CFL condition (3.23) was satisfied by having $\Delta t = Cr\Delta x/\sqrt{g}$ with condition number $Cr = 0.5$. The standard limiting parameter $\theta = 1.2$ was used in the generalised minmod limiter, (3.2) for both FEVM₂ and FDVM₂. Dirichlet boundary conditions were used at both ends as the analytic solution is stationary.

The numerical methods are assessed using the specified lake at rest solution as initial conditions and comparing the numerical solutions of FEVM₂ and FDVM₂ at $t = 10s$ to the analytic solution, which are the initial conditions. To demonstrate the utility of the well-balancing method the results from two versions of FEVM₂ and FDVM₂ are presented, where the well-balancing method described in Chapter 3 is and is not employed.

Example numerical solutions with $\Delta x = 100/2^{10}m \approx 0.0977m$ at $t = 10s$ for all versions of FEVM₂ and FDVM₂ are given in Figure 5.4. The numerical solutions in these figures are indistinguishable from the analytic solutions at this scale and so the analytic solutions have been omitted from the plots.

Examination of the L_2 errors depicted in Figure 5.5 reveals that only the well-balanced methods have accurately recovered the analytic solution. Because, only the well-balanced versions of the methods reproduced h , G and u precisely. However, as Δx decreased the round-off errors accumulated, causing the observed increase in the error in h , G and u for the well-balanced methods.

For methods without well-balancing; the errors are significantly larger, yet they are converging to the analytic solution. However, the order of accuracy of the convergence in u and G has degraded and is not the expected second-order accuracy observed for h . The poor convergence of u and G is a result of the errors in u and G not being damped by the method. Thus, errors generated by the imbalance between the flux and source terms increase over time degrading the order of accuracy. The second-order accuracy in h is retained for the presented Δx values as the errors in u and G are small for a single cell, although for smaller

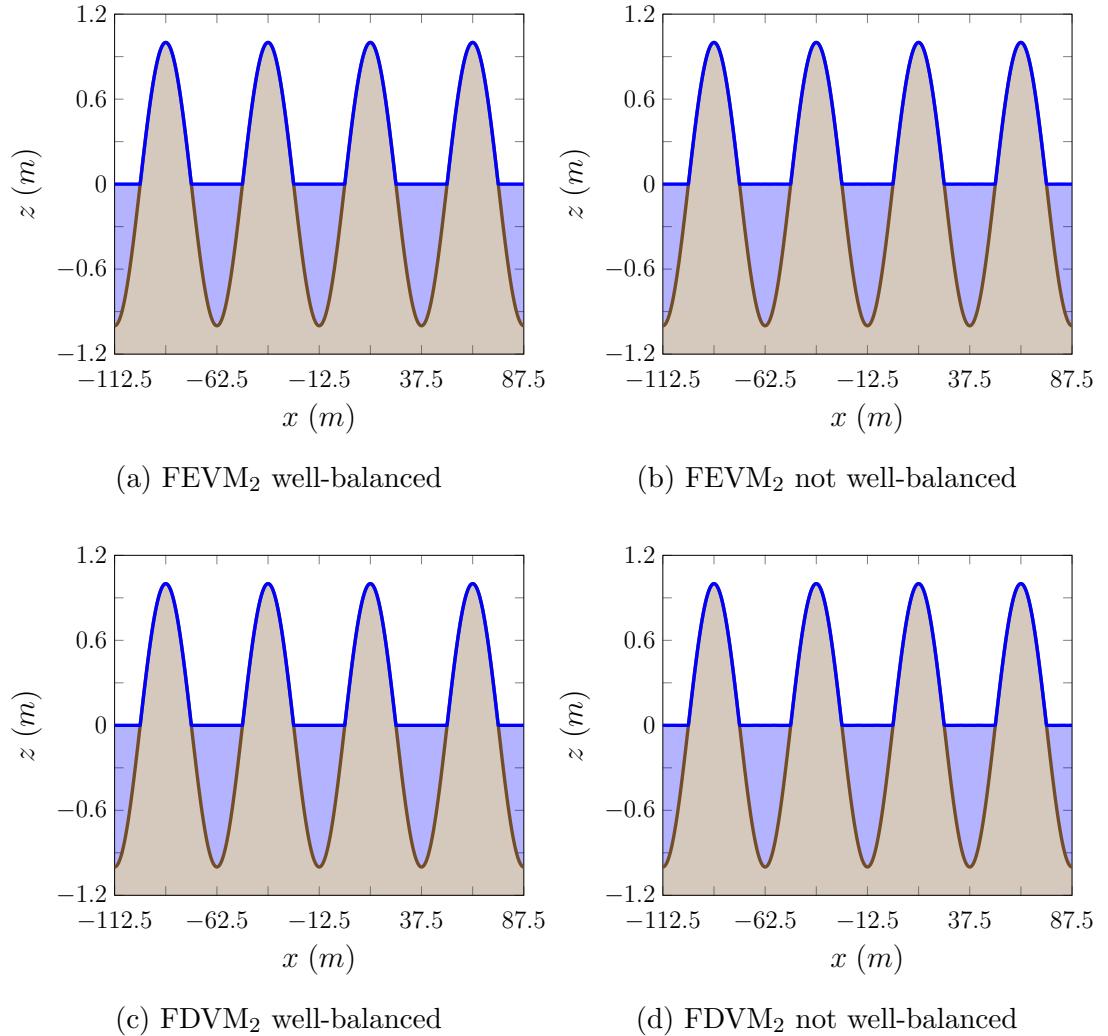


Figure 5.4: Numerical solutions for w (■) and b (□) with $\Delta x = 100/2^{10}m$ for the lake at rest problem at $t = 10s$ for FEVM₂ and FDVM₂.

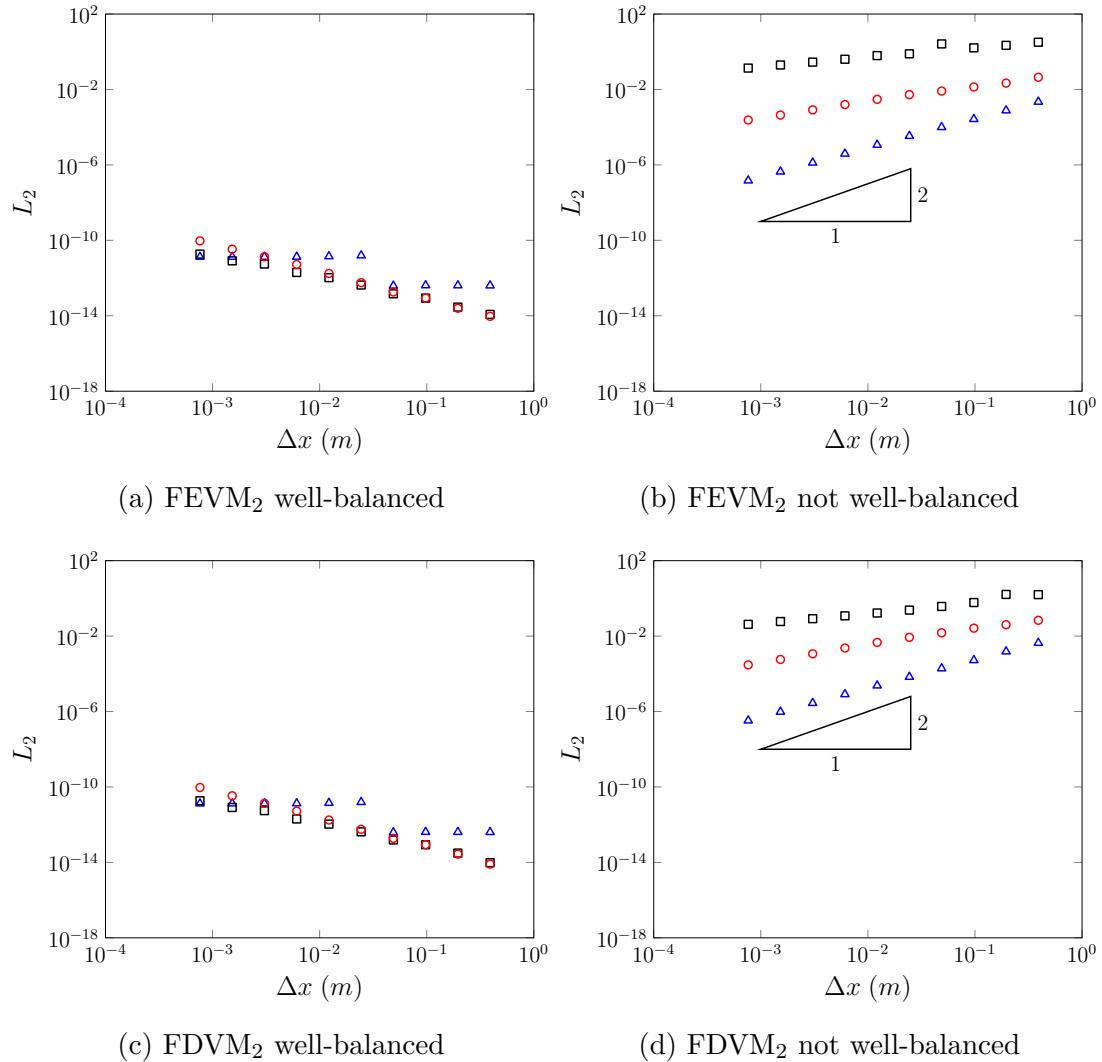


Figure 5.5: Convergence as measured by the L_2 norm against Δx for h (Δ), u (\square) and G (\circ) for the lake at rest problem at $t = 10s$ for FEVM₂ and FDVM₂.

Δx values the errors in u and G will begin to dominate the errors in h .

Using the expressions in Appendix A for the total amounts of the conserved quantities, the conservation error C was calculated for FEVM₂ and FDVM₂ with the results plotted in Figure 5.6. The error in conservation of these methods affirms the superiority of the well-balanced version of the methods. In particular, it is observed that the total amounts of uh and G are only conserved within machine precision when well-balancing is employed. Since uh and G are uniformly zero in the initial conditions the well-balanced methods have only introduced round-off errors into these quantities, whilst without well-balancing large errors in these quantities are introduced in the methods.

The errors in conservation of h and \mathcal{H} for the well-balanced methods are large, but do converge at the order of accuracy of the scheme or better. These errors are caused by the discretisation of the initial conditions, primarily the approximation of the boundaries of the wet regions by the numerical grids. The initial discretisation error can be removed by comparing the total amounts of the conserved quantities in the numerical solution to their numerically calculated total amounts in the initial conditions with C^* as in Figure 5.7. For the completely numerically calculated conservation error C^* all the conserved quantities are conserved at machine precision for the well-balanced methods. Furthermore, \mathcal{H} is conserved exactly for most numerical solutions, hence its disappearance from the log-log plot. The conservation error of \mathcal{H} is small for the lake at rest solution since u is small. Hence, \mathcal{H} is essentially the gravitational potential energy. Since mass is well conserved then so is the gravitational potential energy and therefore \mathcal{H} .

These results demonstrate the need for well-balancing for both numerical methods, as it is only with its inclusion that the lake at rest steady state can be accurately reproduced.

5.4 Forced Solutions

The previous analytic solution validations do not provide a stringent test for all terms present in the Serre equations and there are currently no known analytic solutions that do. To remedy this the forced solutions introduced in Chapter 2 were used to validate the numerical methods. Since the added terms in the modified Serre equations, (2.13) can be determined and accounted for analytically, the only source of error in the numerical solutions are the numerical methods

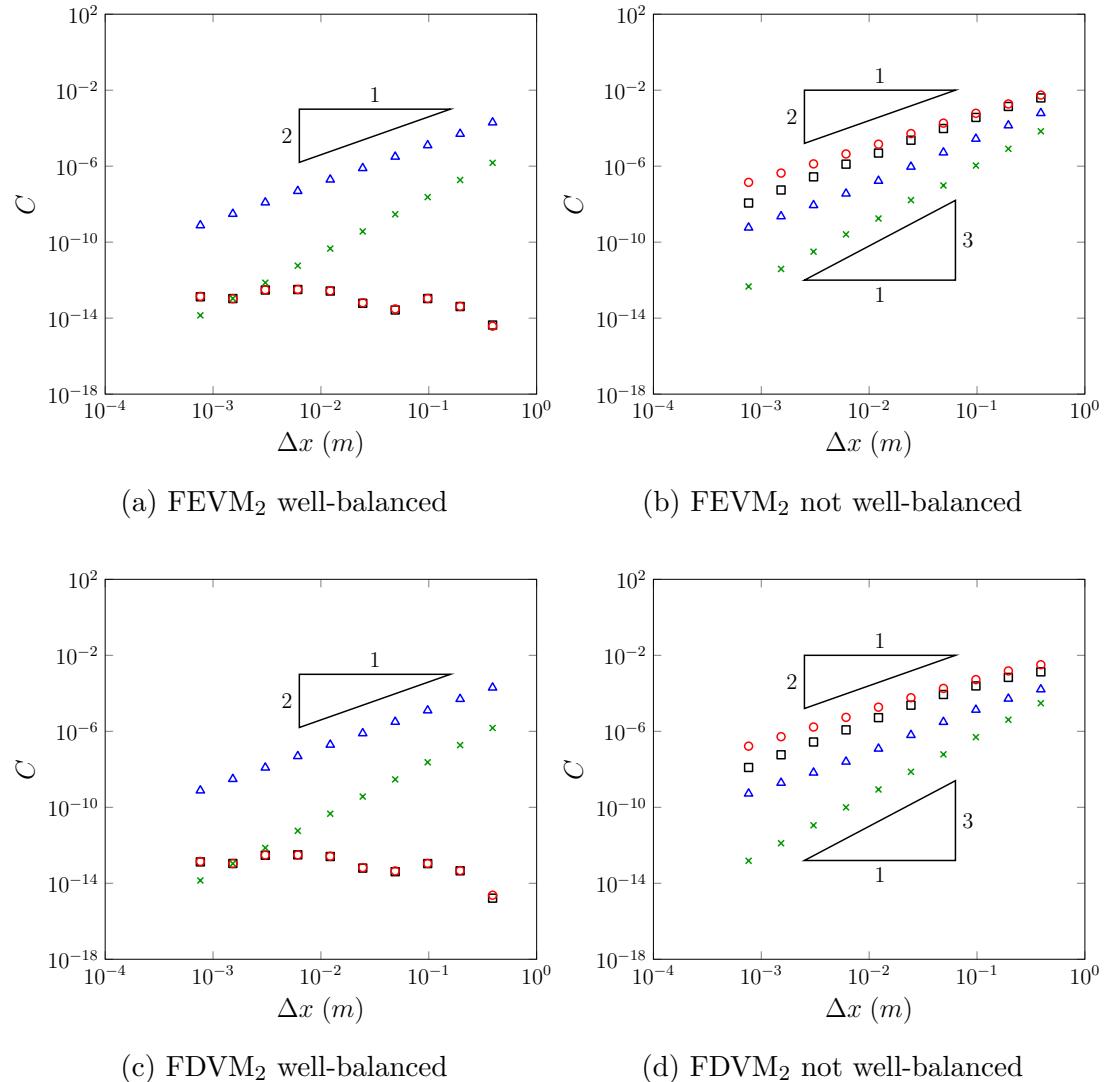


Figure 5.6: Conservation error C against Δx for h (Δ), uh (\square), G (\circ) and \mathcal{H} (\times) for the lake at rest problem at $t = 10s$ for FEVM₂ and FDVM₂.

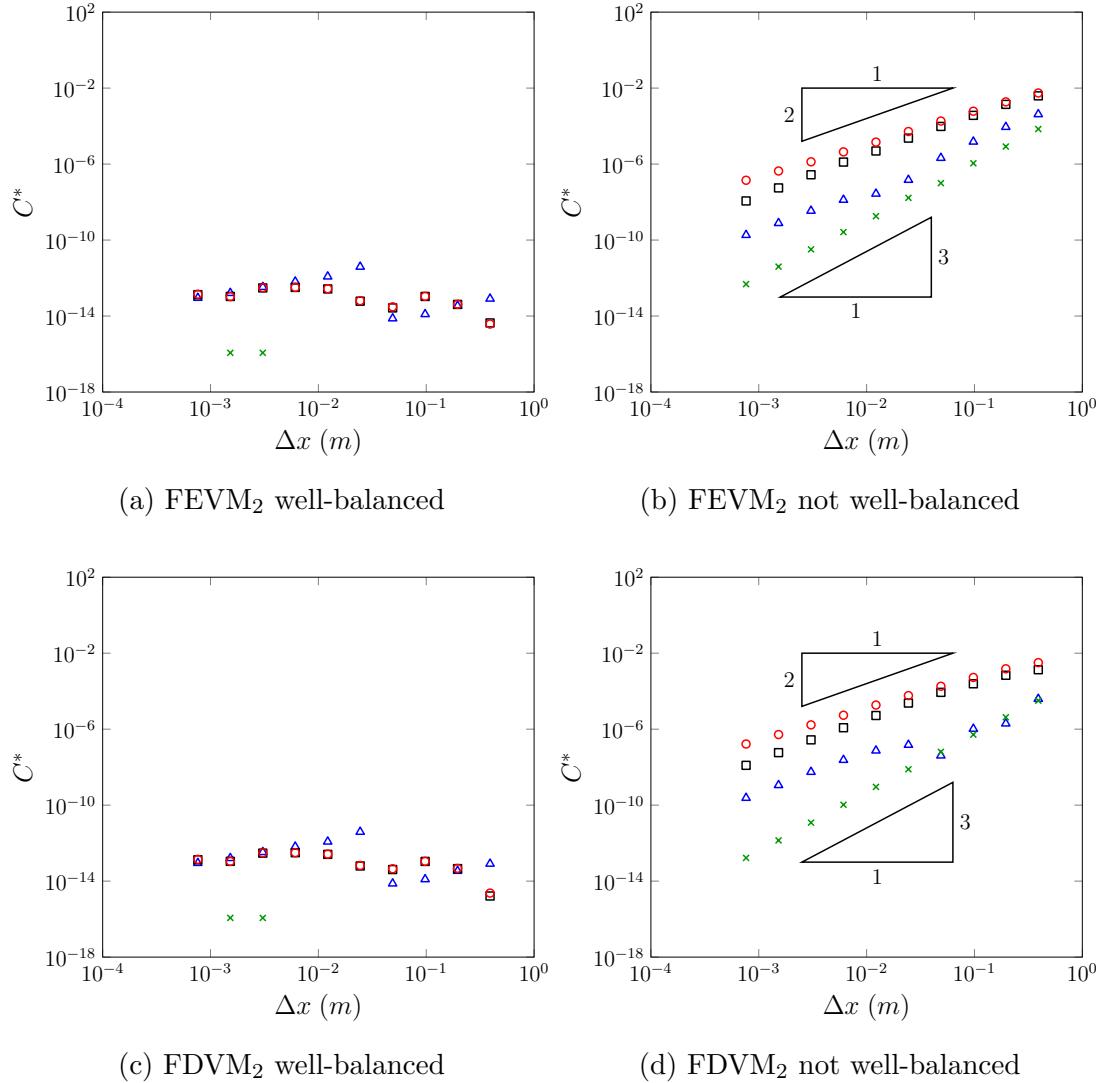


Figure 5.7: Conservation error using only numerical calculations C^* against Δx for h (Δ), uh (\square), G (\circ) and \mathcal{H} (\times) for the lake at rest problem at $t = 10s$ for FEVM₂ and FDVM₂.

themselves and thus the theoretical second-order accuracy of FEVM₂ and FDVM₂ should be recovered.

Validation tests were performed for two forced solutions; one with a finite water depth everywhere and the other with a dry bed to validate and compare the numerical solutions in both situations. To ensure that all terms of the Serre equations were accurately approximated in the numerical method the functions

$$h^*(x, t) = a_0 + a_1 \exp\left(-\frac{[(x - a_2 t) - a_3]^2}{2a_4}\right), \quad (5.3a)$$

$$u^*(x, t) = a_5 \exp\left(-\frac{[(x - a_2 t) - a_3]^2}{2a_4}\right), \quad (5.3b)$$

$$b^*(x) = a_6 \sin(a_7 x) \quad (5.3c)$$

for the primitive variables were chosen. These functions produce a Gaussian bump with an amplitude of a_1 for h and u that travels at a fixed speed a_2 over a periodic bed. Hence, h and u will have constant shape and travel to the right over time. However, this is not the case for G as u and h have constant shape but the bed is periodic. Since the bed varies the bed terms in G (2.7) will transform the shape of G as the Gaussian bump in h and u encounters different bed slopes.

For non-trivial choices of the parameters a_i all terms in the Serre equations vary in space and time and so all terms must be accurately approximated by the numerical method to adequately reproduce the forced solution.

Both validation studies used the values $a_1 = 0.5m$, $a_2 = 2\pi/(10a_7)m/s$, $a_3 = -3\pi/(2a_7)m$, $a_4 = \pi/(16a_7)m^2$, $a_5 = 0.5m/s$, $a_6 = 1.0m$ and $a_7 = \pi/25m^{-1}$ with $a_0 = 1m$ for the finite water depth forced solution and $a_0 = 0m$ for the dry bed forced solution. These parameter values result in a Gaussian bump in h and u that has a width much smaller than the wavelength of the bed profile and travels precisely one wavelength in $10s$.

The domain of the numerical solutions was $x \in [-112.5m, 87.5m]$ with $t \in [0s, 10s]$. The standard gravitational acceleration $g = 9.81m/s^2$ was used. The spatial resolution of the numerical methods was varied like so $\Delta x = 100/2^k m$ with $k \in [8, \dots, 17]$. To satisfy the CFL condition, (3.23) the temporal resolution $\Delta t = Cr\Delta x / (a_2 + a_5 + \sqrt{g(a_0 + a_1)})$ was chosen with condition number $Cr = 0.5$. The value $\theta = 1.2$ was used in the generalised minmod limiter (3.2) for both FEVM₂ and FDVM₂ and Dirichlet boundary conditions were applied at the boundaries of the domain.

5.4.1 Results for a Wet Bed

For the non-zero water depth case where $a_0 = 1m$ an example of the numerical solutions of FEVM₂ and FDVM₂ are given in Figures 5.8 and 5.9 respectively for $\Delta x = 100/2^{10}m \approx 0.0977m$ at various times. The numerical solutions accurately reproduced the forced solution as it travels over the bed. Thus, h and u maintain their constant shape while G does not due to the influence of the periodic bed.

The L_2 errors of h , u and G for the FEVM₂ and FDVM₂ are given in Figure 5.10. Both methods recover the expected second-order accuracy for h , u and G . The additional terms of the forced Serre equations are added analytically and all terms of the Serre equations must be accurately approximated by the method to reproduce this forced solution. Therefore, these results demonstrate that FEVM₂ and FDVM₂ are second-order accurate for all terms of the Serre equations when the bed is wet everywhere, as desired.

5.4.2 Results with a Dry Bed

To demonstrate the capability of the methods to accurately approximate the wetting and drying of a bed, a series of numerical simulations of the forced solutions (5.3a) where $a_0 = 0m$ were conducted using both FEVM₂ and FDVM₂.

Example numerical solutions demonstrating the evolution of the wave are given in Figure 5.11 for FEVM₂ and in Figure 5.12 for FDVM₂ with $\Delta x = 100/2^{10}m \approx 0.0977m$ at various times. The methods accurately reproduce the analytic solution for the stage w , h and G . However, both fail to accurately reproduce u when h is small, particularly behind the Gaussian bump. Consequently, h is the only quantity that maintains a constant shape in the numerical solutions, as G changes due to the periodic bed and u changes due to numerical errors.

The large errors in u when h is small are caused by the values $h_{base} = 10^{-8}$ and $h_{tol} = 10^{-12}$ used in the desingularisation transformation applied to the finite element method (3.12). Choosing larger values of these quantities damps the errors in u [54]. However, if h_{base} and h_{tol} are larger they begin to dominate the L_2 errors in h , G and uh making the convergence less obvious. This trade-off is present in all desingularisation transforms.

For the purposes of this validation the chosen desingularisation transform (3.27) with small h_{base} and h_{tol} values was sufficient, resulting in large observed errors in u when h is small.

The L_2 errors for h , u , uh and G for both methods are given in Figure 5.13. Both methods exhibit second-order convergence in all the quantities except u .

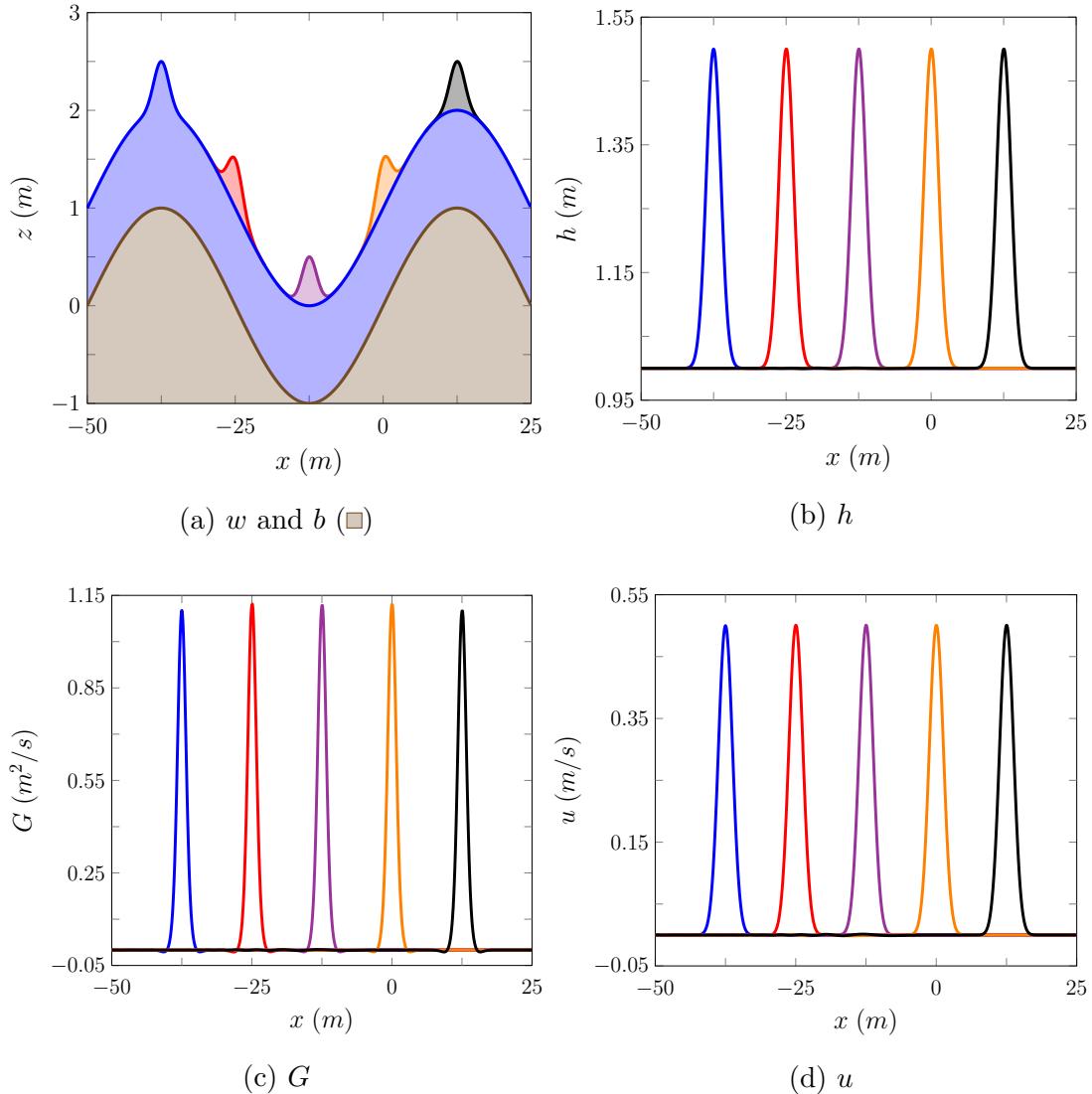


Figure 5.8: Numerical solutions for w , b , h , G and u produced by FEVM₂ with $\Delta x = 100/2^{10}m$ at $t = 0s$ (— / ■), $2.5s$ (— / □), $5.0s$ (— / ▨), $7.5s$ (— / □), $10.0s$ (— / ■) to the wet bed forced solution problem, where $a_0 = 1m$.

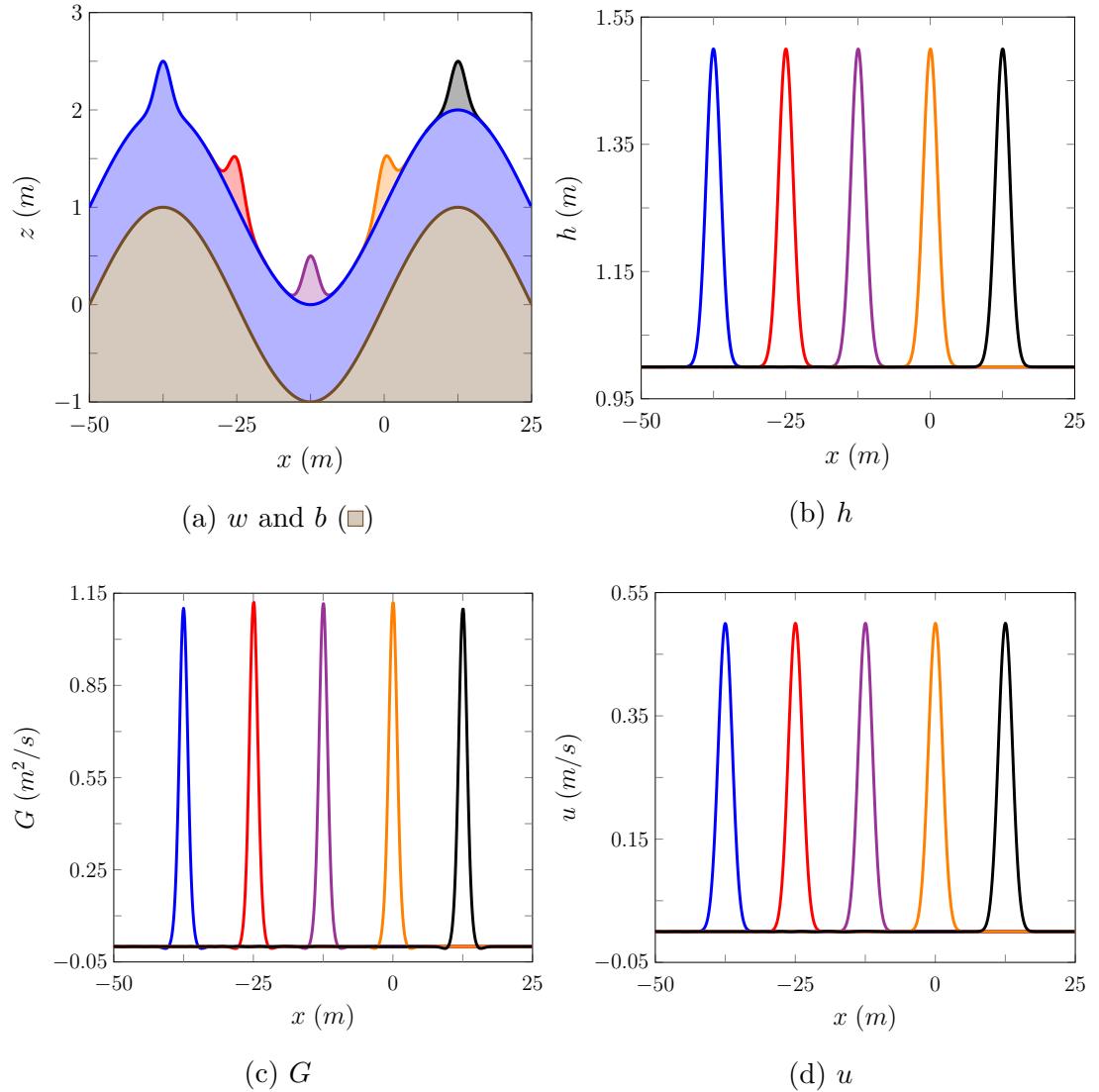


Figure 5.9: Numerical solutions for w , b , h , G and u produced by FDVM₂ with $\Delta x = 100/2^{10}m$ at $t = 0s$ (— / □), $2.5s$ (— / □), $5.0s$ (— / □), $7.5s$ (— / □), $10.0s$ (— / □) to the wet bed forced solution problem, where $a_0 = 1m$.

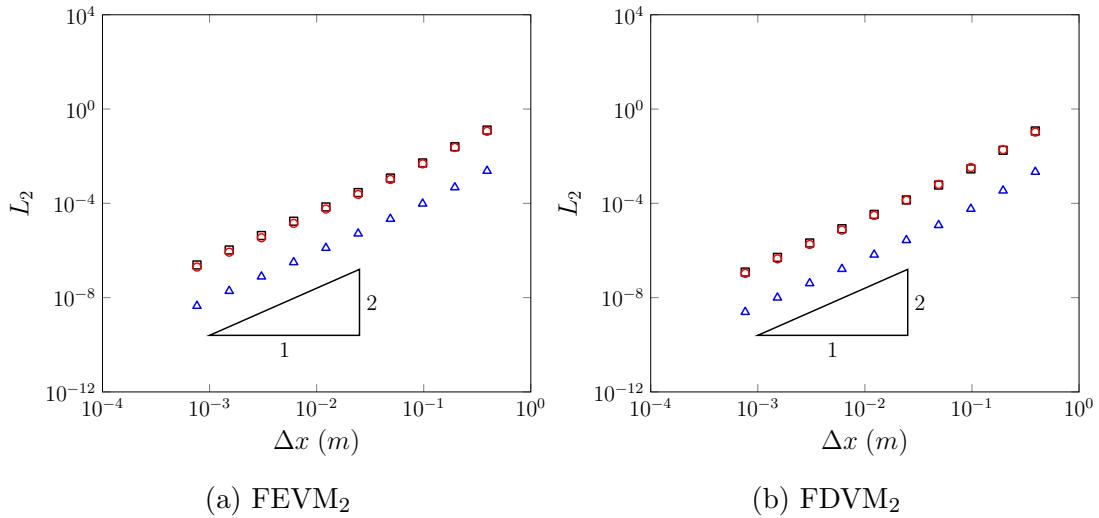


Figure 5.10: Convergence as measured by the L_2 norm against Δx for h (Δ), u (\square) and G (\circ) for the wet bed forced solution problem for FEVM₂ and FDVM₂ at $t = 10s$.

The large errors in u occur when h is small. Therefore, when the L_2 norm is restricted to regions where $h > 10^{-3}m$ the second-order accuracy in u is recovered, as observed in Figure 5.14. Since, all the flux and source terms of the Serre equations (2.6) depend on u multiplied by some power of h ; the large errors in u when h is small do not translate to significant errors in G , h or uh .

These results indicate that the methods accurately approximate the dry bed problem, even with small h_{base} and h_{tol} values. Although, in such cases the velocity may have large errors in regions where h is small. For physical applications where large errors in u when h is small are not acceptable it is recommended that the dry bed handling techniques are modified by increasing the h_{base} and h_{tol} values or altering the desingularisation transformation [54].

In this chapter the analytic and forced solutions were used to assess the numerical methods. It was found that the finite volume based methods performed better than the finite difference methods and that second-order methods were sufficient to accurately reproduce the analytic solutions. Finally, the second-order accuracy of FEVM₂ and FDVM₂ was confirmed for the wetting and drying of variable beds using forced solutions to the Serre equations. The analytically validated FEVM₂ and FDVM₂ will now be validated against experimental results.

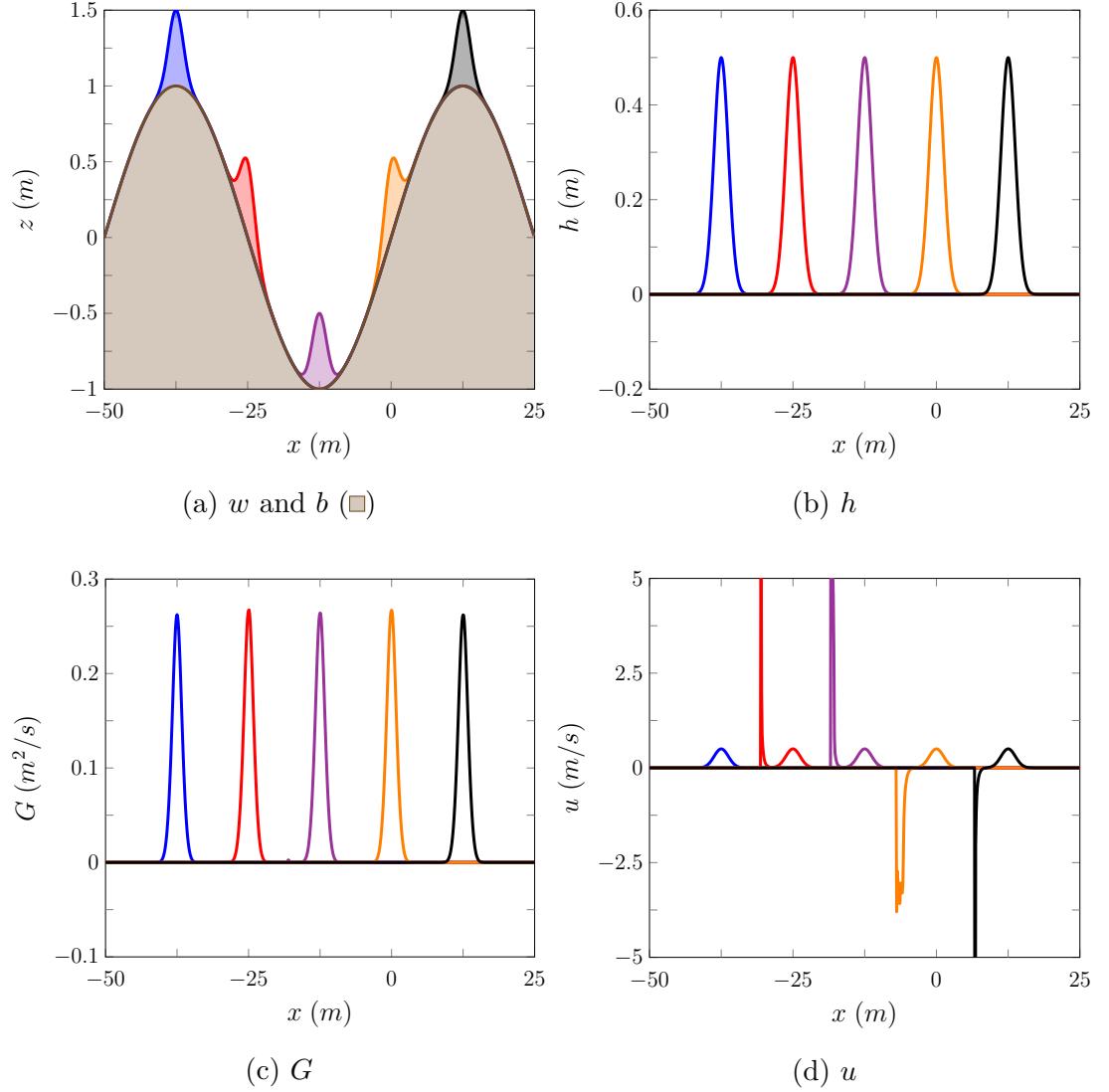


Figure 5.11: Numerical solutions for w , b , h , G and u produced by FEVM₂ with $\Delta x = 100/2^{10}m$ at $t = 0s$ (— / □), $2.5s$ (— / ▢), $5.0s$ (— / ▣), $7.5s$ (— / ▤), $10.0s$ (— / ▥) to the dry bed forced solution problem, where $a_0 = 0m$.

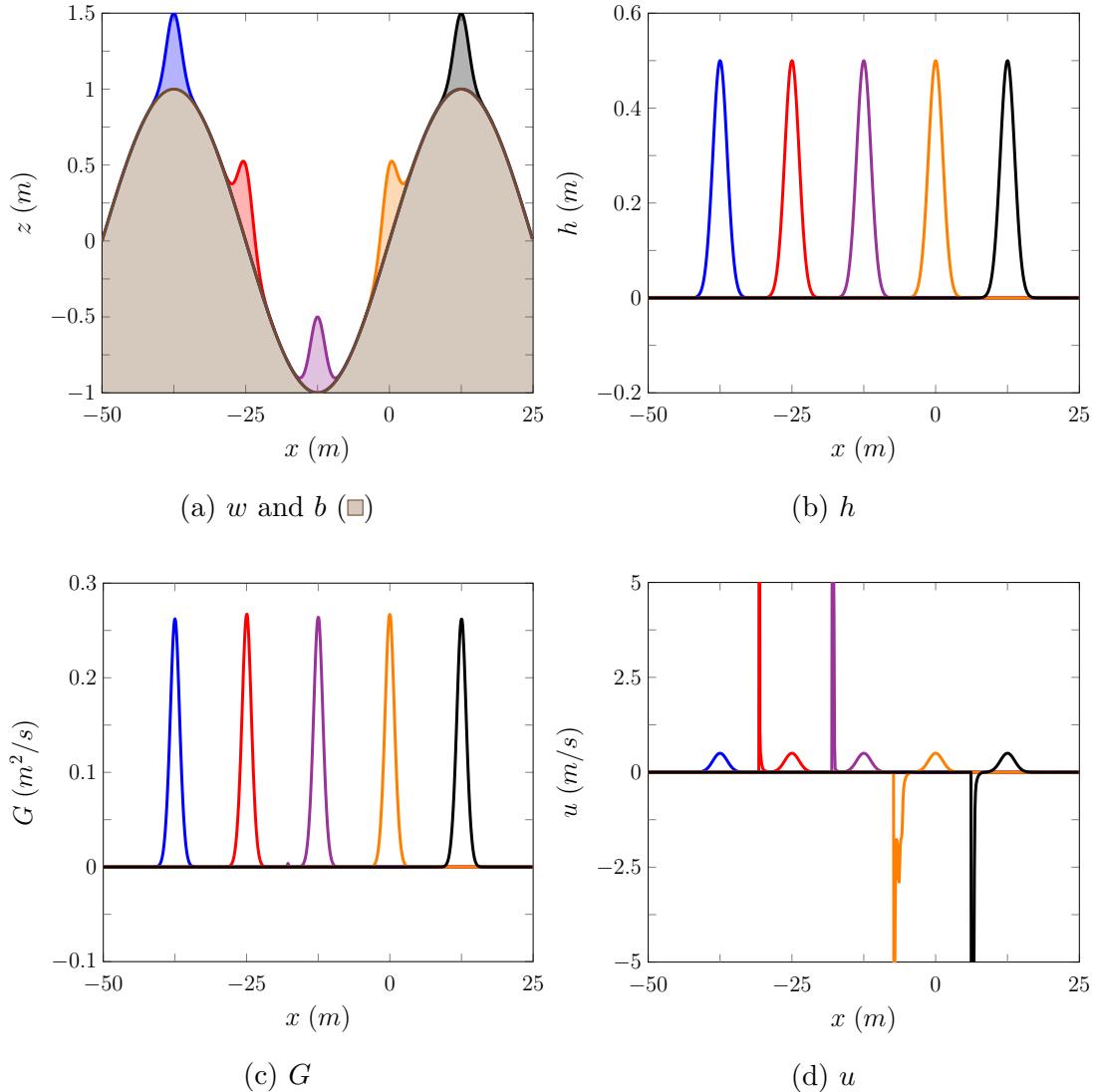


Figure 5.12: Numerical solutions for w , b , h , G and u produced by FDVM₂ with $\Delta x = 100/2^{10}m$ at $t = 0s$ (— / ■), $2.5s$ (— / □), $5.0s$ (— / ▨), $7.5s$ (— / ▤), $10.0s$ (— / ▨) to the dry bed forced solution problem, where $a_0 = 0m$.

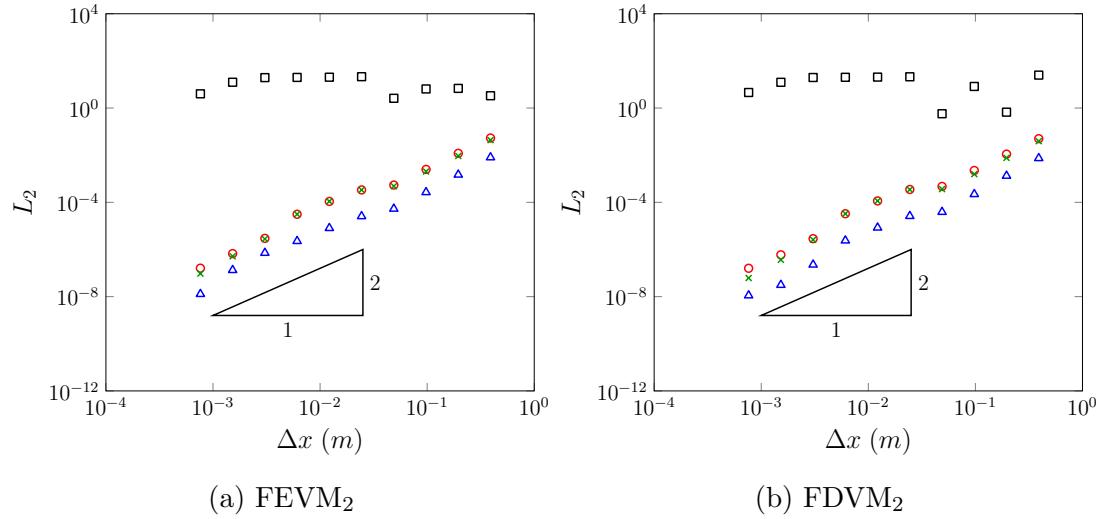


Figure 5.13: Convergence as measured by the L_2 norm against Δx for h (Δ), u (\square), uh (\times) and G (\circ) for the dry bed forced solution problem for FEVM₂ and FDVM₂ at $t = 10s$.

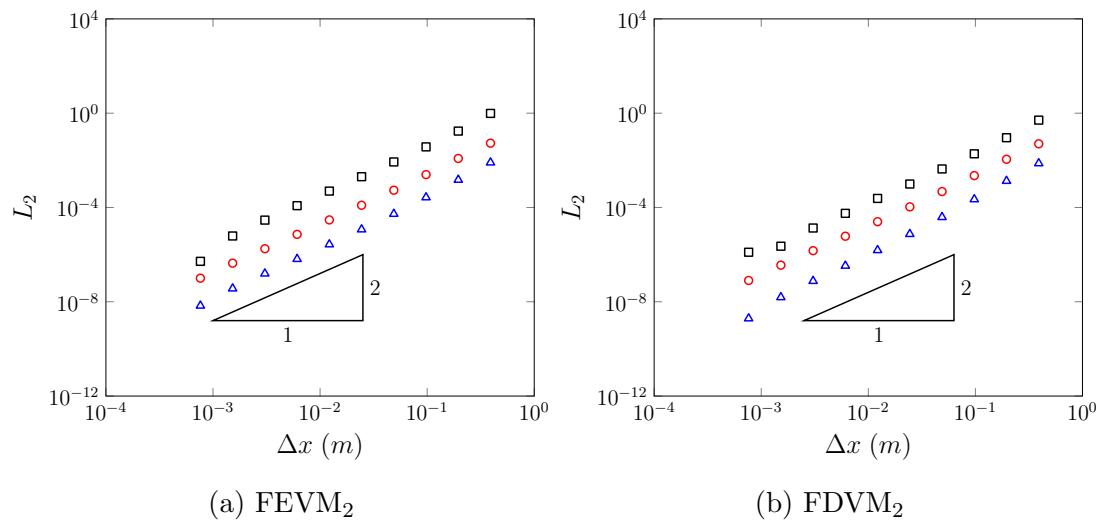


Figure 5.14: Convergence for regions where $h > 10^{-3}m$ as measured by the L_2 norm against Δx for h (Δ), u (\square) and G (\circ) for the dry bed forced solution problem for FEVM₂ and FDVM₂ at $t = 10s$.

Chapter 6

Experimental Validation

The second-order Finite Difference Volume Method (FDVM) denoted FDVM₂ and the second-order Finite Element Volume Method (FEVM) termed FEVM₂ are experimentally validated by comparing their numerical solutions to experimental data. A description of FEVM₂ was provided in Chapter 3 while FDVM₂ was described by Zoppou et al. [15]. Note that the dry-bed handling technique outlined in Chapter 3 for FEVM₂ was also applied to FDVM₂.

The chosen experiments allow the methods capability to model a variety of physical situations to be tested. The situations tested include flows with steep gradients in the free surface, strong dispersive waves interacting with varying bathymetry, shoaling and wave-breaking and finally the inundation of a sloping beach. Therefore, the ability of FEVM₂ and FDVM₂ to robustly reproduce all the experimental results well strongly demonstrates their capability to model a variety of physical situations.

6.1 Evolution of a Negative Rectangular Wave

A series of experiments studying the evolution of a negative rectangular wave in the free surface were conducted by Hammack and Segur [16]. These experiments were performed in a wave tank that was 0.394m wide, 31.6m long and 0.61m high. The rectangular negative waves were generated using a piston 0.61m long with its left edge against the wave tank wall. The 0.1m deep water is initially stationary with a horizontal free surface and the piston in the up position. The experiment begins when the piston suddenly moves down. This creates a sudden negative rectangular wave on the water surface generating a dispersive wave train that is recorded at Wave Gauges (WG) located 0m, 5m, 10m, 15m and 20m away

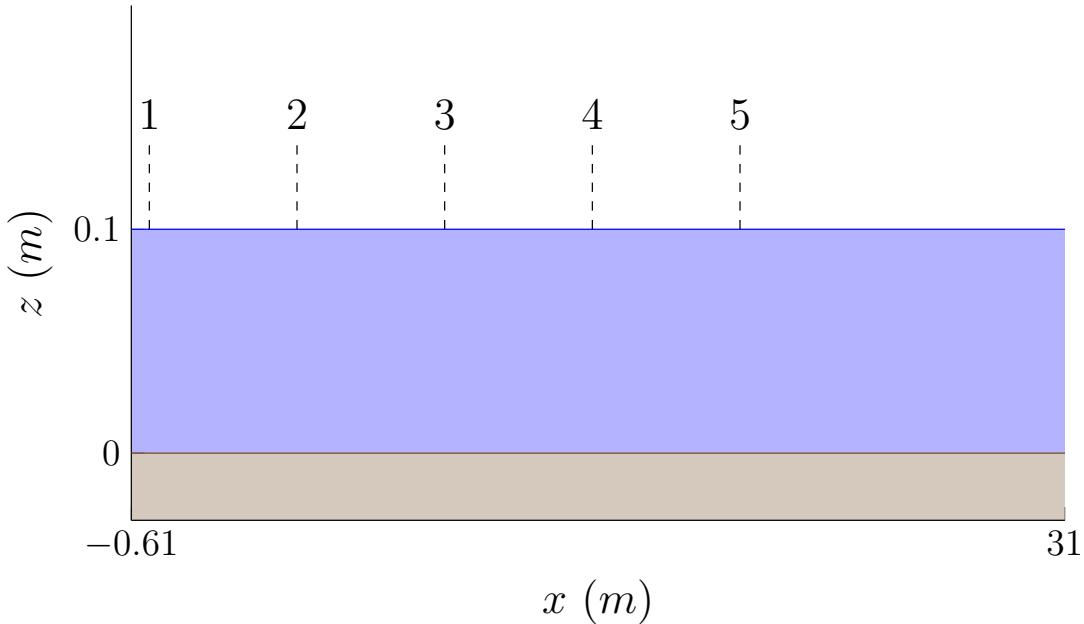


Figure 6.1: Diagram showing a longitudinal section of the wave tank for the negative rectangular wave experiments with the water (blue), the bed (brown) and the WG locations marked.

from the right edge of the piston. A diagram of the longitudinal section of the wave tank with the WG locations marked is given in Figure 6.1.

These experiments provide a good benchmark for the capability of the numerical method to accurately model problems with steep gradients in the free surface. However, these experiments are affected by bed friction, viscosity and the inability of the piston and water to move vertically instantaneously and slip free. The Serre equations do not contain viscosity and bed friction was neglected in this thesis. Additionally, discontinuous initial conditions are used to model the negative rectangular wave. Therefore, it is expected that the numerical solutions of the Serre equations will produce more oscillations in the dispersive wave train than are observed experimentally [18].

Hammack and Segur [16] report the results for two different initial negative wave amplitudes $0.01m$ and $0.03m$, resulting in the non-linearity parameters $\epsilon = 0.1$ and $\epsilon = 0.3$ respectively. Since these non-linearity parameters are relatively small there was no breaking of waves throughout the experiment.

These experiments were modelled numerically using the reflected problem, with the left wall of the wave tank as the axis of symmetry. In the numerical experiments the domain is $[-60m, 60m]$ and the experiment is run for $50s$ with

$g = 9.81 \text{ m/s}^2$. The spatial resolution used was $\Delta x = 0.01 \text{ m}$ while the CFL condition (3.23) was satisfied by using $\Delta t = 0.5\Delta x/\sqrt{g0.1}$. The limiting parameter $\theta = 1.2$ was used in the reconstruction (3.2) in FEVM₂ and FDVM₂.

The numerical results of FDVM₂ for these experiments were published by Zoppou et al. [15]. These results have been extended in this thesis with the inclusion of the conservation error of h , G and uh in the numerical simulation of both experiments.

6.1.1 Results for the 0.01m Negative Rectangular Wave

Comparisons of the numerical and experimental WG data for the 0.01m negative rectangular wave are displayed in Figures 6.2 and 6.3 for FEVM₂ and FDVM₂ respectively. The data is presented in the same dimensionless scales as reported in the original paper [16]. Furthermore, Tables 6.1 and 6.2 display the completely numerically calculated error in conservation C^* .

The numerical solutions have reproduced the experimental results well; particularly for the front of the dispersive wave train. While the total amounts of h , G , uh and \mathcal{H} have been conserved by the numerical methods with little error.

The numerical solutions produce larger and consequently faster waves and generate oscillations in the dispersive wave train which are not present in the experimental data of WG1. Moreover, as expected the methods produce more oscillations than were observed experimentally. These discrepancies can be attributed to the lack of viscosity and the omission of bed friction for the Serre equations in this thesis (2.6). Furthermore, it is highly likely the experiment produced some smooth approximation to a discontinuous jump in the water depth with the down-stroke of the piston. Such a smoothing of the initial conditions will significantly attenuate the high frequency waves in the generated dispersive wave train [18]. Given these differences the numerical methods replicated the experimental behaviour well.

These numerical solutions compare well to those of Zoppou et al. [15] for FDVM₂ with $\Delta x = 0.005 \text{ m}$, $\Delta t = 0.2\Delta x/\sqrt{g0.1} \text{ s}$ and $\theta = 1$. The differences in the resolution and the value of θ had little impact on the WG results. However, the more diffusive value of the limiting parameter θ smoothed the initial discontinuity which negatively impacted the conservation of \mathcal{H} reported by Zoppou et al. [15].

Both FEVM₂ and FDVM₂ have produced indistinguishable results at this scale and have demonstrated very good conservation of all the quantities as demon-

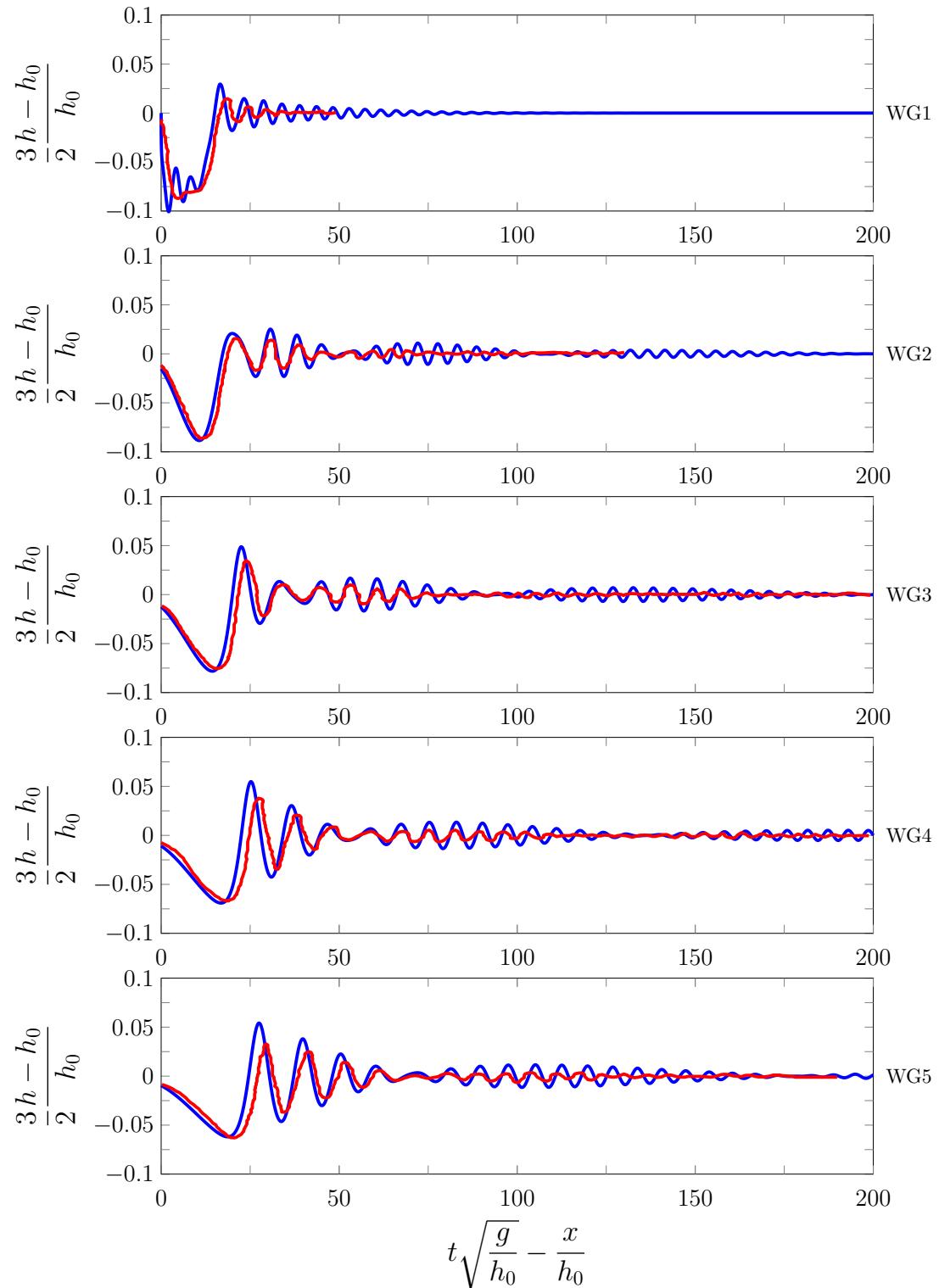


Figure 6.2: Time series of the experimental WG data (—) and numerical results (—) of FEVM₂ for the 0.01m negative rectangular wave.

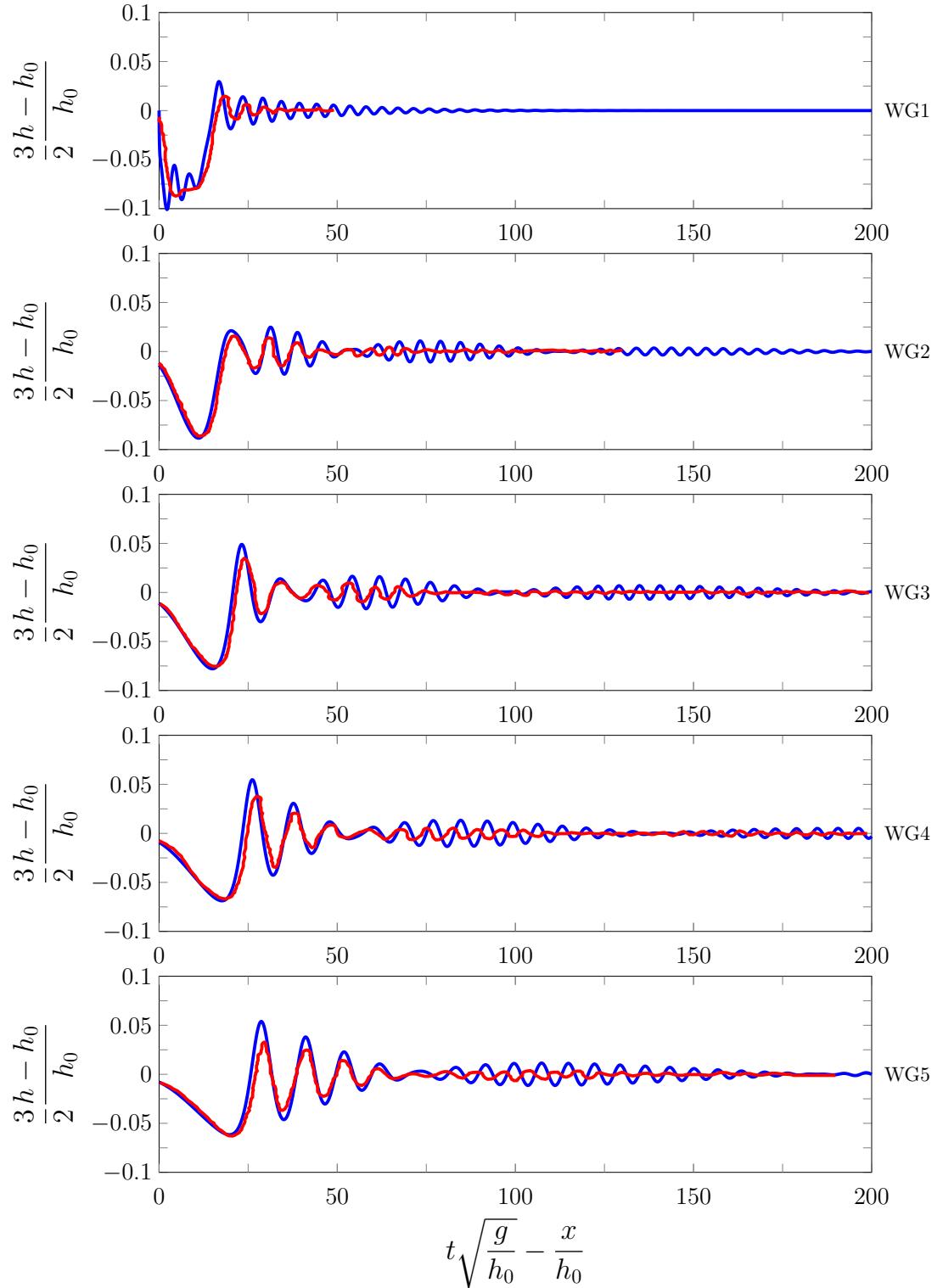


Figure 6.3: Time series of the experimental WG data (—) and numerical results (—) of FDVM₂ for the 0.01m negative rectangular wave.

Quantity	$\mathcal{C}^*(\mathbf{q}^0)$	$\mathcal{C}^*(\mathbf{q}^*)$	$C^*(\mathbf{q}^0, \mathbf{q}^*)$
h	11.9888	11.9888	0
uh	0	7.44×10^{-18}	7.44×10^{-18}
G	0	1.56×10^{-18}	1.56×10^{-18}
\mathcal{H}	5.8751	5.8751	5.70×10^{-6}

Table 6.1: Initial and final total amounts and the conservation error for all conserved quantities for the numerical solution of FEVM₂ for the 0.01m negative rectangular wave.

Quantity	$\mathcal{C}^*(\mathbf{q}^0)$	$\mathcal{C}^*(\mathbf{q}^*)$	$C^*(\mathbf{q}^0, \mathbf{q}^*)$
h	11.9888	11.9888	0
uh	0	-1.19×10^{-17}	-1.19×10^{-17}
G	0	-8.05×10^{-18}	-8.05×10^{-18}
\mathcal{H}	5.8751	5.8751	6.27×10^{-6}

Table 6.2: Initial and final total amounts and the conservation error for all conserved quantities for the numerical solution of FDVM₂ for the 0.01m negative rectangular wave.

strated in Tables 6.1 and 6.2. Given the extensive examination of several numerical schemes for steep gradient problems [18], this indicates that these solutions are representative of true solutions of the Serre equations. However, these results do not demonstrate the superiority of one of these methods over the other.

6.1.2 Results for the 0.03m Negative Rectangular Wave

The WG data for the numerical and experimental results for the evolution of the 0.03m negative rectangular wave are displayed in Figures 6.4 and 6.5 for FEVM₂ and FDVM₂ respectively. These results are reported using the same dimensionless scales the original paper [16]. The completely numerically calculated conservation error C^* for all of the conserved quantities are given in Tables 6.3 and 6.4 for FEVM₂ and FDVM₂ respectively.

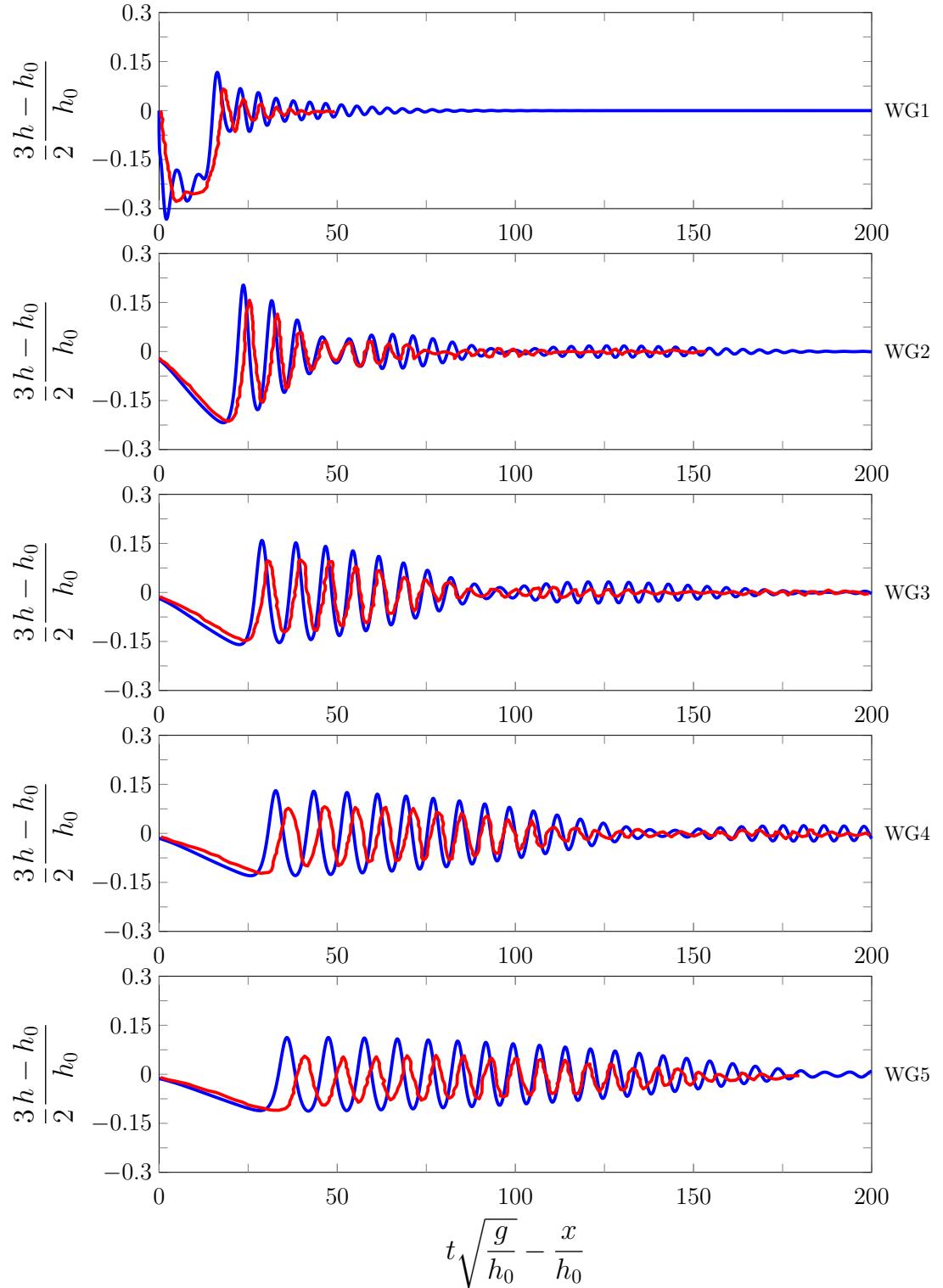


Figure 6.4: Time series of the experimental WG data (—) and numerical results (—) of FEVM₂ for the 0.03m negative rectangular wave.

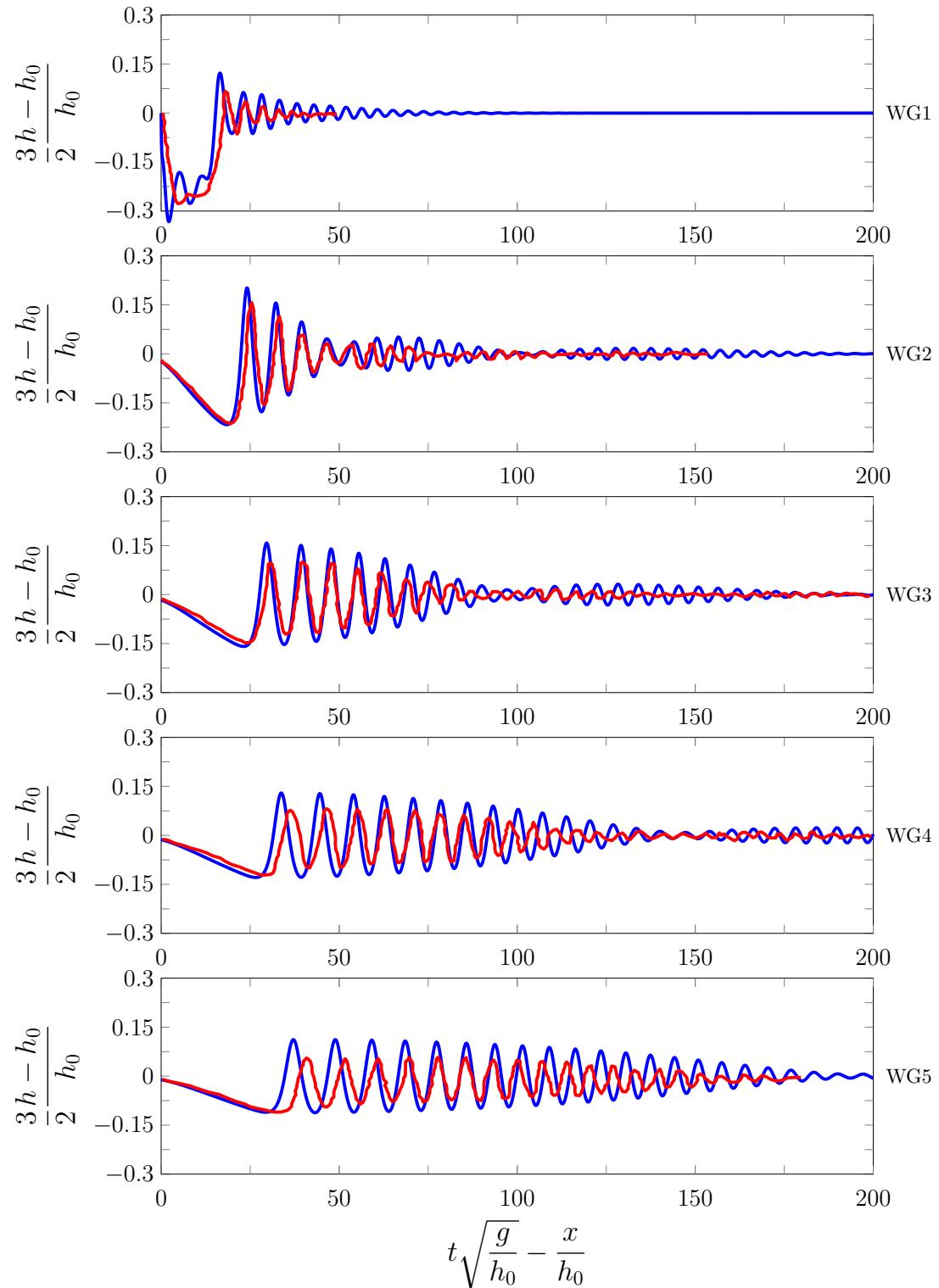


Figure 6.5: Time series of the experimental WG data (—) and numerical results (—) of FDVM₂ for the 0.03m negative rectangular wave.

Both methods reproduce the overall behaviour of this experiment very well. Since this rectangular wave is deeper than the $0.01m$ negative rectangular wave experiment, this experiment provides a more rigorous test of the numerical methods. However, the increased depth also strengthens the effects that produced the discrepancy between the experimental results and the numerical solutions of the Serre equations in the $0.01m$ negative rectangular wave experiment. Hence, the differences observed in the amplitude and speed of the generated waves between the numerical and experimental results.

Since the negative rectangular wave is larger than in the $0.01m$ negative rectangular wave experiment, the numerical methods have a larger error in conservation for all the quantities, except h which is conserved exactly. For G and uh the conservation errors are around machine precision and can be disregarded, thus only the error in conservation of \mathcal{H} was significantly increased by a larger negative wave. However, the error in conservation of \mathcal{H} was small and therefore all quantities were well conserved by the numerical methods.

These results compare well with those by Zoppou et al. [15] for FDVM₂ with $\Delta x = 0.005m$, $\Delta t = 0.2\Delta t/\sqrt{g0.1}s$ and $\theta = 1$. The WG data of those numerical solutions and the solutions displayed here are indistinguishable. The error in conservation in \mathcal{H} is very similar as well, although for the numerical solutions displayed here the numerical grid is coarser. The poorer than expected conservation of \mathcal{H} given the finer grid of the numerical solution is caused by the limiting parameter $\theta = 1$ diffusing the discontinuous jump [18].

Conserving \mathcal{H} is difficult for numerical methods solving steep gradient problems [18], especially when two steep gradients interact with one another over short time spans. Increasing the resolution of the numerical method improves the conservation of \mathcal{H} as demonstrated for the analytic solutions. However, the results for the WG data for higher resolution solutions will be indistinguishable from the results presented here and so the provided solutions are sufficient.

These experiments have been replicated equally well by the numerical methods. Given the resolution, the error in conservation and the extensive study summarised in Chapter 2; these results demonstrate the accuracy of the numerical methods in the presence of steep gradients in the free surface.

Quantity	$\mathcal{C}^*(\mathbf{q}^0)$	$\mathcal{C}^*(\mathbf{q}^*)$	$C^*(\mathbf{q}^0, \mathbf{q}^*)$
h	11.9644	11.9644	0
uh	0	-7.75×10^{-17}	-7.75×10^{-17}
G	0	-3.33×10^{-16}	-3.33×10^{-16}
\mathcal{H}	5.8560	5.8552	1.24×10^{-4}

Table 6.3: Initial and final total amounts and the conservation error for all conserved quantities for the numerical solution of FEVM₂ for the 0.03m negative rectangular wave.

Quantity	$\mathcal{C}^*(\mathbf{q}^0)$	$\mathcal{C}^*(\mathbf{q}^*)$	$C^*(\mathbf{q}^0, \mathbf{q}^*)$
h	11.9644	11.9644	0
uh	0	-9.09×10^{-17}	-9.09×10^{-17}
G	0	-1.16×10^{-16}	-1.16×10^{-16}
\mathcal{H}	5.8560	5.8552	1.30×10^{-4}

Table 6.4: Initial and final total amounts and the conservation error for all conserved quantities for the numerical solution of FDVM₂ for the 0.03m negative rectangular wave.

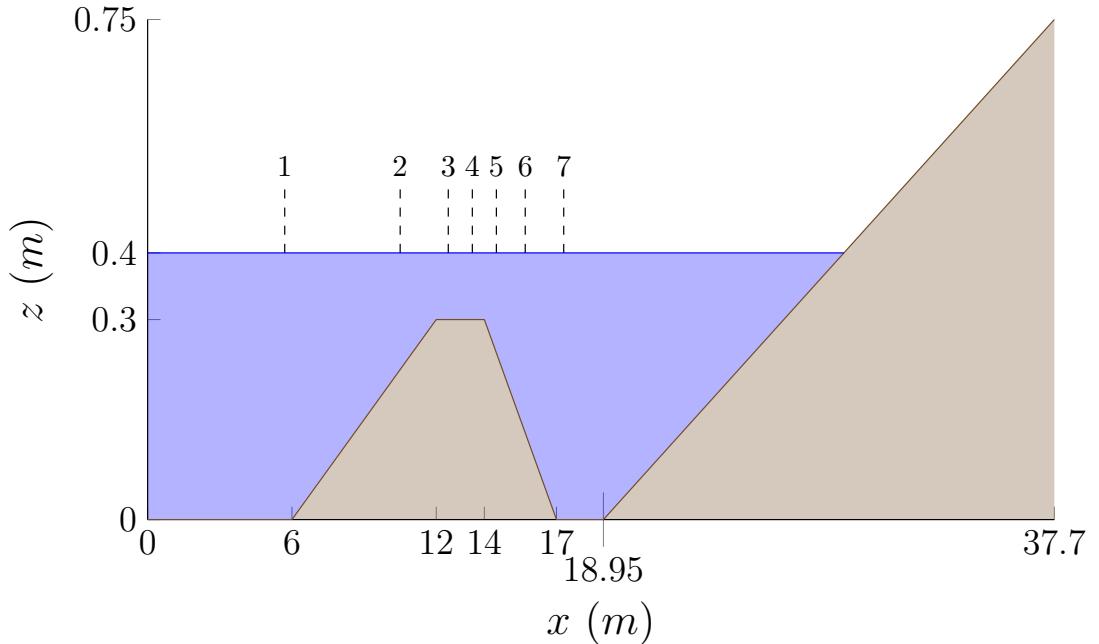


Figure 6.6: Diagram showing a longitudinal section of the wave tank for the periodic waves over a submerged bar experiments with the water (blue), the bed (brown) and the WG locations displayed.

6.2 Periodic Waves Over A Submerged Bar

Beji and Battjes conducted a series of experiments investigating the effect of submerged bars on the propagation of periodic waves [17, 60]. The behaviour observed in these experiments was mainly driven by the dispersion properties of the waves and their interaction with variations in bathymetry. Therefore, these experiments serve as a benchmark for the ability of the numerical schemes to accurately model the interaction of dispersive waves with variable bathymetry. In this thesis only the monochromatic wave experiments of Beji and Battjes [17] were simulated.

The experiments of Beji and Battjes [17] were conducted in a wave tank 37.7m long, 0.8m wide and 0.75m high. A diagram of the longitudinal section of the wave tank is given in Figure 6.6. There are seven WG at the following locations; 5.7m, 10.5m, 12.5m, 13.5m, 14.5m, 15.7m and 17.3m. Waves are generated from a piston-type wave maker located at 0m and travel on the initially still water 0.4m deep to the right, over the submerged trapezoidal bar and are absorbed by a sloping beach.

Two monochromatic non-breaking wave experiments were performed, a low

frequency one with a wavelength $\lambda \approx 3.69m$ and a period of $T = 2s$, and a high frequency one with $\lambda \approx 2.05m$ and a period of $T = 1.25s$. Both experiments had a wave amplitude of $0.01m$ and so both had the same small non-linearity parameter $\epsilon = 0.01/0.4 = 0.025$.

For the numerical solutions the spatial domain was $[5.7m, 150m]$ with $\Delta x = 0.1/2^4m \approx 0.0063m$ and $\Delta t = Sp/2^5s \approx 0.0012s$ where $Sp = 0.039s$ is the experimental sampling period. These Δx and Δt values satisfy the CFL condition, (3.23). In the numerical experiments only the submerged trapezoidal bar is present, and the sloping beach is replaced with a very long horizontal bed that ensures that the enforced Dirichlet boundary conditions at the downstream boundary do not affect the experimental results.

To simulate incoming waves at the upstream boundary WG1 and linear extrapolation was used to calculate h in the left ghost cell and thus produce incoming waves. The velocity distribution over the left ghost cell was calculated from the height values in the same way as Beji and Battjes [17] using

$$u(x, t) = \frac{\omega}{k} \frac{h(x, t) - h_0}{h(x, t)}$$

where ω is the angular frequency of the linearised Serre equations (2.9) and k is the wavenumber of the wave. Finally the boundary conditions for G were calculated using the boundary values for h and u .

The results of FDVM₂ with the same resolution and limiting parameter were published by Zoppou et al. [15]. These results have been reproduced here to serve as a comparison for the results of FEVM₂.

6.2.1 Low Frequency Results

A comparison of the wave heights η of the experimental and numerical results are located in Figures 6.7 and 6.8 for FEVM₂ and Figures 6.9 and 6.10 for FDVM₂.

These results demonstrate the ability of the numerical methods to reproduce the experimental results, particularly for WG1 to WG5 where the best agreement between experimental and numerical results is observed. The results for these WG validate the numerical schemes for simulating the shoaling of dispersive waves as these WG are all located on the windward side of the submerged bar.

The numerical results for WG6 and WG7 on the leeward side capture some of the wave behaviour but their agreement with the experimental results is worse.

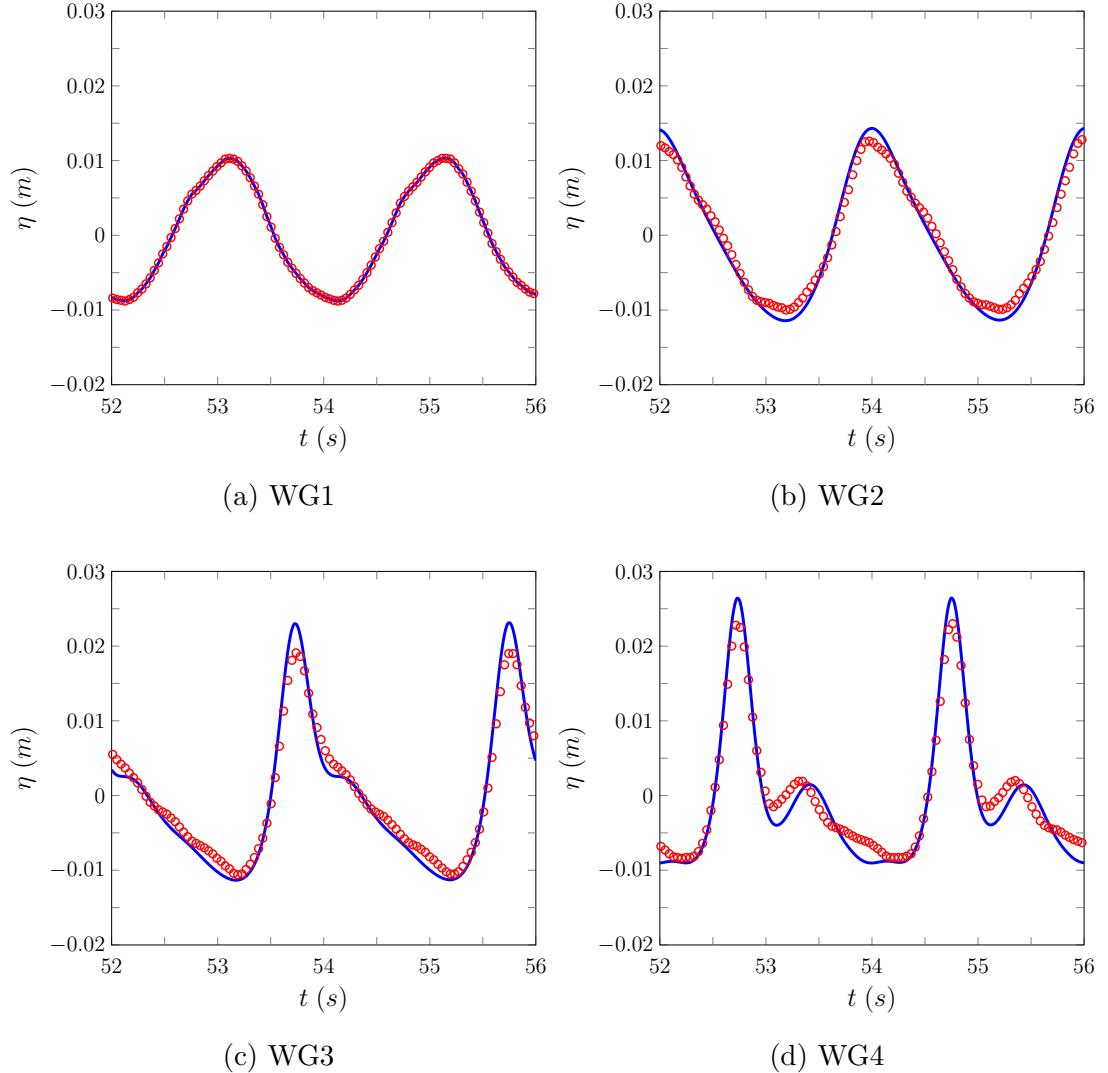


Figure 6.7: Time series of the wave heights η of the numerical results of FEVM₂ (—) and the experimental results (○) for WG1 - WG4 for the low frequency experiment.

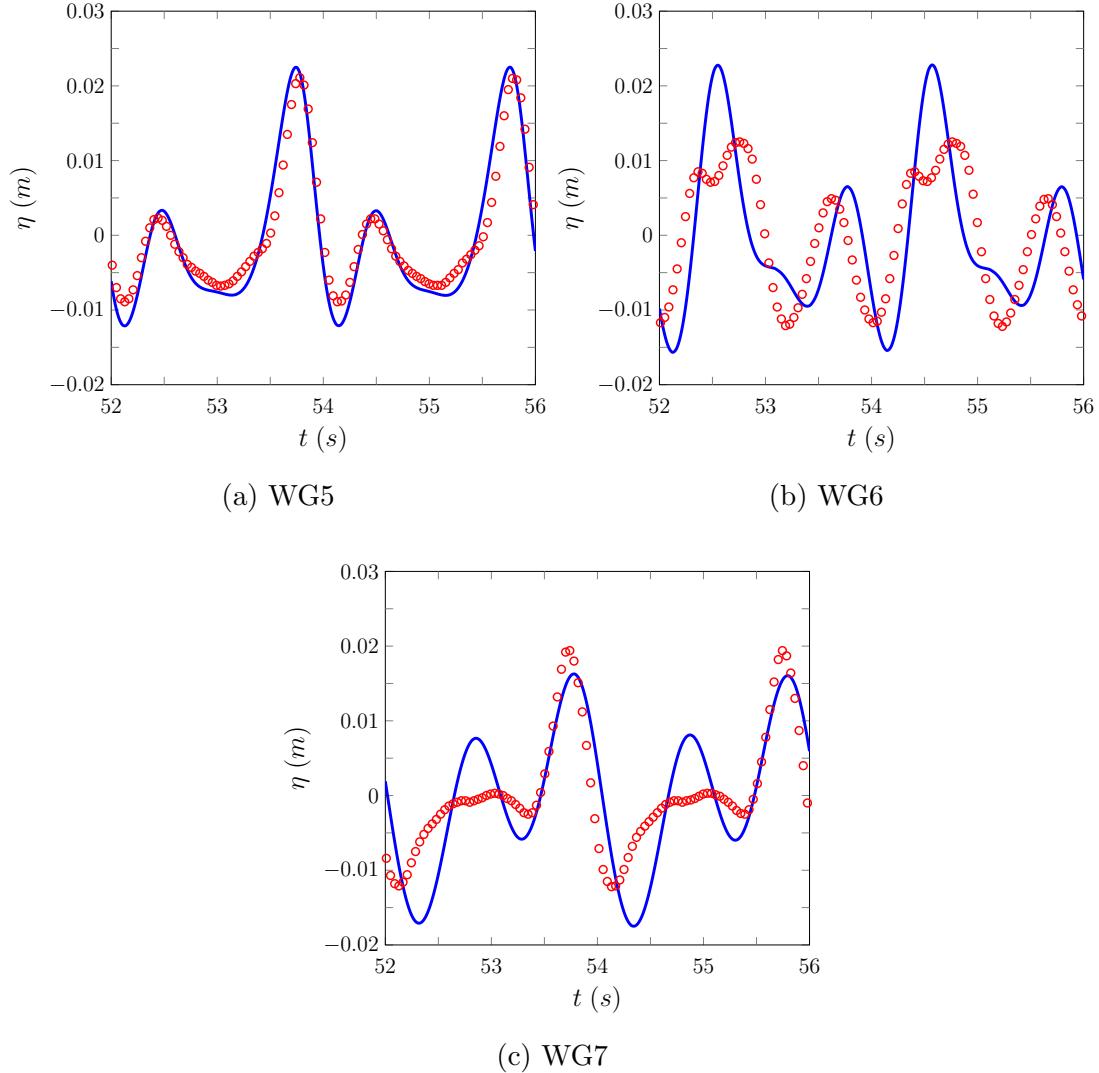


Figure 6.8: Time series of the wave heights η of the numerical results of FEVM₂ (—) and the experimental results (○) for WG5 - WG7 for the low frequency experiment.

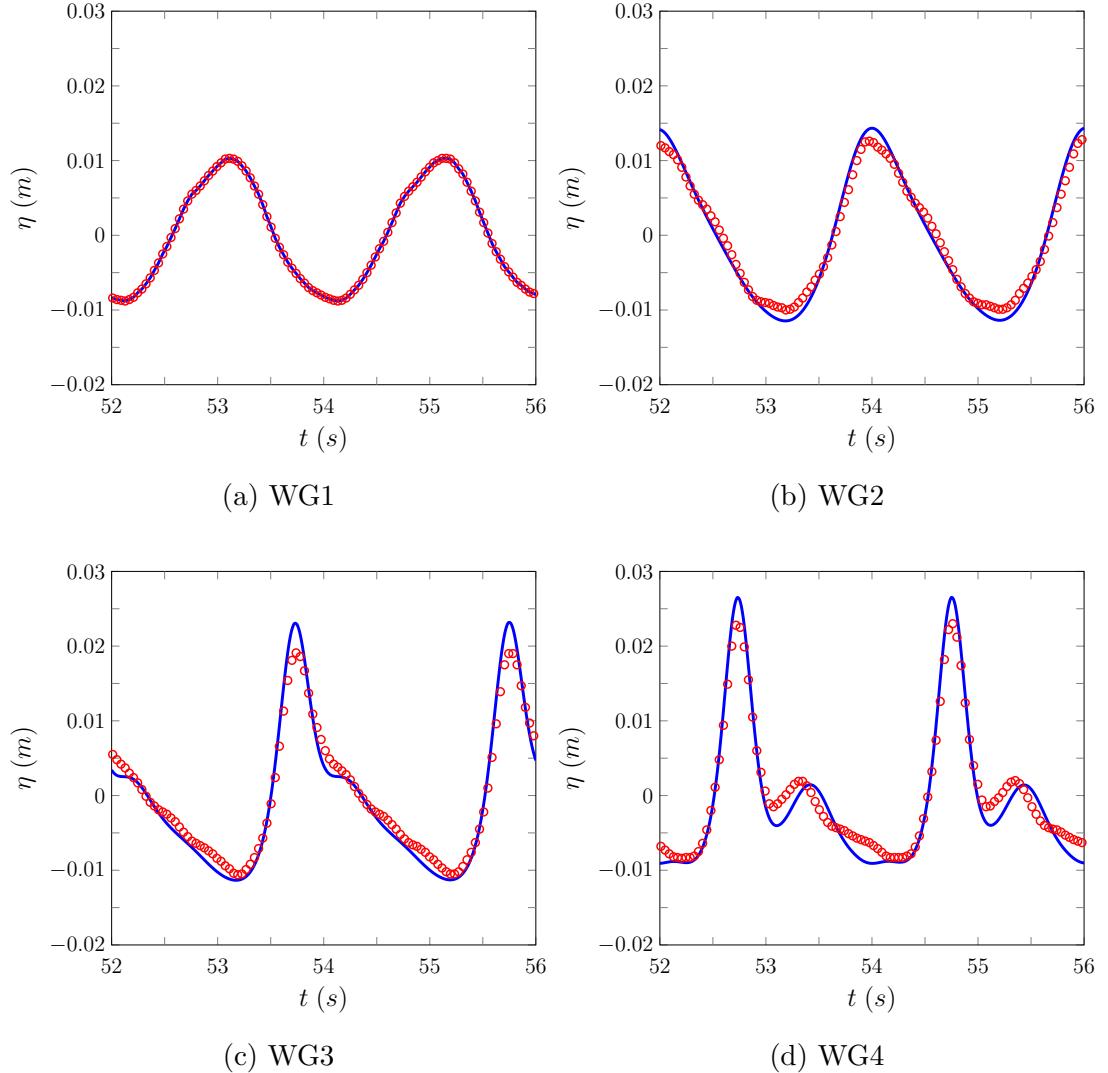


Figure 6.9: Time series of the wave heights η of the numerical results of FDVM₂ (—) and the experimental results (○) for WG1 - WG4 for the low frequency experiment.

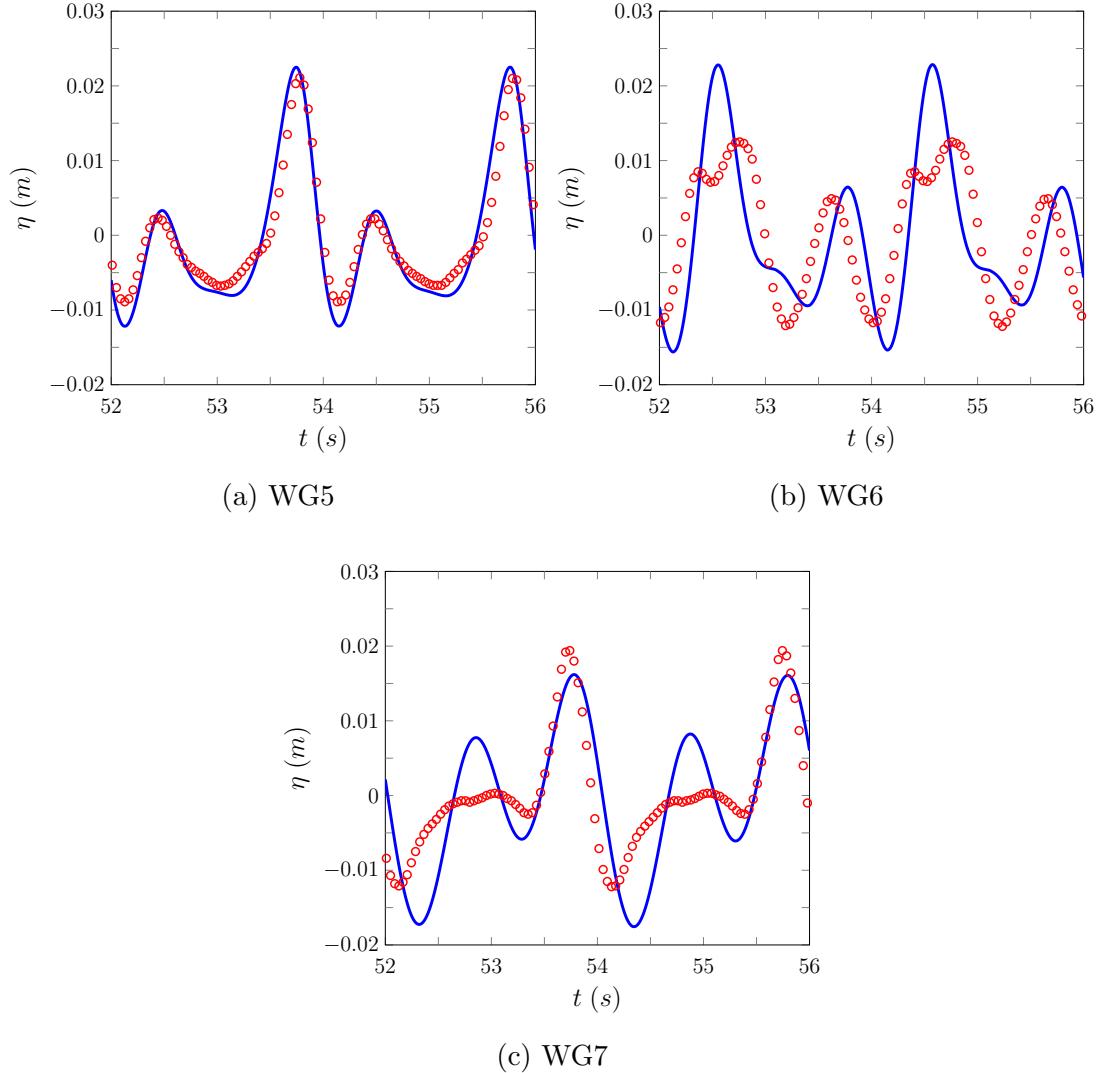


Figure 6.10: Time series of the wave heights η of the numerical results of FDVM₂ (—) and the experimental results (○) for WG5 - WG7 for the low frequency experiment.

The inadequacy of the numerical results here appears to be caused by the discrepancy between the dispersion properties of the Serre equations and actual water waves [17, 61]. Although for WG2 to WG5 the dispersion properties of the Serre equations approximate the dispersion properties of actual water waves well enough to accurately reproduce the experimental results. Dispersion is vital to reproducing these experimental results as the non-dispersive SWWE are not capable of accurately simulating this experiment [62].

Both FEVM₂ and FDVM₂ have accurately reproduced the experimental results, accounting for the shortcomings of the underlying Serre equations [17, 61]. However, both numerical schemes have produced indistinguishable results for all WG and so this experiment does not discriminate between these two methods.

6.2.2 High Frequency Results

The wave heights of the experimental and numerical results are given in Figures 6.11 and 6.12 for FEVM₂. While the wave height results for FDVM₂ are given in Figures 6.13 and 6.14.

As in the low frequency experiment it is observed that the numerical results perform well on the windward side of the slope for WG1 to WG4 but perform poorly for the leeward side of the slope for WG5 to WG7. Although, in the high frequency experiment the divergence between the numerical and experimental results occurs earlier than in the low frequency experiment, as evident in the results of WG5. As in the low frequency example this is caused by the difference in the dispersion relations of the Serre equations and the linear theory for water waves [17, 61]. Because the difference between the dispersion relation of the Serre equations and water waves is largest for higher frequency waves [24] the earlier divergence between experimental and numerical results is expected.

The numerical results for FDVM₂ and FEVM₂ agree well with other numerical results for weakly dispersive equations for the simulation of periodic waves over a submerged bar in the literature [17, 28, 61, 63]. Therefore, without changing the underlying partial differential equations, the numerical methods perform as well as other numerical schemes in the literature at recreating the experimental results of Beji and Battjes [17]. Finally, as in the low frequency experiment FEVM₂ and FDVM₂ produce indistinguishable results for all WG and so this benchmark does not discriminate between these two methods.

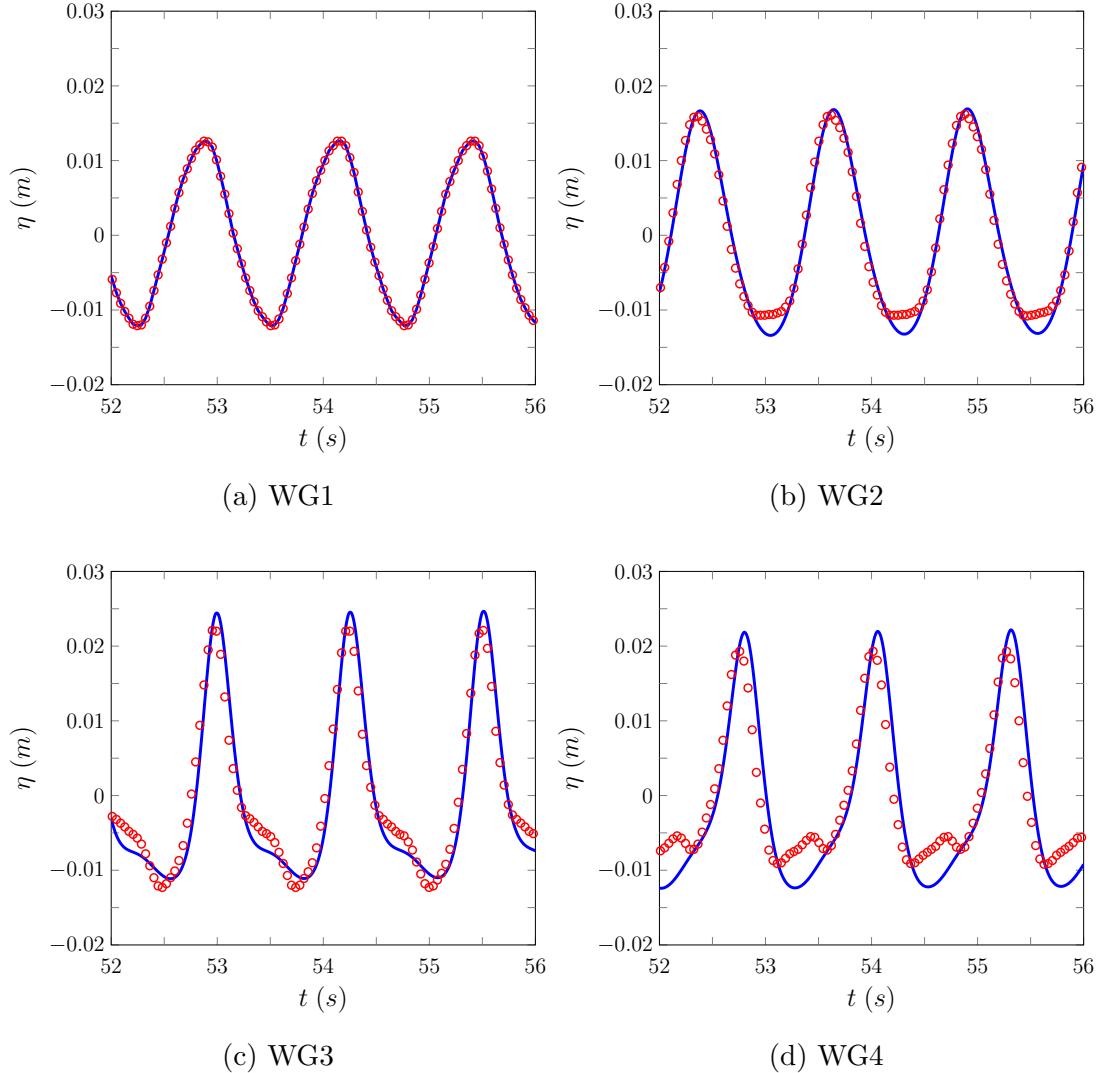


Figure 6.11: Time series of the wave heights η of the numerical results of FEVM₂ (—) and the experimental results (○) for WG1 - WG4 for the high frequency experiment.

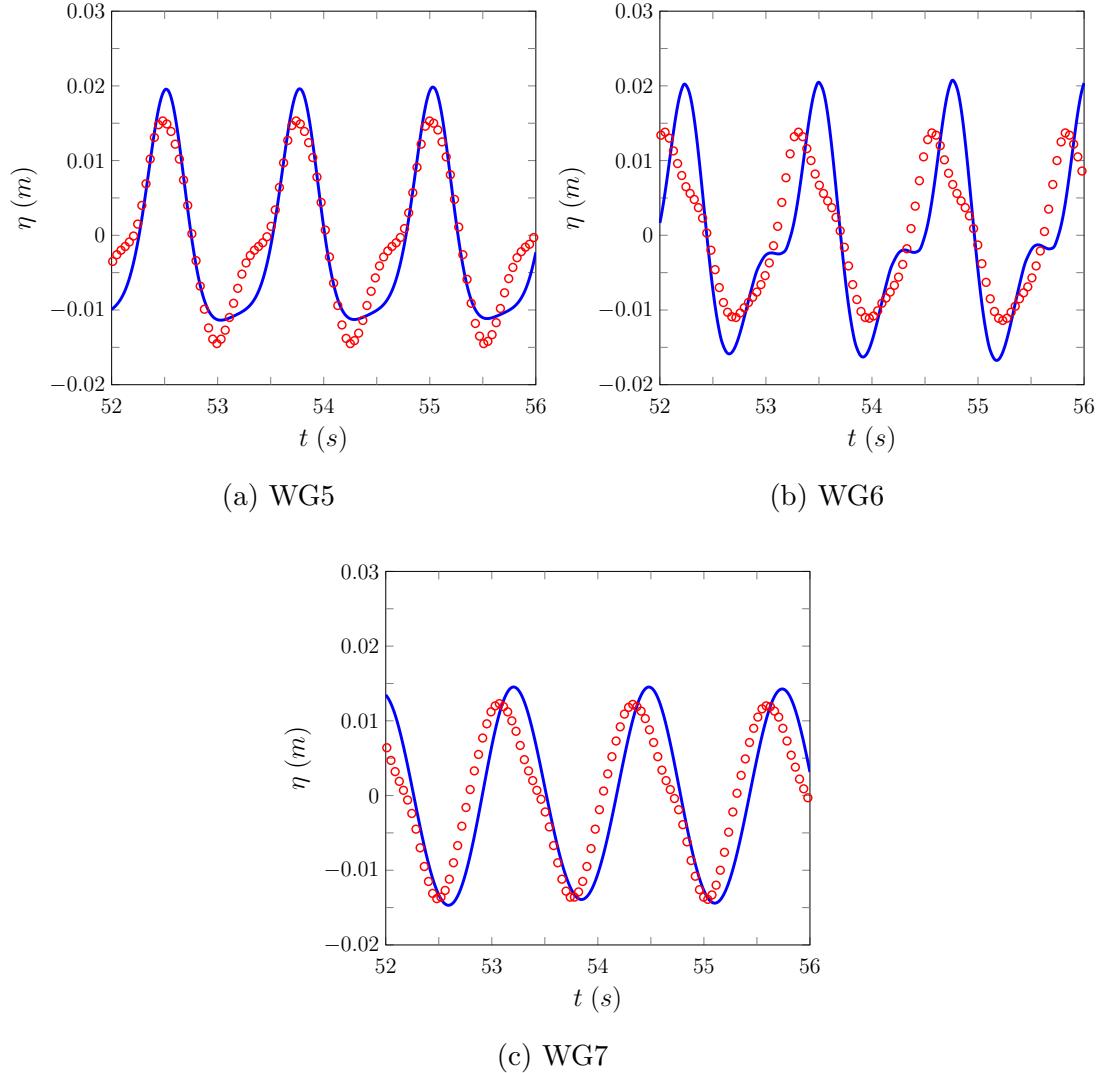


Figure 6.12: Time series of the wave heights η of the numerical results of FEVM₂ (—) and the experimental results (○) for WG5 - WG7 for the high frequency experiment.

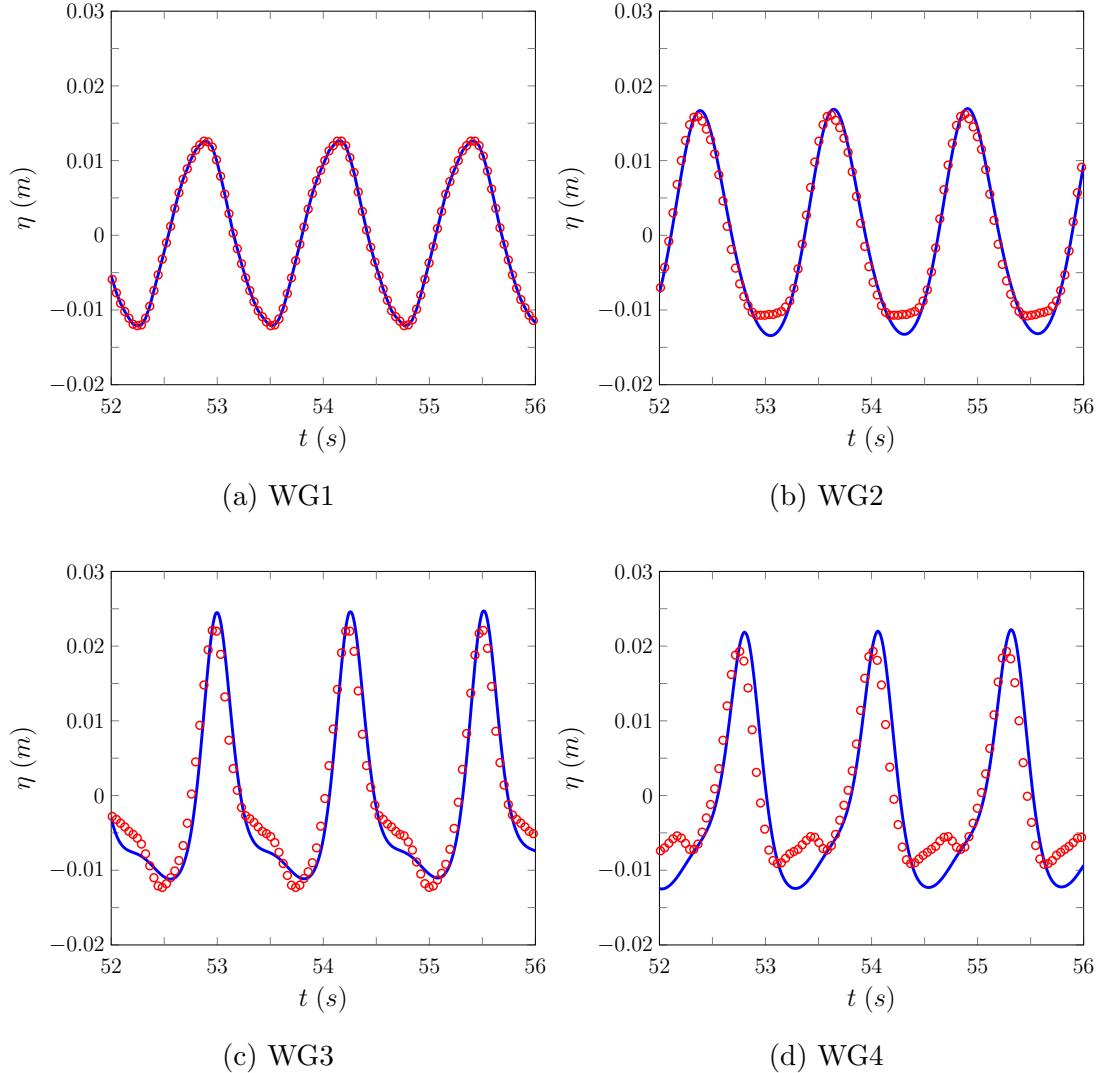


Figure 6.13: Time series of the wave heights η of the numerical results of FDVM₂ (—) and the experimental results (○) for WG1 - WG4 for the high frequency experiment.

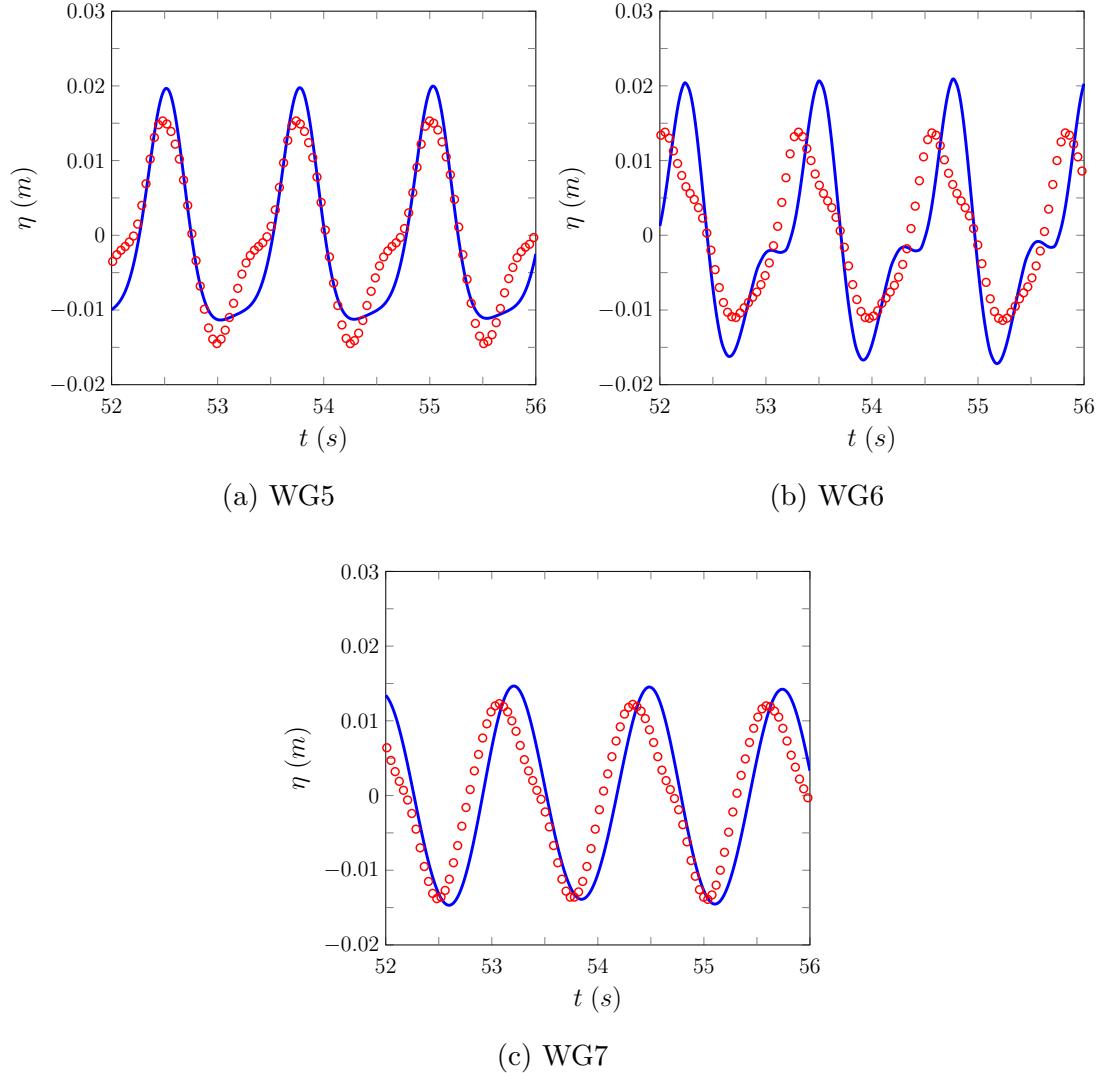


Figure 6.14: Time series of the wave heights η of the numerical results of FDVM₂ (—) and the experimental results (○) for WG5 - WG7 for the high frequency experiment.

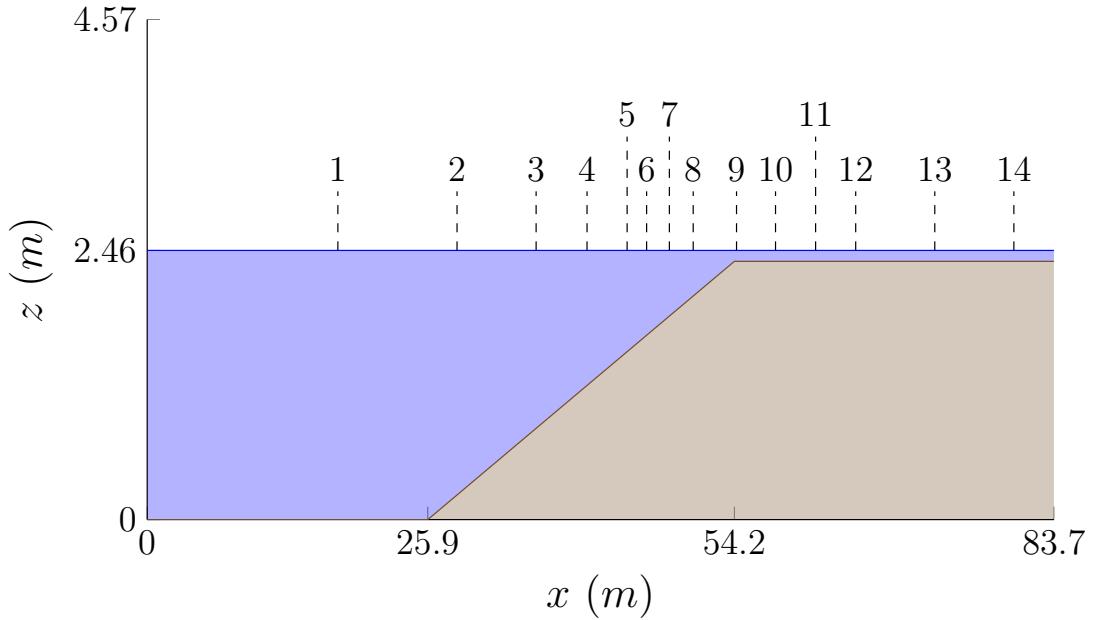


Figure 6.15: Diagram showing a longitudinal section of the wave tank for the solitary wave over a fringing reef experiment with the water (■), the bed (□) and the WG locations marked.

6.3 Solitary Wave Over a Fringing Reef

To study the evolution of waves on fringing reefs a series of experiments were conducted by Roeber [64]. These experiments were performed in a wave tank 3.66m wide, 83.7m long and 4.57m high with a removable bed that allowed for the wide range of experiments reported by Roeber [64]. The experiment with the bathymetry displayed in Figure 6.15 was numerically simulated in this thesis. In this experiment a solitary wave is generated from the wave maker at 0m and is recorded at the WG located 17.6m, 28.6m, 35.9m, 40.6m, 44.3m, 46.1m, 48.2m, 50.4m, 54.4m, 58.0m, 61.7m, 65.4m, 72.7m and 80.0m downstream of the wave maker.

This experiment investigates the behaviour of a wave with high non-linearity $\epsilon \approx 1.23/2.46 = 0.5$ as it shoals over a linear bed into a very shallow body of water with a depth of only 0.1m. Given the high non-linearity of this wave, it is not surprising that as it shoals it becomes a plunging breaker by $t \approx 32s$ with an elliptical air cavity observed at $t \approx 33s$ [64]. As with other depth averaged equations, the Serre equations are only appropriate up to wave-breaking so this experiment is not an entirely appropriate test of the numerical methods, particularly after $t = 32s$.

This experiment was numerically modelled on the domain $[17.6m, 400m]$ and was run until $t = 60s$ after which the reflections from the downstream end of the tank significantly effect the experiment results. The beginning of the domain was chosen so that WG1 could be used as the left boundary condition by employing the same incoming wave technique as in Section 6.2 for the periodic waves over a submerged bar experiment. The spatial resolution was $\Delta x = 0.025m$ and the temporal resolution was $\Delta t = Sp/8s = 0.0025$ where $Sp = 0.02s$ was the sampling period of the WG. The employed spatial and temporal resolutions satisfy the CFL condition (3.23) for this experiment. The right edge of the domain used Dirichlet boundary conditions, since the domain was large no effects from the downstream boundary were observed throughout the numerical simulation.

The WG results comparing the numerical and experimental data are displayed in Figures 6.16, 6.17 and 6.18 for FEVM₂ and 6.19 and 6.20 for FDVM₂.

Both methods accurately reproduce the shoaling of the solitary wave, particularly in WG1 through WG8 which record the wave before breaking begins. The behaviour of the trailing waves is not as well replicated, with the numerical solutions overestimating the amplitude and speed of these waves. As in the negative rectangular wave experiment these observed differences can be attributed to the lack of bottom friction and viscosity in the Serre equations. Furthermore, the reflected wave can be observed in the experimental WG results and since the numerical simulation did not have reflective boundaries these waves are not replicated in the numerical solutions.

When breaking begins the numerical solutions perform much worse as expected; most notably FDVM₂ becomes unstable and the solution blows up. Because of this the numerical solution of FDVM₂ was only plotted until $t = 34s$. The instability is caused by the appearance of a very steep gradient with a large jump in the water depth compared to the depth of water that surrounds it as the wave breaks. The appearance of this instability implies that FDVM₂ is not robust and therefore, not well suited to large scale tsunami modelling on super computers with little oversight. The finite element based FEVM₂ does not suffer from these instability issues, but due to the limitations of the Serre equations does produce a dispersive wave train with amplitudes far exceeding the observed amplitudes in the experiment.

Even with the limitations of the underlying Serre equations the results for FEVM₂ are robust and accurately model the shoaling of the solitary wave. The greater robustness of FEVM₂ over FDVM₂ makes this a more promising method for further development. However, these results indicate the need to incorpo-

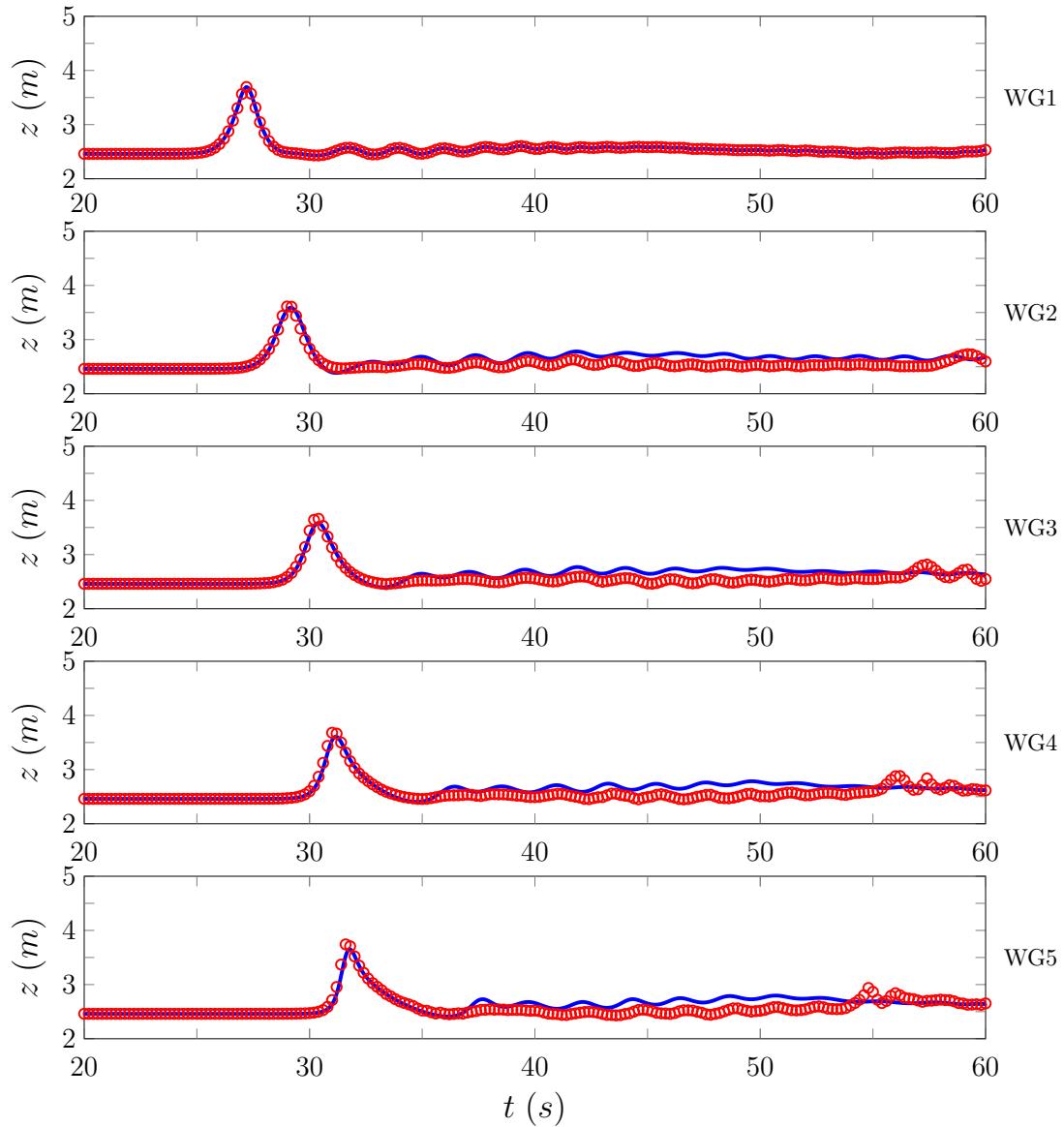


Figure 6.16: Time series of the experimental (○) and numerical (—) WG data produced by FEVM₂ for WG1 to WG5.

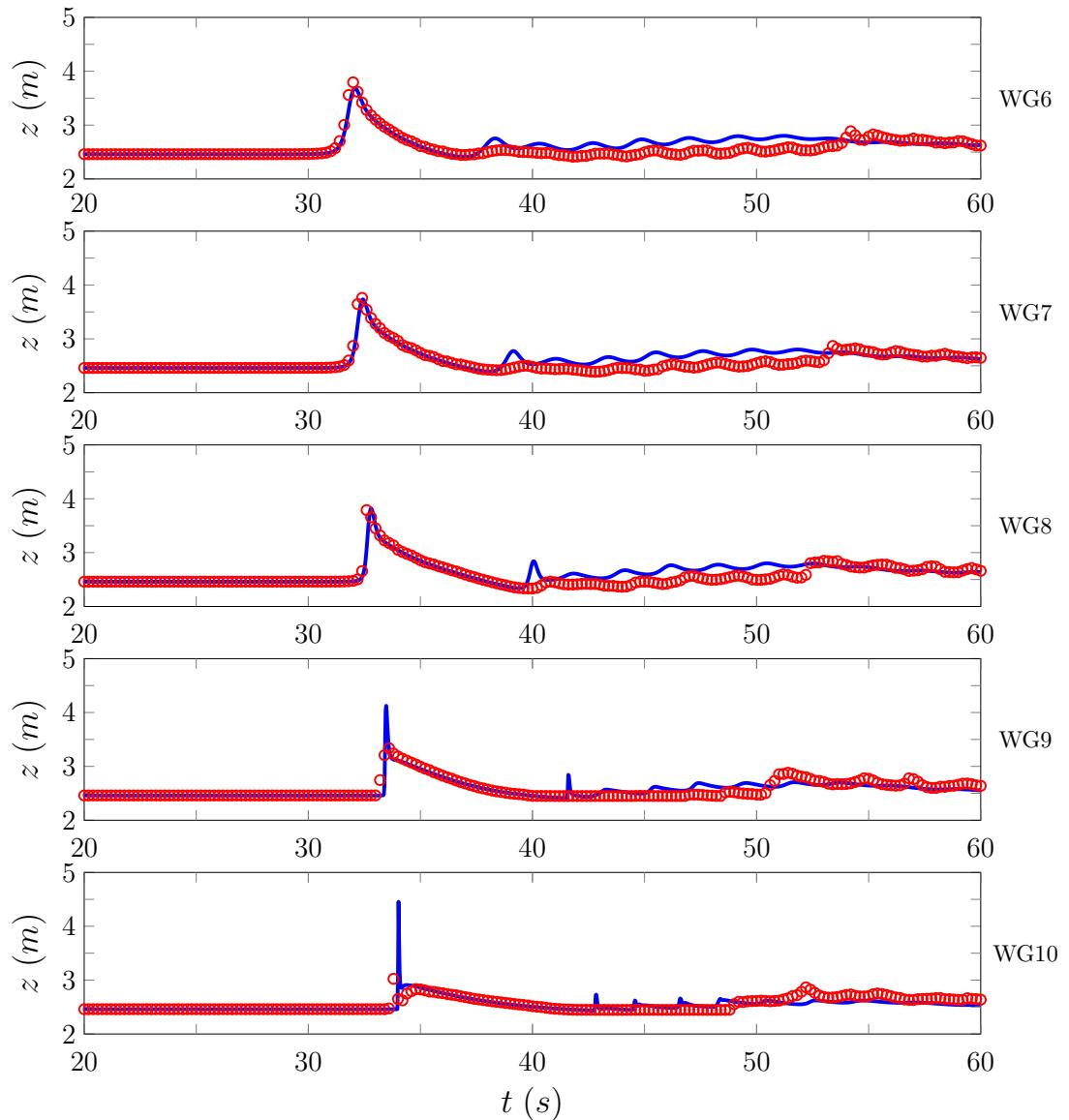


Figure 6.17: Time series of the experimental (○) and numerical (—) WG data produced by FEVM₂ for WG6 to WG10.

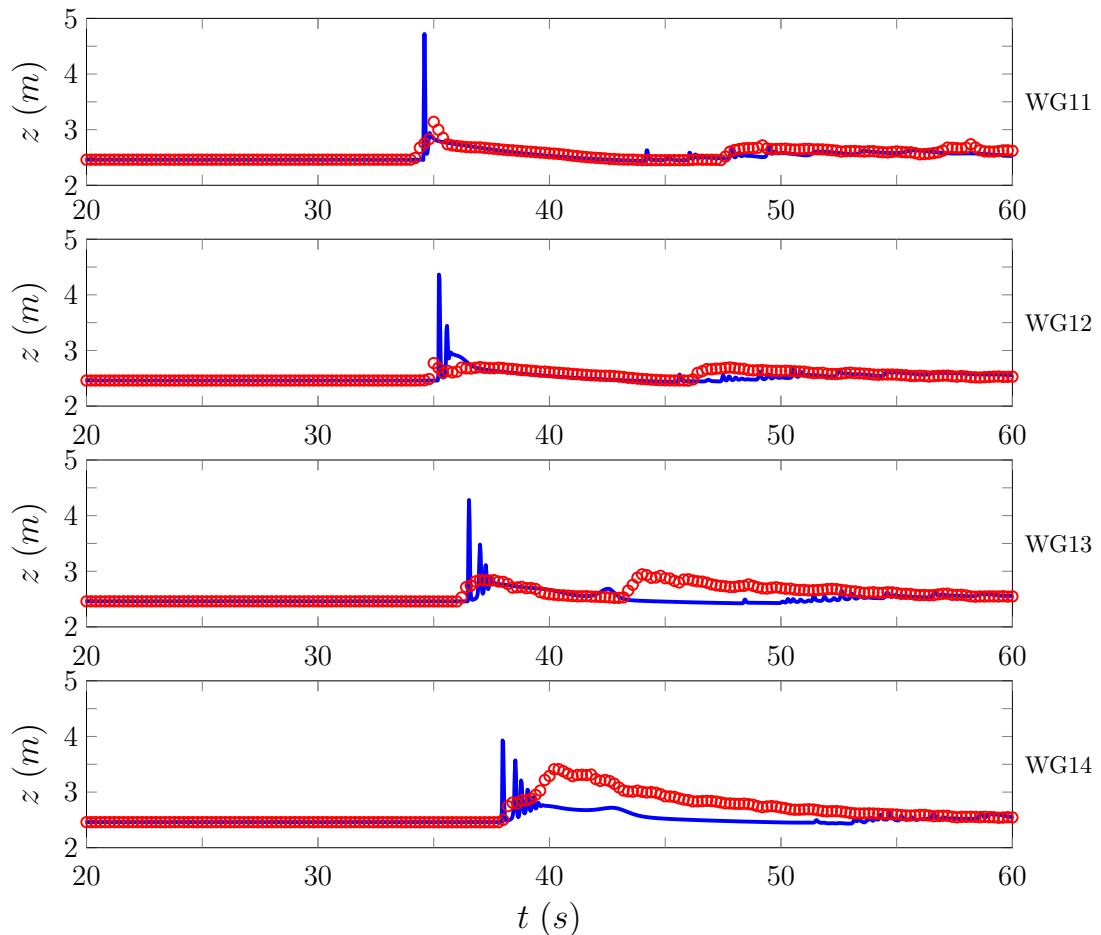


Figure 6.18: Time series of the experimental (○) and numerical (—) WG data produced by FEVM₂ for WG11 to WG14.

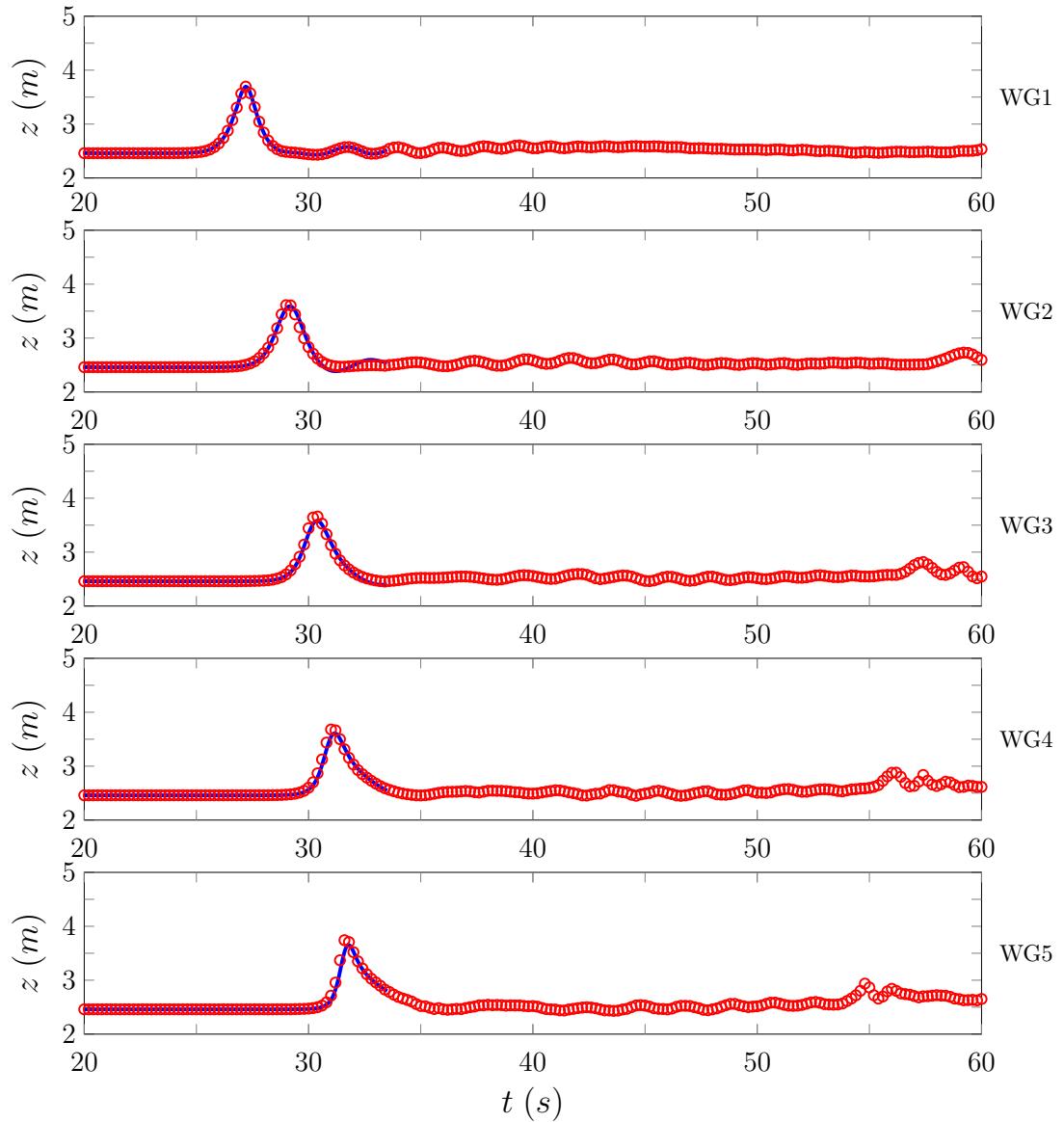


Figure 6.19: Time series of the experimental (○) and numerical (—) WG data produced by FDVM₂ for WG1 to WG7.

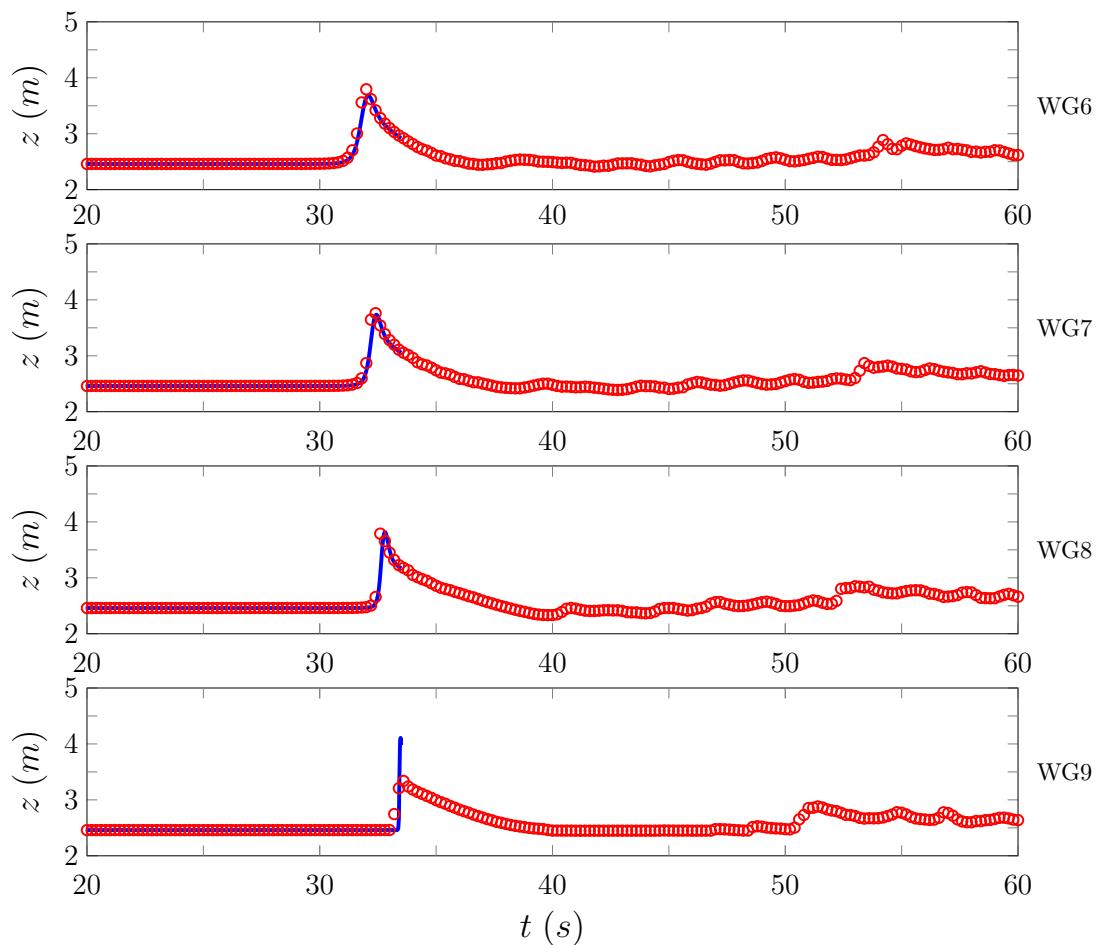


Figure 6.20: Time series of the experimental (○) and numerical (—) WG data produced by FDVM₂ for WG6 to WG9.

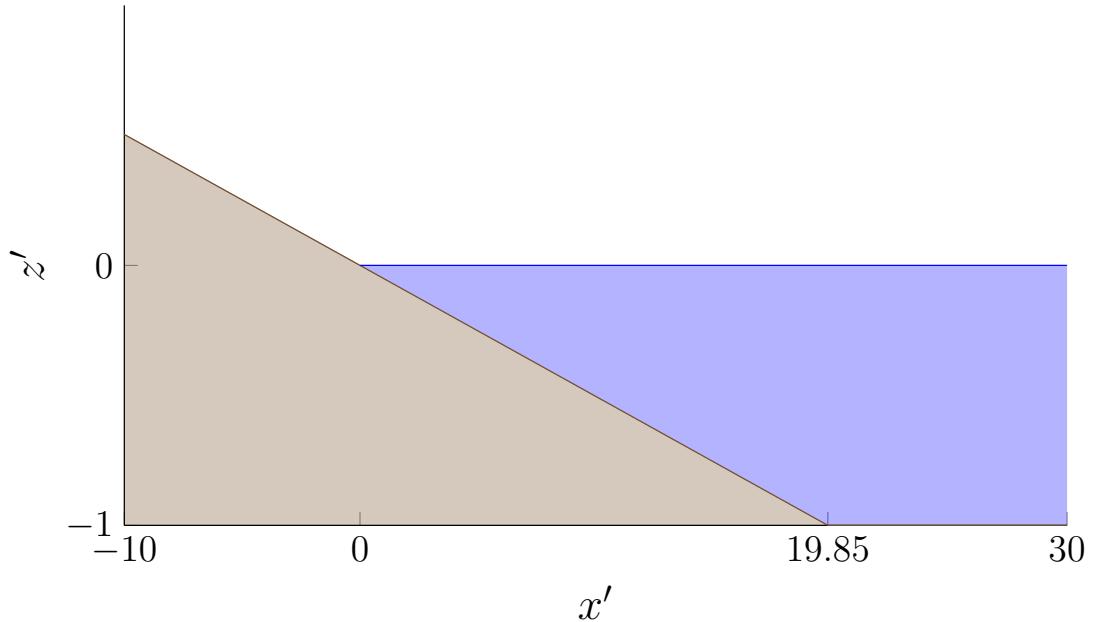


Figure 6.21: Diagram showing a longitudinal section of the wave tank for the run-up experiment with the water (■) and the bed (■) where the coordinates have been non-dimensionalised [65].

rate breaking wave modelling techniques to more accurately model some physical situations.

6.4 Run-up Experiment

To study the run-up of incoming waves on linear beaches a series of experiments were conducted by Synolakis [65]. These experiments consisted of a number of run-up events for a wide array of breaking and non-breaking waves where snapshots of the entire water surface were taken at certain times. These experiments were all performed on the beach profile depicted in Figure 6.21, where all the quantities are non-dimensionalised [65]. To denote that a quantity is non-dimensionalised a prime is used; for example for a generic quantity q its non-dimensionalised version is q' . The computational methods were used to model the run-up of a non-breaking solitary wave with a non-linearity parameter of $\epsilon = 0.0185$.

This experiment allows the inundation behaviour of the numerical methods to be compared with experimental results. For this experiment the effect of dispersion on the run-up behaviour is minimal, and there is good agreement

between numerical solutions of the SWWE and this particular experiment [66]. Therefore, the effect of the extra dispersive terms included by the Serre equations on the inundation process is not well tested by this experiment. However, this experiment does demonstrate the robustness of the numerical methods during the wetting and drying of the bed.

The numerical experiments used the non-dimensionalised quantities reported by Synolakis [65] to reproduce the experiment. The spatial domain was $x' \in [-30, 150]$ with a resolution of $\Delta x = 0.05$ and was run until $t' = 70$ with the CFL condition (3.23) satisfied by setting $\Delta t = 0.1\Delta x$. The spatial reconstruction used the input parameter $\theta = 1.2$ and the acceleration due to gravity $g = 1$ was chosen to match the non-dimensionalisation.

The non-dimensionalised water surface data is given at the various times in Figure 6.23 for FDVM₂ and 6.22 for FEVM₂. The error in conservation of h' and \mathcal{H}' as measured by C^* are given in Tables 6.5 and 6.6 for FEVM₂ and FDVM₂ respectively.

The results for FEVM₂ and FDVM₂ are indistinguishable and replicate the incoming wave properties and the maximum run-up well. The experimental wave appears to be more skewed towards the shoreline, but this shape difference has all but disappeared as the wave begins to inundate the shore. Additionally, the numerical solutions run-down further than the experimental results, due to the omission of bed friction in the Serre equations in this thesis.

Both h' and \mathcal{H}' are well conserved by the method throughout the run-up and run-down of the wave, particularly h' . However, the total amount of \mathcal{H}' appears to have slightly increased in the simulation during the run-up process due to the methods handling of the dry bed problem. During this experiment kinetic energy is converted into gravitational potential energy and then back again as the wave is reflected, therefore $u'h'$ and G' will only be conserved in this experiment after the wave has completely reflected from the beach. Full reflection of the wave has not occurred by $t' = 70$ and so the conservation results for uh and G were omitted from Tables 6.5 and 6.6.

The numerical solutions demonstrate good agreement with the experimental results and display the capability of FEVM₂ and FDVM₂ to model the inundation of non-breaking waves.

In this chapter FEVM₂ and FDVM₂ were validated using experimental data. It was found that for most experiments the solutions of FEVM₂ and FDVM₂ were indistinguishable although FEVM₂ is the preferred method due to its greater

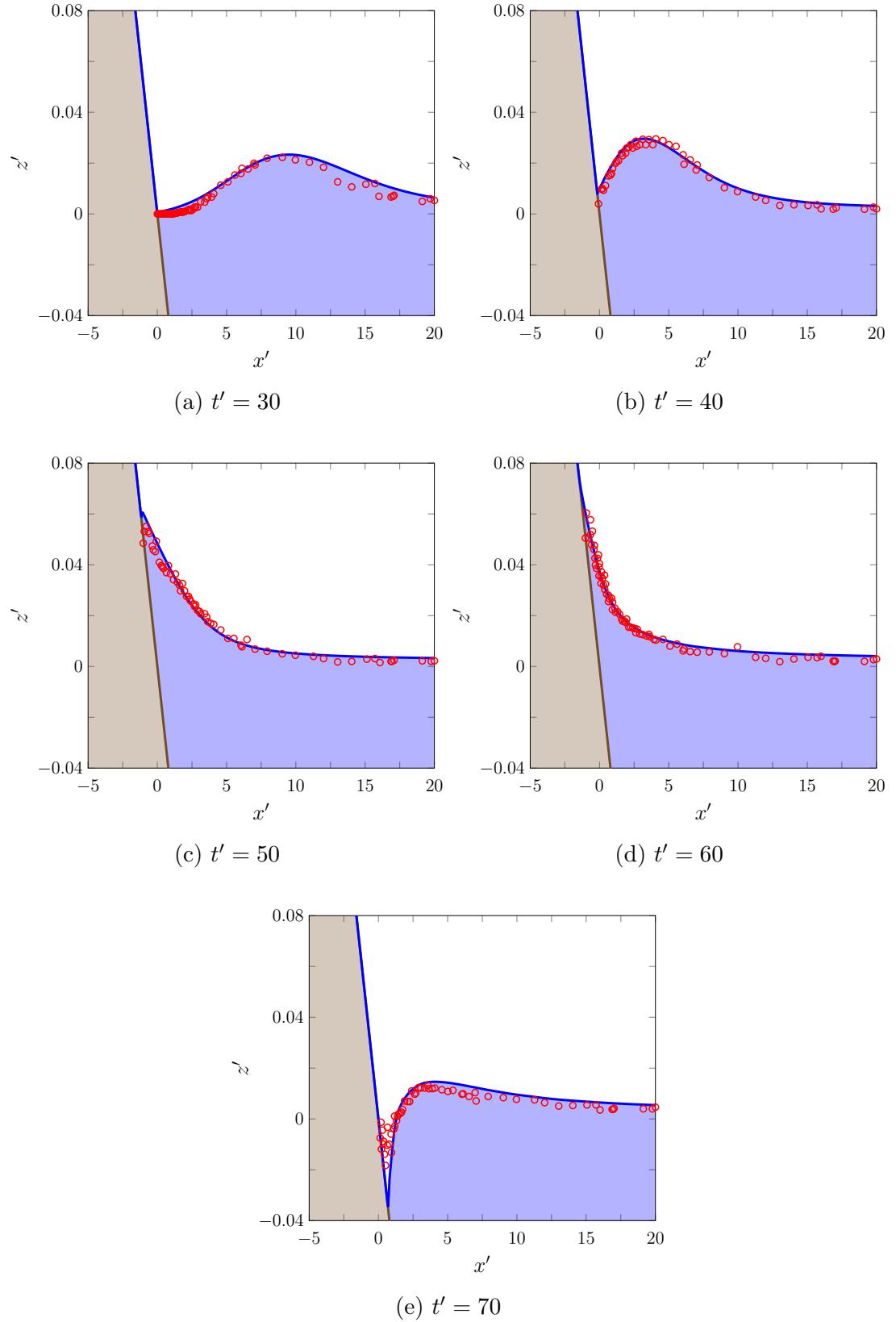


Figure 6.22: A comparison of the water surface profiles $w'(x', t')$ for the experiment (○) and the numerical solution (■) produced by FEVM₂ over the bed (■) at various times.

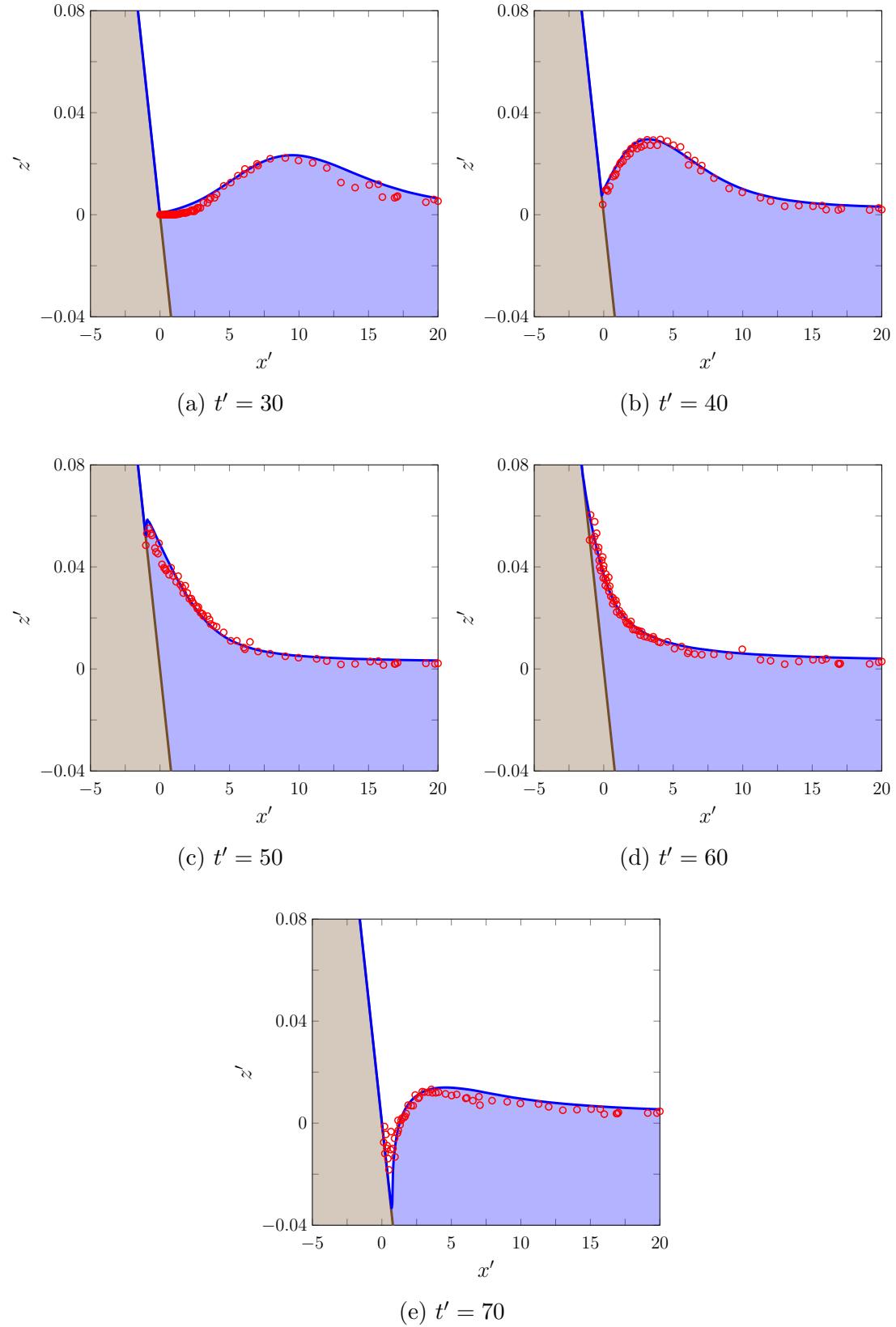


Figure 6.23: A comparison of the water surface profiles $w'(x', t')$ for the experiment (○) and the numerical solution (■) produced by FDVM₂ over the bed (■) at various times.

Quantity	$\mathcal{C}^*(\mathbf{q}^0)$	$\mathcal{C}^*(\mathbf{q}^*)$	$C^*(\mathbf{q}^0, \mathbf{q}^*)$
h'	140.4170	140.4170	7.65×10^{-12}
\mathcal{H}'	68.3900	68.3914	2.16×10^{-5}

Table 6.5: Initial and final total amounts and the conservation error for h' and \mathcal{H}' for the numerical solution of FEVM₂ for the run-up experiment.

Quantity	$\mathcal{C}^*(\mathbf{q}^0)$	$\mathcal{C}^*(\mathbf{q}^*)$	$C^*(\mathbf{q}^0, \mathbf{q}^*)$
h'	140.4170	140.4170	1.11×10^{-7}
\mathcal{H}'	68.3900	68.3914	2.16×10^{-5}

Table 6.6: Initial and final total amounts and the conservation error for h' and \mathcal{H}' for the numerical solution of FDVM₂ for the run-up experiment.

robustness.

Chapter 7

Conclusion

The evolution of the dam-break problem for the Serre equations was comprehensively studied using various numerical methods. The variety of numerical solutions provided a justification for the newly observed growth structure, the main feature of which is the growth in amplitude in the dispersive wave train around the contact discontinuity. Furthermore, it was found that the effects of smoothing of the initial conditions, the resolution and the diffusive error of the numerical method all contributed to produce different structures in the numerical solutions. These effects were the cause of the different structures previously published in the literature.

A well-balanced second-order Finite Element Volume Method (FEVM) termed FEVM_2 was described for the one-dimensional Serre equations. The method uses a finite element and a finite volume method and thus is robust to steep gradients in the conserved variables h and G . The use of these methods also makes the FEVM well suited to solve the two-dimensional Serre equations using unstructured meshes with parallelised code.

A linear analysis of the convergence and dispersion properties of FEVM_2 was performed and the results presented. The results of the linear analysis for the Finite Difference Volume Methods of first- (FDVM_1), second- (FDVM_2) and third-order (FDVM_3) described by Zoppou et al. [15] and the second-order finite difference methods \mathcal{D} and \mathcal{W} described by Pitt et al. [18] were also provided. The analysis demonstrated that FDVM_1 , FDVM_2 , FDVM_3 , FEVM_2 and \mathcal{D} are convergent methods, while \mathcal{W} is only convergent when the mean background flow velocity is zero. The dispersion analysis demonstrated that all methods approximated the dispersion relation of the Serre equations with the expected order of accuracy. The presented linear analysis extended a previous analysis of the

dispersion relationships of numerical methods [40] by allowing non-zero mean flow, combining the spatial and temporal analyses and comparing the real and imaginary parts of the dispersion error.

A comparison of the solutions of various numerical methods for the analytic solitary travelling wave solution of the Serre equations was performed. The order of accuracy and the conservation properties of the methods were as expected. Furthermore, the results demonstrated that the increase in accuracy achieved by a third-order method over a second-order method did not warrant the extra computational effort, justifying the further development of second-order methods over third-order methods for future work. For this reason only the second-order FDVM₂ and FEVM₂ were developed further to allow varying bathymetry and dry beds.

The second-order FDVM₂ and FEVM₂ were then validated against the lake at rest steady state and the forced solutions. The validation against the lake at rest steady state demonstrated that both methods are well-balanced, as they both accurately reproduced the lake at rest steady state. The necessity of the well-balancing modifications outlined in Chapter 3 was also demonstrated, as without these modifications the lake at rest steady state was not accurately reproduced. The validation using forced solutions demonstrated that both methods accurately approximated all terms in the Serre equations in the presence of dry beds.

Finally, the second-order FDVM₂ and FEVM₂ were compared to experimental data; demonstrating their modelling capabilities across a wide array of physical scenarios. The experimental comparison results established the greater robustness of FEVM₂; as FDVM₂ was found to be unstable for the solitary wave over a fringing reef experiment due to the presence of steep gradients in the water surface as the wave broke. Due to the greater robustness of FEVM₂ and its potential to be extended to unstructured meshes, FEVM₂ is the most promising of the methods presented in this thesis for solving the two-dimensional Serre equations in the future. For these reasons the FEVM₂ satisfies the overarching goal of the thesis as it is a method for the one-dimensional Serre equations that is robust in the presence of steep gradients in the free surface and during the inundation of a beach and it is extendable to the two-dimensional Serre equations using unstructured meshes.

To summarise the major contributions of the research underpinning this thesis are:

- Description and implementation of the third-order FDVM. The description

of the third-order FDVM was primarily produced by me and was published by Zoppou et al. [15].

- Observation and justification of a new structure in the solution of the Serre equations in the presence of steep gradients in the free surface. This work was published by Pitt et al. [18] and was summarised in Chapter 2.
- Extension of the second-order FDVM to allow for dry beds. A description of the second-order FDVM was published by Zoppou et al. [15]. To extend the second-order FDVM to allow for dry beds the desingularisation transformation in Chapter 3 was used, as well as the banded diagonal matrix solver of Press et al. [48]. Given the available description of the core components of the method, a complete description is not provided in this thesis.
- Development and description of the well-balanced second-order FEVM that is capable of modelling flows over dry beds. The second-order FEVM which is well-balanced and capable of modelling flows over dry beds is described in Chapter 3 and its desired properties are validated in Chapters 5 and 6.
- A linear analysis of convergence for all developed finite volume based methods as well as some finite difference methods was performed in Chapter 4.
- A complete linear analysis of the dispersion properties for all developed finite volume based methods as well as some finite difference methods was performed in Chapter 4.
- In Chapter 5 a validation of FEVM and the second-order FDVM using forced solutions where all terms of the Serre equations are present for both wet and dry beds was performed.
- Comparison of the numerical solutions of FEVM and the second-order FDVM with experimental results in the presence of dry beds and with wave breaking was presented in Chapter 6.

7.1 Future Work

Following the work conducted in this thesis; some possible extensions are:

- Inclusion of wave-breaking in the model.
- Implementation of different boundary conditions.
- Incorporation of discontinuous bed profiles.
- Incorporation of bed friction in FEVM₂.
- A complete analysis of the convergence properties of FEVM₂.
- Extension of FEVM₂ to the two-dimensional Serre equations on unstructured meshes.

7.1.1 Including Wave-breaking

In the presence of wave-breaking the Serre equations are no longer appropriate as was demonstrated in Chapter 6. One technique to improve numerical models based on the Serre equations in the presence of wave-breaking, is to revert to a Shallow Water Wave Equation (SWWE) solver under some criteria, typically based on wave steepness [40, 67, 68]. These numerical methods benefit from the splitting technique employed, resulting in a method which solves the hyperbolic part of the Serre equations, the SWWE followed by a method to solve the remaining dispersive part. Therefore, when the wave-breaking criteria is met the dispersive part is neglected, producing a solution of the SWWE for the breaking wave.

The method in this thesis is not based on such a splitting technique. However, by introducing a quantity β to the Serre equations in conservation law form (2.6) as follows

$$\begin{aligned} \frac{\partial h}{\partial t} + \frac{\partial(uh)}{\partial x} &= 0, \\ \frac{\partial G}{\partial t} + \frac{\partial}{\partial x} \left(uG + \frac{gh^2}{2} - \beta \left[\frac{2}{3}h^3 \left[\frac{\partial u}{\partial x} \right]^2 + h^2u \frac{\partial u}{\partial x} \frac{\partial b}{\partial x} \right] \right) \\ &+ \beta \left(\frac{1}{2}h^2u \frac{\partial u}{\partial x} \frac{\partial^2 b}{\partial x^2} - hu^2 \frac{\partial b}{\partial x} \frac{\partial^2 b}{\partial x^2} \right) + gh \frac{\partial b}{\partial x} = 0, \\ G &= uh + \beta \left[uh \left(\frac{\partial h}{\partial x} \frac{\partial b}{\partial x} + \frac{1}{2}h \frac{\partial^2 b}{\partial x^2} + \left[\frac{\partial b}{\partial x} \right]^2 \right) - \frac{\partial}{\partial x} \left(\frac{1}{3}h^3 \frac{\partial u}{\partial x} \right) \right] \end{aligned}$$

such a technique could be employed by setting $\beta = 0$ in the presence of wave-breaking and $\beta = 1$ otherwise. Since when $\beta = 0$, these modified Serre equations are equivalent to the SWWE.

7.1.2 Implementation of Different Boundary Conditions

When imposing boundary conditions other than the Dirichlet boundary conditions implemented in this thesis special care must be taken to ensure that the appropriate weak form of (2.7) is used. The presented weak form (3.5) is only valid because the surface integral of the terms were zero due to the use of Dirichlet boundary conditions.

The general weak form of (2.7) is

$$\begin{aligned} \int_{\Omega} Gv \, dx &= \int_{\Omega} uh \left(1 + \left[\frac{\partial b}{\partial x} \right]^2 \right) v \, dx + \int_{\Omega} \frac{1}{3} h^3 \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} \, dx - \int_{\Omega} \frac{1}{2} uh^2 \frac{\partial b}{\partial x} \frac{\partial v}{\partial x} \, dx \\ &\quad - \int_{\Omega} \frac{1}{2} h^2 \frac{\partial b}{\partial x} \frac{\partial u}{\partial x} v \, dx - \left[\frac{1}{3} h^3 \frac{\partial u}{\partial x} v \right]_{x_{-1/2}}^{x_{m+1/2}} + \left[\frac{1}{2} uh^2 \frac{\partial b}{\partial x} v \right]_{x_{-1/2}}^{x_{m+1/2}} \end{aligned} \quad (7.1)$$

where $\Omega = [x_{-1/2}, x_{m+1/2}]$ as in Chapter 3. Given the appropriate representation of h , G and b in the ghost cells, any appropriate boundary condition problem can be solved with the general weak form (7.1).

7.1.3 Incorporation of Discontinuous Bed Profiles

In the current derivation of the Serre equations, the bed profile is assumed to be smooth. To be able to model dispersive waves with non-smooth bed profiles requires a new derivation of the Serre equations. Although for most practical purposes smoothing the bed sufficiently will produce adequate results.

7.1.4 Incorporation of Bed Friction

The Serre equations presented in this thesis (2.6) assume no bottom friction. To include bottom friction approximation there are many options as noted by do Carmo et al. [68]. Although, all bottom friction approximations introduce some quadratic friction term to (2.4b), such as the Manning's bottom roughness [40, 67, 68]. An equivalent formulation for the momentum equation in conserva-

tion law form (2.6b) including Manning's bottom friction [69] is

$$\begin{aligned} \frac{\partial G}{\partial t} + \frac{\partial}{\partial x} \left(uG + \frac{gh^2}{2} - \frac{2}{3}h^3 \left[\frac{\partial u}{\partial x} \right]^2 + h^2 u \frac{\partial u}{\partial x} \frac{\partial b}{\partial x} \right) \\ + \frac{1}{2}h^2 u \frac{\partial u}{\partial x} \frac{\partial^2 b}{\partial x^2} - hu^2 \frac{\partial b}{\partial x} \frac{\partial^2 b}{\partial x^2} + gh \left(\frac{\partial b}{\partial x} + \frac{n^2 u |u|}{h^{4/3}} \right) = 0. \end{aligned} \quad (7.2)$$

Since the second-order FEVM in its current form already possesses u and h over every cell, this extra term could be added into the centred source term approximation in subsection 3.2.4.

7.1.5 Complete Convergence Analysis

The presented convergence analysis in Chapter 4 is only performed for the linearised Serre equations with a horizontal bed. Given the difficulty of proving convergence for Finite Volume Methods for the SWWE [70] this task currently seems out of reach for the Serre equations but remains of great interest.

7.1.6 Extension to two-dimensional Serre equations on unstructured meshes

The two-dimensional Serre equations have been written in conservation law form with a source term [14]. Given the experience in solving equations in conservation law form with a source term [8, 10] using unstructured meshes and the utility of the Finite Element Method to solve the elliptic equation for the new conserved quantity on unstructured meshes, producing an extension to the FEVM seems straightforward, but is not without its challenges. Since ANUGA [10] uses a second-order reconstruction and the same flux approximation scheme as the FEVM, the real challenge lies in extending the reconstruction of the bed profile and the elliptic solver to the two-dimensional Serre equations. For the reconstruction of the bed profile splines seem an obvious choice, as they would permit the use of unstructured meshes and allow continuity across the volume edges to be enforced as is also required for the two-dimensional Serre equations. While for the elliptic solver, an investigation into the derivatives required to approximate the flux and source functions would need to be performed to determine the relevant basis function spaces.

Appendix A

Additional Conservation Information

To calculate the conservation errors requires an analytic expression for the total amount of h , uh , G and \mathcal{H} present in the initial conditions. Therefore, to facilitate the validation tests against the analytic solutions performed in Chapter 5 analytic expressions for total amounts of these quantities contained in the initial conditions of the solitary travelling wave and the lake at rest solutions described in Chapter 2 are presented here. To allow for the simple calculation of the integrals in a concise way for any domain the integrals are provided in indefinite form.

A.1 Solitary Travelling Wave Solution

The solitary wave solution (2.11) is

$$h(x, t) = a_0 + a_1 \operatorname{sech}^2(\kappa [x - ct]),$$

$$u(x, t) = c \left(1 - \frac{a_0}{h(x, t)} \right),$$

$$b(x) = 0.$$

When $t = 0$ the indefinite spatial integrals of the conserved quantities are

$$\int h(x, 0) dx = a_0 x + \frac{a_1}{\kappa} \tanh(\kappa x) + \text{constant}, \quad (\text{A.1a})$$

$$\int u(x, 0)h(x, 0) dx = \frac{a_1 c}{\kappa} \tanh(\kappa x) + \text{constant}, \quad (\text{A.1b})$$

$$\begin{aligned} \int G(x, 0) dx &= \frac{ca_1}{3\kappa} \left(3 + 2a_0^2\kappa^2 \operatorname{sech}^2(\kappa x) \right. \\ &\quad \left. + 2a_0a_1\kappa^2 \operatorname{sech}^4(\kappa x) \right) \tanh(\kappa x) + \text{constant}, \end{aligned} \quad (\text{A.1c})$$

$$\begin{aligned} \int \mathcal{H}(x, 0) dx &= \frac{1}{2} \left(\int g [h(x, 0)]^2 dx + \int h(x, 0) [u(x, 0)]^2 dx \right. \\ &\quad \left. + \int [h(x, 0)]^3 \left[\frac{\partial u(x, 0)}{\partial x} \right]^2 dx \right) \end{aligned} \quad (\text{A.1d})$$

where the integrals making up \mathcal{H} are

$$\begin{aligned} \int g [h(x, 0)]^2 dx &= \frac{g}{12\kappa} \operatorname{sech}^3(\kappa x) \left(9a_0^2\kappa x \cosh(\kappa x) \right. \\ &\quad + 4a_1 [3a_0 + 2a_1 + (3a_0 + a_1) \cosh(2\kappa x)] \sinh(\kappa x) \\ &\quad \left. + 3a_0^2\kappa x \cosh(3\kappa x) \right) + \text{constant}, \end{aligned}$$

$$\begin{aligned} \int h(x, 0) [u(x, 0)]^2 dx &= \frac{\sqrt{a_1}c^2}{\kappa} \left(- \frac{a_0}{\sqrt{a_0 + a_1}} \operatorname{arctanh} \left(\frac{\sqrt{a_1} \tanh(\kappa x)}{\sqrt{a_0 + a_1}} \right) \right. \\ &\quad \left. + \frac{\sqrt{a_1}}{\kappa} \tanh(\kappa x) \right) + \text{constant}, \end{aligned}$$

$$\begin{aligned} \int [h(x, 0)]^3 \left[\frac{\partial u(x, 0)}{\partial x} \right]^2 dx &= \frac{2a_0^2c^2\kappa}{9\sqrt{a_1} (a_0 + a_1 \operatorname{sech}^2(\kappa x))} \\ &\quad \times (a_0 + 2a_1 + a_0 \cosh(2\kappa x)) \operatorname{sech}^2(\kappa x) \\ &\quad \times \left(- 3a_0\sqrt{a_0 + a_1} \operatorname{arctanh} \left(\frac{\sqrt{a_1} \tanh(\kappa x)}{\sqrt{a_0 + a_1}} \right) \right. \\ &\quad \left. + \sqrt{a_1} [3a_0 + a_1 - a_1 \operatorname{sech}^2(\kappa x)] \tanh(\kappa x) \right) \\ &\quad + \text{constant}. \end{aligned}$$

From these expressions the total amount of h , uh , G and \mathcal{H} for the solitary travelling wave solution (2.11) can be calculated over any domain at $t = 0$.

A.2 Lake At Rest Solution

The lake at rest solution (2.12) for an arbitrary bed profile $b(x)$ is

$$\begin{aligned} h(x, t) &= \max \{a_0 - b(x), 0\}, \\ u(x, t) &= 0, \\ G(x, t) &= 0. \end{aligned}$$

The total amount of uh and G in the system are straightforward to calculate as both are zero everywhere and thus it follows that

$$\begin{aligned} \int u(x, 0)h(x, 0) dx &= 0 + \text{constant}, \\ \int G(x, 0) dx &= 0 + \text{constant}. \end{aligned}$$

To calculate the total amount of h and \mathcal{H} in the solution the domain is partitioned into wet regions where $b(x) < a_0$ and dry regions where $b(x) \geq a_0$. For the dry regions h and \mathcal{H} are zero everywhere and therefore

$$\begin{aligned} \int h(x, 0) dx &= 0 + \text{constant}, \\ \int \mathcal{H}(x, 0) dx &= 0 + \text{constant} \end{aligned}$$

whilst in the wet regions

$$\int h(x, 0) dx = a_0 x - \int b(x) dx + \text{constant}, \quad (\text{A.2a})$$

$$\int \mathcal{H}(x, 0) dx = \frac{g}{2} \left(a_0^2 x - 2a_0 \int b(x) dx + \int b(x)^2 dx \right) + \text{constant}. \quad (\text{A.2b})$$

By summing all the wet regions in a given domain together the total amount of h and \mathcal{H} in the system can be calculated from these expressions for an arbitrary bed profile $b(x)$.

Appendix B

Finite Element Method Definitions

The definitions of the basis functions of the finite element method used by FEVM₂ described in Chapter 3 and the function spaces mentioned in Chapter 3 are provided here. Beginning with the basis function definitions.

B.1 Basis Function Definitions

Since all integrals of the basis functions are calculated with respect to the variable ξ , the basis functions are given in terms of ξ . The mapping from the x -space of the numerical grid to the canonical ξ -space is

$$x = x_j + \xi \frac{\Delta x}{2}.$$

This mapping takes the j^{th} cell $[x_{j-1/2}, x_{j+1/2}]$ in the x -space to the interval $[-1, 1]$ in the ξ -space.

The basis functions ψ for h and G shown in Figure B.1 are

$$\psi_{j-1/2}^+ = \begin{cases} \frac{1}{2}(1 - \xi) & -1 \leq \xi \leq 1 \\ 0 & \text{otherwise,} \end{cases} \quad (\text{B.1a})$$

$$\psi_{j+1/2}^- = \begin{cases} \frac{1}{2}(1 + \xi) & -1 \leq \xi \leq 1 \\ 0 & \text{otherwise.} \end{cases} \quad (\text{B.1b})$$

The basis functions ϕ for u and the test function v displayed in Figure B.2

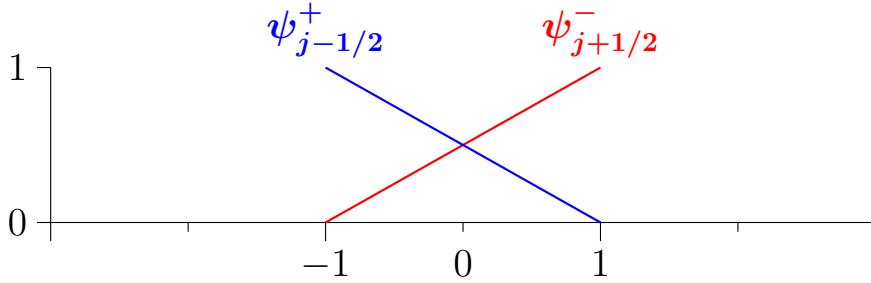


Figure B.1: Support of the discontinuous linear basis functions ψ which are non-zero over the j^{th} cell plotted in the ξ -space.

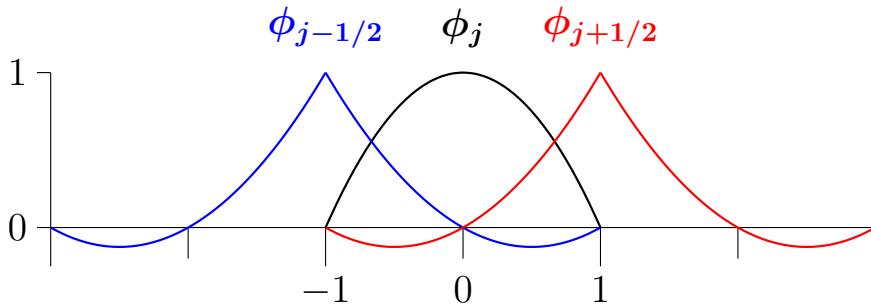


Figure B.2: Support of the continuous piecewise quadratic basis functions ϕ which are non-zero over the j^{th} cell plotted in the ξ -space.

are given by

$$\phi_{j-1/2} = \begin{cases} 2\left(\xi + \frac{3}{2}\right)(\xi + 2) & -2 \leq \xi \leq -1 \\ \frac{1}{2}\xi(\xi - 1) & -1 \leq \xi \leq 1 \\ 0 & \text{otherwise,} \end{cases} \quad (\text{B.2a})$$

$$\phi_j = \begin{cases} -(\xi - 1)(\xi + 1) & -1 \leq \xi \leq 1 \\ 0 & \text{otherwise,} \end{cases} \quad (\text{B.2b})$$

$$\phi_{j+1/2} = \begin{cases} \frac{1}{2}\xi(\xi + 1) & -1 \leq \xi \leq 1 \\ 2(\xi - 2)\left(\xi - \frac{3}{2}\right) & 1 \leq \xi \leq 2 \\ 0 & \text{otherwise.} \end{cases} \quad (\text{B.2c})$$

Finally the basis functions γ for the bed profile b displayed in Figure B.3 are

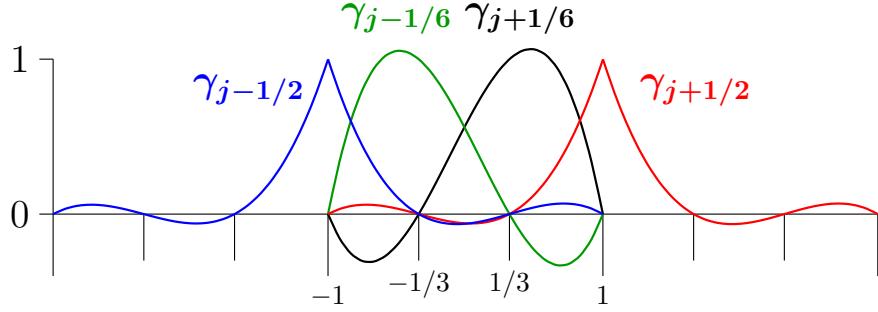


Figure B.3: Support of the continuous piecewise cubic basis functions γ which are non-zero over the j^{th} cell plotted in the ξ -space.

given by

$$\gamma_{j-1/2} = \begin{cases} \frac{9}{2} (\xi + \frac{4}{3}) (\xi + \frac{5}{3}) (\xi + 2) & -2 \leq \xi \leq -1 \\ \frac{9}{16} (\xi - 1) (\xi - \frac{1}{3}) (\xi + \frac{1}{3}) & -1 \leq \xi \leq 1 \\ 0 & \text{otherwise,} \end{cases} \quad (\text{B.3a})$$

$$\gamma_{j-1/6} = \begin{cases} \frac{27}{16} (\xi - 1) (\xi - \frac{1}{3}) (\xi + 1) & -1 \leq \xi \leq 1 \\ 0 & \text{otherwise,} \end{cases} \quad (\text{B.3b})$$

$$\gamma_{j+1/6} = \begin{cases} -\frac{27}{16} (\xi - 1) (\xi + \frac{1}{3}) (\xi + 1) & -1 \leq \xi \leq 1 \\ 0 & \text{otherwise,} \end{cases} \quad (\text{B.3c})$$

$$\gamma_{j+1/2} = \begin{cases} \frac{9}{16} (\xi + 1) (\xi - \frac{1}{3}) (\xi + \frac{1}{3}) & -1 \leq \xi \leq 1 \\ -\frac{9}{2} (\xi - \frac{4}{3}) (\xi - \frac{5}{3}) (\xi - 2) & 1 \leq \xi \leq 2 \\ 0 & \text{otherwise.} \end{cases} \quad (\text{B.3d})$$

The calculation of the derivatives of these basis functions with respect to ξ are straightforward and hence omitted.

B.2 Function Space Definitions

The function spaces mentioned in Chapter 3 are the space of square integrable functions $\mathbb{L}^2(\Omega)$ and the Sobolev space $\mathbb{W}^{k,2}(\Omega)$. To be precise a definition of these function spaces is provided here.

A function $f(x)$ is in $\mathbb{L}^2(\Omega)$ if

$$\left(\int_{\Omega} f(x)^2 dx \right)^{\frac{1}{2}} < \infty.$$

While $f(x)$ is in $\mathbb{W}^{k,2}(\Omega)$ if

$$\left(\int_{\Omega} f(x)^2 dx + \sum_{j=1}^k \int_{\Omega} [D^j f(x)]^2 dx \right)^{\frac{1}{2}} < \infty$$

where $D^j f(x)$ is the j^{th} weak derivative of $f(x)$.

Appendix C

Linear Analysis Results

In this appendix additional details supplementing the linear analysis results reported in Chapter 4 are provided. In particular, the evolution matrices \mathbf{E} for all methods in this thesis, details on how to calculate the convergence and dispersion properties from the evolution matrix where appropriate and the consistency tables not displayed in Chapter 4 are all presented here.

The linear analysis in Chapter 4 studies the convergence and dispersion properties of the numerical methods solving the linearised Serre equations with a horizontal bed (4.3). The linearisation of the Serre equations assumes that there is a small perturbation η on the mean water depth H and a small perturbation μ on the mean background flow velocity U and then smaller terms are neglected. The resultant linearised Serre equations with a horizontal bed (4.3) are

$$\frac{\partial \eta}{\partial t} + H \frac{\partial \mu}{\partial x} + U \frac{\partial \eta}{\partial x} = 0, \quad (\text{C.1a})$$

$$H \frac{\partial \mu}{\partial t} + gH \frac{\partial \eta}{\partial x} + UH \frac{\partial \mu}{\partial x} - \frac{H^3}{3} \left(U \frac{\partial^3 \mu}{\partial x^3} + \frac{\partial^3 \mu}{\partial x^2 \partial t} \right) = 0 \quad (\text{C.1b})$$

which can be written in conservation law form (4.4) for η and G as

$$\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x} (H\mu + U\eta) = 0, \quad (\text{C.2a})$$

$$\frac{\partial G}{\partial t} + \frac{\partial}{\partial x} (UG + UH\mu + gH\eta) = 0 \quad (\text{C.2b})$$

where

$$G = UH + U\eta + H\mu - \frac{H^3}{3} \frac{\partial^2 \mu}{\partial x^2}. \quad (\text{C.2c})$$

The linear analysis assumes that η and μ are Fourier modes (4.5). A generic quantity q is a Fourier mode if

$$q(x, t) = q(0, 0) e^{i(\omega \pm t + kx)} \quad (\text{C.3})$$

where ω^\pm is the frequency of the wave given by the dispersion relation (2.9) and k is the wavenumber. There are two possible frequencies of waves corresponding to upwind and downwind travelling waves these are denoted with the $-$ and $+$ superscripts respectively. Since η and μ are Fourier modes (C.3) then G as well as the cell averages $\bar{\eta}$, $\bar{\mu}$ and \bar{G} are also Fourier modes.

For a fixed grid such as those used by all numerical methods in this thesis the Fourier mode quantities at different nodal values can be related to one another. Hence, for a generic Fourier mode q

$$q_{j+l}^{n+m} = q_j^n e^{i(m\omega^\pm \Delta t + lk\Delta x)} \quad (\text{C.4})$$

where Δt is the length of a time step and Δx is the length of a cell.

Applying the numerical method to the linearised Serre equations (C.1) and using (C.4) one can derive an expression relating the quantities at the current and earlier times to the quantities at the next time via the evolution matrix \mathbf{E} of the numerical method. This evolution matrix can then be analysed to obtain the convergence and dispersion properties of the numerical method.

This appendix provides the process to calculate \mathbf{E} for the first-, second- and third-order Finite Difference Volume Methods (FDVM) which are named FDVM₁, FDVM₂ and FDVM₃ respectively. Descriptions of these methods were published by Zoppou et al. [15]. Additionally, the evolution matrices for the two second-order Finite Difference Methods (FDM) \mathcal{D} and \mathcal{W} are also provided, descriptions of these methods were published by Pitt et al. [18].

The linear analysis for the second-order finite element volume method FEVM₂ was performed in Chapter 4. Given that work the evolution matrix for other finite volume based methods can be written in terms of the factors given by the reconstruction operators and the velocity solve operators of the method. The process to apply Chapter 4 for other finite volume based methods is outlined here, along with the expressions for the required factors. Hence, this appendix allows others to reproduce the evolution matrices for all finite volume based methods in this thesis without reproducing the same process multiple times.

For \mathcal{D} and \mathcal{W} the evolution matrix itself is provided. Since these FDM use the quantities at previous time steps, the shape of the evolution matrix is different. Hence, the process to obtain the convergence and dispersion properties from these particular evolution matrices is explained.

Finally, the consistency results for FDVM₁, FDVM₂, FDVM₃, \mathcal{D} and \mathcal{W} are provided. While the results for FEVM₂ were provided in Chapter 4.

Method	E
FDVM ₁	$\mathbf{I} - \Delta t \mathbf{F}$
FDVM ₂ and FEVM ₂	$\mathbf{I} - \Delta t \mathbf{F} + \frac{1}{2} \Delta t^2 \mathbf{F}^2$
FDVM ₃	$\mathbf{I} - \Delta t \mathbf{F} + \frac{1}{2} \Delta t^2 \mathbf{F}^2 - \frac{1}{6} \Delta t^3 \mathbf{F}^3$

Table C.1: Formula for **E** given **F** determined by the SSP Runge-Kutta time stepping method.

C.1 Finite Difference Volume Methods

For the finite volume based methods the evolution matrix is a 2×2 matrix that gives the following relationship (4.1)

$$\begin{bmatrix} \bar{\eta} \\ \bar{G} \end{bmatrix}_j^{n+1} = \mathbf{E} \begin{bmatrix} \bar{\eta} \\ \bar{G} \end{bmatrix}_j^n.$$

This evolution matrix **E** is obtained by applying the numerical method to (C.2) and making use of (C.4) for all quantities.

In Chapter 4 the evolution matrix for FEVM₂ is given in terms of the flux matrix **F** obtained from the flux approximation of the method. This expression relating **E** and **F** is derived from the employed SSP Runge-Kutta time stepping method. The expressions for all employed SSP Runge-Kutta time stepping schemes and their associated numerical methods are summarised in Table C.1. Since FDVM₂ and FEVM₂ both use second-order SSP Runge-Kutta time stepping, their expressions are the same.

The flux matrix **F** from the finite volume methods approximation to the flux terms (4.12) is

$$\mathbf{F} = -\frac{(1 - e^{-ik\Delta x})}{\Delta x} \begin{bmatrix} \mathcal{F}^{\eta,\eta} & \mathcal{F}^{\eta,G} \\ \mathcal{F}^{G,\eta} & \mathcal{F}^{G,G} \end{bmatrix}. \quad (\text{C.5})$$

Where $\mathcal{F}^{\eta,\eta}$, $\mathcal{F}^{\eta,G}$, $\mathcal{F}^{G,\eta}$, $\mathcal{F}^{G,G}$ can be written in terms of factors given by the reconstruction and the velocity solve operators. These expressions written in

Froude Number	$\mathcal{F}^{\eta,\eta}$
$Fr < -1$	$H\mathcal{G}^\eta + U\mathcal{R}_{j+1/2}^+$
$-1 \leq Fr \leq 1$	$H\mathcal{G}^\eta + \frac{U}{2} \left(\mathcal{R}_{j+1/2}^- + \mathcal{R}_{j+1/2}^+ \right) - \frac{\sqrt{gH}}{2} \left(\mathcal{R}_{j+1/2}^+ - \mathcal{R}_{j+1/2}^- \right)$
$1 < Fr$	$H\mathcal{G}^\eta + U\mathcal{R}_{j+1/2}^-$

Table C.2: Factor $\mathcal{F}^{\eta,\eta}$ that multiples η in the flux function for η for all finite volume based methods.

terms of the constituent factors are the same for all finite volume based methods in the thesis, as they all used the flux approximation of Kurganov et al. [49].

The term $\mathcal{F}^{\eta,G}$ in the flux matrix does not depend on the Froude number $Fr = U/gH$ and is

$$\mathcal{F}^{\eta,G} = H\mathcal{G}^G.$$

The expressions for the other terms of \mathbf{F} which do depend on the Froude number are summarised in Tables C.2-C.4 for all values of Fr with $\mathcal{F}^{\eta,\eta}$ in Table C.2, $\mathcal{F}^{G,\eta}$ in Table C.3 and $\mathcal{F}^{G,G}$ in Table C.4.

Using the appropriate expressions for the factors generated by the reconstruction operators \mathcal{R}_j , $\mathcal{R}_{j-1/2}^+$, $\mathcal{R}_{j+1/2}^-$ and the velocity solve operators \mathcal{G}^η and \mathcal{G}^G used by the method, one can obtain all the terms of \mathbf{F} (C.5) for the finite volume based methods.

The expressions for these fundamental factors of all finite volume based methods are summarised in Table C.5 for \mathcal{R}_j , Table C.6 for $\mathcal{R}_{j-1/2}^+$, Table C.7 for $\mathcal{R}_{j+1/2}^-$ and Table C.8 for \mathcal{G}^G . Since $\mathcal{G}^\eta = -U\mathcal{G}^G$ only the table for \mathcal{G}^G is provided. Furthermore, because all the perturbations are Fourier modes (4.5) the reconstruction factor $\mathcal{R}_{j+1/2}^+ = e^{ik\Delta x} \mathcal{R}_{j-1/2}^+$.

Tables C.5-C.8 also include the factors for FEVM₂ summarising the work in Chapter 4. The analytic value of the factors for an exact method are also provided. Additionally, the lowest order term of the Taylor series of the difference between the factor of the numerical method and the exact factor are also provided. The reported Taylor series results demonstrate that all methods use

Froude Number	$\mathcal{F}^{G,\eta}$
$Fr < -1$	$UH\mathcal{G}^\eta + gH\mathcal{R}_{j+1/2}^+$
$-1 \leq Fr \leq 1$	$\frac{U\sqrt{gH}}{2} (\mathcal{R}_{j+1/2}^- - \mathcal{R}_{j+1/2}^+) + UH\mathcal{G}^\eta + \frac{gH}{2} (\mathcal{R}_{j+1/2}^- + \mathcal{R}_{j+1/2}^+)$
$1 < Fr$	$UH\mathcal{G}^\eta + gH\mathcal{R}_{j+1/2}^-$

Table C.3: Factor $\mathcal{F}^{G,\eta}$ that multiples η in the flux function for G for all finite volume based methods.

Froude Number	$\mathcal{F}^{G,G}$
$Fr < -1$	$U\mathcal{R}_{j+1/2}^+ + UH\mathcal{G}^G$
$-1 \leq Fr \leq 1$	$UH\mathcal{G}^G + \frac{U}{2} (\mathcal{R}_{j+1/2}^- + \mathcal{R}_{j+1/2}^+) - \frac{\sqrt{gH}}{2} (\mathcal{R}_{j+1/2}^+ - \mathcal{R}_{j+1/2}^-)$
$1 < Fr$	$U\mathcal{R}_{j+1/2}^+ + UH\mathcal{G}^G$

Table C.4: Factor $\mathcal{F}^{G,G}$ that multiples G in the flux function for G for all finite volume based methods

Method	\mathcal{R}_j	Lowest Order Term of Method - Exact
Exact	$\frac{k\Delta x}{2 \sin\left(k\frac{\Delta x}{2}\right)}$	-
FDVM ₁	1	$-\frac{1}{24}k^2\Delta x^2$
FDVM ₂ and FEVM ₂	1	$-\frac{1}{24}k^2\Delta x^2$
FDVM ₃	$\frac{26 - 2 \cos(k\Delta x)}{24}$	$-\frac{3}{640}k^4\Delta x^4$

Table C.5: Factor \mathcal{R}_j from the reconstruction of the nodal value at the midpoint and the lowest order term of the Taylor series of the factor in the method minus the exact factor for all finite volume based methods.

approximations with the appropriate order of accuracy or better.

The terms of the flux matrix \mathbf{F} (C.5) can be calculated from these expressions for the factors summarised in Tables C.5 - C.8. The evolution matrix \mathbf{E} can then be calculated from \mathbf{F} based on the SSP Runge-Kutta time stepping expressions summarised in Table C.1. Thus all the evolution matrices of the finite volume based methods can be obtained, as desired.

Method	$\mathcal{R}_{j-1/2}^+$	Lowest Order Term of Method - Exact
Exact	$\frac{k\Delta x}{2 \sin\left(\frac{k\Delta x}{2}\right)} e^{-\frac{ik\Delta x}{2}}$	-
FDVM ₁	1	$\frac{i}{2}k\Delta x$
FDVM ₂ and FEVM ₂	$1 - \frac{i \sin(k\Delta x)}{2}$	$\frac{1}{12}k^2\Delta x^2$
FDVM ₃	$\frac{1}{6}(5 + 2e^{-ik\Delta x} - e^{ik\Delta x})$	$\frac{i}{12}k^3\Delta x^3$

Table C.6: Factor $\mathcal{R}_{j-1/2}^+$ from the reconstruction of η and G at $x_{j+1/2}$ from the $(j+1)^{th}$ cell and the lowest order term of the Taylor series of the factor in the method minus the exact factor for all finite volume based methods.

Method	$\mathcal{R}_{j+1/2}^-$	Lowest Order Term of Method - Exact
Exact	$\frac{k\Delta x}{2 \sin\left(\frac{k\Delta x}{2}\right)} e^{\frac{ik\Delta x}{2}}$	-
FDVM ₁	1	$-\frac{i}{2}k\Delta x$
FDVM ₂ and FEVM ₂	$1 + \frac{i \sin(k\Delta x)}{2}$	$\frac{1}{12}k^2\Delta x^2$
FDVM ₃	$\frac{1}{6}(5 - e^{-ik\Delta x} + 2e^{ik\Delta x})$	$-\frac{i}{12}k^3\Delta x^3$

Table C.7: Factor $\mathcal{R}_{j+1/2}^-$ from the reconstruction of η and G at $x_{j+1/2}$ using the j^{th} cell and the lowest order term of the Taylor series of the factor in the method minus the exact factor for all finite volume based methods.

Method	\mathcal{G}^G	Lowest Order Term of Method - Exact
Exact	$\frac{3}{3H + H^3 k^2} \frac{k\Delta x}{2 \sin\left(\frac{k\Delta x}{2}\right)} e^{\frac{ik\Delta x}{2}}$	-
FDVM ₁	$\frac{3\Delta x^2 (1 + e^{ik\Delta x})}{6\Delta x^2 H - 2H^3 (2 \cos(k\Delta x) - 2)}$	$-\frac{6 + H^2 k^2}{4H (3 + H^2 k^2)^2} k^2 \Delta x^2$
FDVM ₂	$\frac{3\Delta x^2 (1 + e^{ik\Delta x})}{6\Delta x^2 H - 2H^3 (2 \cos(k\Delta x) - 2)}$	$-\frac{6 + H^2 k^2}{4H (3 + H^2 k^2)^2} k^2 \Delta x^2$
FEVM ₂	$\begin{aligned} & \frac{\Delta x}{6} \left(1 + \frac{i \sin(k\Delta x)}{2} + e^{ik\Delta x} \left[1 - \frac{i \sin(k\Delta x)}{2} \right] \right) \\ & \div \left(H \frac{\Delta x}{30} \left[4 \cos\left(\frac{k\Delta x}{2}\right) - 2 \cos(k\Delta x) + 8 \right] \right. \\ & \left. + \frac{H^3}{9\Delta x} \left[-16 \cos\left(\frac{k\Delta x}{2}\right) + 2 \cos(k\Delta x) + 14 \right] \right) \end{aligned}$	$\frac{12 + 5H^2 k^2}{40H (3 + H^2 k^2)^2} k^2 \Delta x^2$
FDVM ₃	$\frac{9\Delta x^2 (-e^{-ik\Delta x} + 9e^{ik\Delta x} - e^{2ik\Delta x} + 9)}{144\Delta x^2 H - 4H^3 (32 \cos(k\Delta x) - 2 \cos(2k\Delta x) - 30)}$	$-\frac{243 + 49H^2 k^2}{960H (3 + H^2 k^2)^2} k^4 \Delta x^4$

Table C.8: Factor \mathcal{G}^G that multiples G given by solving (4.3c) for $v_{j+1/2}$ and the lowest order term of the Taylor series of the factor in the method minus the exact factor for all finite volume based methods.

C.2 Finite Difference Methods

The Finite Difference Methods (FDM) solve the linearised Serre equations with a horizontal bed in non-conservative form (C.1). The FDM rely on previous time steps as well as the current time step to update the quantities. One particular way of expressing this is with a 4×4 evolution matrix \mathbf{E} producing the following relationship

$$\begin{bmatrix} \eta^{n+1} \\ \mu^{n+1} \\ \eta^n \\ \mu^n \end{bmatrix}_j = \mathbf{E} \begin{bmatrix} \eta^n \\ \mu^n \\ \eta^{n-1} \\ \mu^{n-1} \end{bmatrix}_j \quad (\text{C.6})$$

where the time superscript was brought inside the vector to make clear the time step at which the different elements are placed. Since the FDM are used to calculate η_j^{n+1} and μ_j^{n+1} given η_j^n , μ_j^n , η_j^{n-1} and μ_j^{n-1} their evolution matrices have the following structure

$$\mathbf{E} = \begin{bmatrix} E_{0,0} & E_{0,1} & E_{0,2} & E_{0,3} \\ E_{1,0} & E_{1,1} & E_{1,2} & E_{1,3} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (\text{C.7})$$

Because $\eta_j^{n-1} = e^{-i\omega^\pm \Delta t} \eta_j^n$ and $\mu_j^{n-1} = e^{-i\omega^\pm \Delta t} \mu_j^n$ as η and μ are Fourier modes (C.3) then (C.6) can be rewritten as

$$\begin{bmatrix} \eta \\ \mu \end{bmatrix}_j^{n+1} = \mathbf{E}^{(2 \times 2)} \begin{bmatrix} \eta \\ \mu \end{bmatrix}_j^n \quad (\text{C.8})$$

where $\mathbf{E}^{(2 \times 2)}$ is a 2×2 matrix that depends on the elements of \mathbf{E} (C.7) in the following way

$$\mathbf{E}^{(2 \times 2)} = \begin{bmatrix} E_{0,0} + e^{-i\omega^\pm \Delta t} E_{0,2} & E_{0,1} + e^{-i\omega^\pm \Delta t} E_{0,3} \\ E_{1,0} + e^{-i\omega^\pm \Delta t} E_{1,2} & E_{1,1} + e^{-i\omega^\pm \Delta t} E_{1,3} \end{bmatrix}. \quad (\text{C.9})$$

Since the evolution matrix can be represented in two ways, it will now be stated whether \mathbf{E} or $\mathbf{E}^{(2 \times 2)}$ were used in the convergence and dispersion analysis of the FDM.

Using the Lax-Equivalence theorem the convergence of the methods were analysed by analysing their stability and consistency separately. The stability analysis was performed by finding the spectral radius of the naive evolution matrix \mathbf{E} of

the FDM (C.6). The consistency analysis was performed by comparing the 2×2 evolution matrix $\mathbf{E}^{(2 \times 2)}$ (C.9) to the exact evolution matrix $e^{i\omega^\pm \Delta t} \mathbf{I}$ for (C.8). Finally, the dispersion error was derived using the eigenvalues of \mathbf{E} (C.6), this matrix has an additional two eigenvalues beyond the ones given by $e^{i\omega^+ \Delta t}$ and $e^{i\omega^- \Delta t}$ that were ignored. It was found that the methods had the same stability and dispersion properties when $\mathbf{E}^{(2 \times 2)}$ was investigated. The 4×4 evolution matrices for \mathcal{D} and \mathcal{W} are now presented. Given these matrices the corresponding 2×2 evolution matrix $\mathbf{E}^{(2 \times 2)}$ can be calculated using (C.9).

By using (C.4) all the derivative approximations in the finite difference methods \mathcal{D} and \mathcal{W} can be written as factors that are constant in j and n as was done for the finite volume based methods.

The evolution matrix for \mathcal{D} is

$$\mathbf{E} = \begin{bmatrix} E_{0,0} & E_{0,1} & 1 & 0 \\ E_{1,0} & E_{1,1} & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (\text{C.10})$$

with

$$E_{0,0} = -\frac{2i\Delta t}{\Delta x} U \sin(k\Delta x),$$

$$E_{0,1} = -\frac{2i\Delta t}{\Delta x} H \sin(k\Delta x),$$

$$E_{1,0} = -\frac{6gi\Delta x\Delta t}{3\Delta x^2 - 2H^2(\cos(k\Delta x) - 1)} \sin(k\Delta x),$$

$$E_{1,1} = -\frac{2i\Delta t}{\Delta x} U \sin(k\Delta x).$$

While for \mathcal{W} the evolution matrix is

$$\mathbf{E} = \begin{bmatrix} E_{0,0} & E_{0,1} & 0 & E_{0,3} \\ E_{1,0} & E_{1,1} & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (\text{C.11})$$

with

$$E_{0,0} = 1 - \frac{\Delta t}{\Delta x} \left(-\frac{6gi\Delta x\Delta t}{3\Delta x^2 - 2H^2(\cos(k\Delta x) - 1)} \sin(k\Delta x) \right) H \frac{i \sin(k\Delta x)}{2}$$

$$- \frac{\Delta t}{\Delta x} U \left(i \sin(k\Delta x) - \frac{\Delta t}{\Delta x} U (\cos(k\Delta x) - 1) \right),$$

$$E_{0,1} = - \frac{\Delta t}{\Delta x} \left(H \frac{i \sin(k\Delta x)}{2} \left[1 - \frac{2i\Delta t}{\Delta x} U \sin(k\Delta x) \right] \right. \\ \left. - U \left[\frac{\Delta t}{\Delta x} H (\cos(k\Delta x) - 1) \right] \right),$$

$$E_{0,3} = - \frac{\Delta t}{\Delta x} H \frac{i \sin(k\Delta x)}{2},$$

$$E_{1,0} = - \frac{6gi\Delta x\Delta t}{3\Delta x^2 - 2H^2(\cos(k\Delta x) - 1)} \sin(k\Delta x),$$

$$E_{1,1} = - \frac{2i\Delta t}{\Delta x} U \sin(k\Delta x).$$

C.3 Consistency Results

The consistency results for FDVM₁, FDVM₂, FDVM₃, \mathcal{D} and \mathcal{W} are provided here. The consistency results for FEVM₂ were provided in Chapter 4.

To demonstrate the consistency of FDVM₁, FDVM₂ and FDVM₃ it is sufficient to demonstrate that all the lowest order terms of the Taylor series of $\mathbf{E} - e^{i\omega^\pm \Delta t} \mathbf{I}$ contain a factor of Δt . Since the results are similar for ω^- and ω^+ only the results for ω^+ are provided. All the Taylor series terms are presented in Tables C.9 and C.10 for FDVM₁, Table C.11 for FDVM₂ and Tables C.12 and C.13 for FDVM₃.

From Tables C.9-C.13 it is clear that all the lowest order terms of the Taylor series of $\mathbf{E} - e^{i\omega^\pm \Delta t} \mathbf{I}$ for the FDVM contain a factor of Δt . Therefore, FDVM₁, FDVM₂ and FDVM₃ are consistent numerical methods.

To demonstrate that \mathcal{D} and \mathcal{W} are consistent it is sufficient to demonstrate that all the lowest order terms of the Taylor series of $\mathbf{E}^{(2 \times 2)} - e^{i\omega^\pm \Delta t} \mathbf{I}$ contain a factor of Δt . Since the results are similar for ω^- and ω^+ only the results for ω^+ are presented. The results for ω^+ are displayed in Table C.14 for \mathcal{D} and Table

Element	Lowest Order Δx Term of $\mathbf{E} - e^{i\omega^+\Delta t}\mathbf{I}$ for FDVM ₁		
	$Fr < -1$	$-1 < Fr < 1$	$Fr > 1$
$E_{0,0} - e^{i\omega^+\Delta t}$	$\frac{1}{2}k^2U\Delta t\Delta x$	$-\frac{1}{2}\sqrt{gH}k^2\Delta t\Delta x$	$-\frac{1}{2}k^2U\Delta t\Delta x$
$E_{0,1}$	$\frac{1}{2}gHk^2\Delta t\Delta x$	$\frac{3+\beta}{4\beta^2}ik^3\Delta t\Delta x^2$	$\frac{1}{2}gHk^2\Delta t\Delta x$
$E_{1,0}$	$-\frac{1}{2}\sqrt{gH}k^2\Delta t\Delta x$	$-\frac{1}{2}\sqrt{gH}k^2\Delta t\Delta x$	$-\frac{1}{2}\sqrt{gH}k^2\Delta t\Delta x$
$E_{1,1} - e^{i\omega^+\Delta t}$	$\frac{1}{2}k^2U\Delta t\Delta x$	$-\frac{1}{2}\sqrt{gH}k^2\Delta t\Delta x$	$-\frac{1}{2}k^2U\Delta t\Delta x$

Table C.9: Lowest order Δx term of the Taylor series for the elements of $\mathbf{E} - e^{i\omega^+\Delta t}\mathbf{I}$ for FDVM₁. Here $\beta = 3 + k^2H^2$.

C.15 for \mathcal{W} . Note that \mathcal{D} and \mathcal{W} do not depend on Fr .

From Tables C.14 and C.15 it can be seen that all the lowest order terms of the Taylor series of $\mathbf{E}^{(2 \times 2)} - e^{i\omega^\pm\Delta t}\mathbf{I}$ for the FDM contain a factor of Δt . Therefore, \mathcal{D} and \mathcal{W} are consistent numerical methods.

Element	Lowest Order Δt Term of $\mathbf{E} - e^{i\omega^+ \Delta t} \mathbf{I}$ for FDVM ₁
$E_{0,0} - e^{i\omega^+ \Delta t}$	$\frac{\sqrt{3gH\beta} + 3U}{\beta} ik\Delta t$
$E_{0,1}$	$-\frac{3}{\beta} ik\Delta t$
$E_{1,0}$	$\left(-gH + \frac{3U^2}{\beta}\right) ik\Delta t$
$E_{1,1} - e^{i\omega^+ \Delta t}$	$\frac{\sqrt{3gH\beta} - 3U}{\beta} ik\Delta t$

Table C.10: Lowest order Δt term of the Taylor series for the elements of $\mathbf{E} - e^{i\omega^+ \Delta t} \mathbf{I}$ for FDVM₁ for all values of Fr . Here $\beta = 3 + k^2 H^2$.

Element	Lowest Order Terms of $\mathbf{E} - e^{i\omega^+ \Delta t} \mathbf{I}$ for FDVM ₂	
	Δx	Δt
$E_{0,0} - e^{i\omega^+ \Delta t}$	$-\frac{i(27 + 9H^2k^2 + H^4k^4)}{12\beta^2} U k^3 \Delta t \Delta x^2$	$\frac{\sqrt{3gH\beta} + 3U}{\beta} ik\Delta t$
$E_{0,1}$	$\frac{3 + \beta}{4\beta^2} ik^3 \Delta t \Delta x^2$	$-\frac{3}{\beta} ik\Delta t$
$E_{1,0}$	$-\left(gH + \frac{3U^2}{\beta} + \frac{9U^2}{\beta^2}\right) \frac{k^3}{12} \Delta t \Delta x^2$	$\left(-gH + \frac{3U^2}{\beta}\right) ik\Delta t$
$E_{1,1} - e^{i\omega^+ \Delta t}$	$\frac{-9 + H^2k^2\beta}{\beta^2} \frac{k^3}{12} iU \Delta t \Delta x^2$	$\frac{\sqrt{3gH\beta} - 3U}{\beta} ik\Delta t$

Table C.11: Lowest order terms of the Taylor series for the elements of $\mathbf{E} - e^{i\omega^+ \Delta t} \mathbf{I}$ for FDVM₂ for all values of Fr . Here $\beta = 3 + k^2 H^2$.

Element	Lowest Order Δx Term of $\mathbf{E} - e^{i\omega^+\Delta t}\mathbf{I}$ for FDVM ₃		
	$Fr < -1$	$-1 < Fr < 1$	$Fr > 1$
$E_{0,0} - e^{i\omega^+\Delta t}$	$\frac{1}{12}k^4U\Delta t\Delta x^3$	$-\frac{1}{12}\sqrt{gH}k^4\Delta t\Delta x^3$	$-\frac{1}{12}k^4U\Delta t\Delta x^3$
$E_{0,1}$	$\frac{1}{4\beta}iUk^5\Delta t^2\Delta x^3$	$\frac{\sqrt{gH}}{4\beta}ik^5\Delta t^2\Delta x^3$	$-\frac{1}{4\beta}iUk^5\Delta t^2\Delta x^3$
$E_{1,0}$	$\frac{1}{12}gHk^4\Delta t^2\Delta x^3$	$-\frac{1}{12}\sqrt{gH}k^4\Delta t\Delta x^3$	$-\frac{1}{12}gHk^4\Delta t^2\Delta x^3$
$E_{1,1} - e^{i\omega^+\Delta t}$	$\frac{1}{12}k^4U\Delta t\Delta x^3$	$-\frac{1}{12}\sqrt{gH}k^4\Delta t\Delta x^3$	$-\frac{1}{12}k^4U\Delta t\Delta x^3$

Table C.12: Lowest order Δx term of the Taylor series for the elements of $\mathbf{E} - e^{i\omega^+\Delta t}\mathbf{I}$ for FDVM₃. Here $\beta = 3 + k^2H^2$.

Element	Lowest Order Δt Term of $\mathbf{E} - e^{i\omega^+\Delta t}\mathbf{I}$ for FDVM ₃	
$E_{0,0} - e^{i\omega^+\Delta t}$		$\frac{\sqrt{3gH\beta} + 3U}{\beta}ik\Delta t$
$E_{0,1}$		$-\frac{3}{\beta}ik\Delta t$
$E_{1,0}$		$\left(-gH + \frac{3U^2}{\beta}\right)ik\Delta t$
$E_{1,1} - e^{i\omega^+\Delta t}$		$\frac{\sqrt{3gH\beta} - 3U}{\beta}ik\Delta t$

Table C.13: Lowest order Δt term of the Taylor series for the elements of $\mathbf{E} - e^{i\omega^+\Delta t}\mathbf{I}$ for FDVM₃ for all values of Fr . Here $\beta = 3 + k^2H^2$.

Element	Lowest Order Terms of $\mathbf{E}^{(2 \times 2)} - e^{i\omega^+ \Delta t} \mathbf{I}$ for \mathcal{D}	
	Δx	Δt
$E_{0,0}^{(2 \times 2)} - e^{i\omega^+ \Delta t}$	$\frac{ik^3}{3} U \Delta t \Delta x^2$	$\sqrt{\frac{3gH}{\beta}} 2ik \Delta t$
$E_{0,1}^{(2 \times 2)}$	$\frac{iHk^3}{3} \Delta t \Delta x^2$	$-2Hik \Delta t$
$E_{1,0}^{(2 \times 2)}$	$\frac{ig(3 + \beta)}{2\beta^2} k^3 \Delta t \Delta x^2$	$-\frac{6igk}{\beta} \Delta t$
$E_{1,1}^{(2 \times 2)} - e^{i\omega^+ \Delta t}$	$\frac{ik^3}{3} U \Delta t \Delta x^2$	$\sqrt{\frac{3gH}{\beta}} 2ik \Delta t$

Table C.14: Lowest order terms of the Taylor series for the elements of $\mathbf{E}^{(2 \times 2)} - e^{i\omega^\pm \Delta t} \mathbf{I}$ for \mathcal{D} . Here $\beta = 3 + k^2 H^2$.

Element	Lowest Order Terms of $\mathbf{E}^{(2 \times 2)} - e^{i\omega^+ \Delta t} \mathbf{I}$ for \mathcal{W}	
	Δx	Δt
$E_{0,0}^{(2 \times 2)} - e^{i\omega^+ \Delta t}$	$\frac{ik^3}{6} U \Delta t \Delta x^2$	$\sqrt{\frac{3gH}{\beta}} ik \Delta t$
$E_{0,1}^{(2 \times 2)}$	$\frac{iHk^3}{6} \Delta t \Delta x^2$	$-Hik \Delta t$
$E_{1,0}^{(2 \times 2)}$	$\frac{ig(3 + \beta)}{2\beta^2} k^3 \Delta t \Delta x^2$	$-\frac{6igk}{\beta} \Delta t$
$E_{1,1}^{(2 \times 2)} - e^{i\omega^+ \Delta t}$	$\frac{ik^3}{3} U \Delta t \Delta x^2$	$\sqrt{\frac{3gH}{\beta}} 2ik \Delta t$

Table C.15: Lowest order terms of the Taylor series for the elements of $\mathbf{E}^{(2 \times 2)} - e^{i\omega^+ \Delta t} \mathbf{I}$ for \mathcal{W} . Here $\beta = 3 + k^2 H^2$.

Appendix D

Publications

This appendix lists the publications my research contributed to in chronological order. This list expands the earlier list in Chapter 1 by providing the abstracts for the paper in place of the summary. My contribution to the paper and the relevance of the paper to this thesis are also provided.

A Solution of the Conservation Law Form of the Serre Equations

Australia and New Zealand Industrial and Applied Mathematics Journal (2016)
C. Zoppou, S.G. Roberts and J. Pitt

Abstract:

The nonlinear and weakly dispersive Serre equations contain higher-order dispersive terms. These include mixed spatial and temporal derivative flux terms which are difficult to handle numerically. These terms can be replaced by an alternative combination of equivalent temporal and spatial terms, so that the Serre equations can be written in conservation law form. The water depth and new conserved quantities are evolved using a second-order finite-volume scheme. The remaining primitive variable, the depth-averaged horizontal velocity, is obtained by solving a second-order elliptic equation using simple finite differences. Using an analytical solution and simulating the dam-break problem, the proposed scheme is shown to be accurate, simple to implement and stable for a range of problems, including flows with steep gradients. It is only slightly more computationally

expensive than solving the shallow water wave equations.

My Contribution:

I produced an independent reproduction of the method of my coauthors, verifying their results. This method was used to compare the computational cost of solving the Serre equations and the Shallow Water Wave Equations (SWWE).

Relevance to Thesis:

The method described in this paper was the foundation of the second-order finite difference volume method whose results are reported in this thesis. The main extensions of this method during my research were the inclusion of varying bathymetry [15] and dry beds. The results of a linear analysis of the second-order finite difference volume method is provided in Chapter 4 and Appendix C. Since the linear analysis is performed for a completely submerged horizontal bed, these results apply to the method described in this paper. Finally, the second-order finite difference volume method was validated against analytic and forced solutions and experimental results in Chapter 5 and Chapter 6, respectively.

Numerical Solution of the Fully Non-Linear Weakly Dispersive Serre Equations for Steep Gradient Flows

Applied Mathematical Modelling (2017)

C. Zoppou, J. Pitt and S.G. Roberts

Abstract:

We demonstrate a numerical approach for solving the one-dimensional non-linear weakly dispersive Serre equations. By introducing a new conserved quantity the Serre equations can be written in conservation law form, where the velocity is recovered from the conserved quantities at each time step by solving an auxiliary elliptic equation. Numerical techniques for solving equations in conservative law form can then be applied to solve the Serre equations. We demonstrate how this is achieved. The system of conservation equations are solved using the finite volume method and the associated elliptic equation for the velocity is solved using a finite difference method. This robust approach allows us to accurately solve problems with steep gradients in the flow, such as those generated by discontinuities in the initial conditions.

The method is shown to be accurate, simple to implement and stable for a range of problems including flows with steep gradients and variable bathymetry.

My Contribution:

This paper was based on research produced by me in collaboration with my coauthors which implemented the methods, performed the dispersion analysis and produced the numerical solutions. These results were then written up by my coauthors.

Relevance to Thesis:

The results of the first-, second- and third-order finite difference volume methods described in this paper are the methods whose results are reported in this thesis. The results of a linear analysis for these methods can be found in Chapter 4 and Appendix C. A validation of some of these methods against analytic and forced solutions and experimental results can be found in Chapter 5 and Chapter 6, respectively.

The linear analysis of the dispersion properties of the methods in this paper was extended in this thesis by allowing for a non-zero background mean flow

velocity. The analytic solution validation of this paper was reproduced in this thesis and extended by studying the convergence and conservation properties of more quantities. The experimental results of the second-order finite difference volume method in this paper for the negative rectangular wave experiment [16] and periodic waves over a submerged bar experiment [17] were reproduced in this thesis.

Importance of Dispersion for Shoaling Waves

22nd International Congress on Modelling and Simulation (2017)

J. Pitt, C. Zoppou and S.G. Roberts

Abstract:

A tsunami has four main stages of its evolution; in the first stage the tsunami is generated, most commonly by seismic activity near subduction zones. The second stage is the tsunamis propagation through the ocean far from the coast, where variation in bathymetry is slight and gradual. The third stage is the shoaling and interaction of the tsunami with bathymetry as it approaches the coastline. Finally the tsunami reaches and inundates the shore. For our purposes the hydrodynamic models we are interested in deal with the final three stages of the evolution of a tsunami.

The propagation of a tsunami with wavelength λ through water that is H deep is well understood when $\lambda/H \leq 1/20$, which we call shallow water as noted by Sorensen (2006). The wavelengths for tsunamis range from a few to hundreds of kilometres, while the maximum water depth is 11km at the Marianas trench, so that most tsunamis occur in shallow water. This stage of tsunami behaviour is adequately modelled using the shallow water wave equations. Current research into tsunamis focuses around more complex approximations to the Euler equations for the third and fourth stages. In this paper we focused on the Serre equations as they are considered a very good model for fluid behaviour up to the shoreline, and they reduce to the shallow water wave equations for large wavelengths.

Although more complicated, the Serre equations provide a better description of the fluid behaviour than the shallow water wave equations and are therefore more computationally expensive to solve numerically. In particular for the methods of this work, we find that the Serre equations have a run-time 50% longer than our equivalent finite volume method for the shallow water wave equations in the one dimensional case. To simulate tsunamis as efficiently as possible it is important to know when using the more complicated Serre equations leads to more accurate predictions of the evolution of a tsunami than the shallow water wave equations. To investigate this we have numerically simulated a laboratory experiment of periodic waves propagating over a submerged bar, and the propagation of a small amplitude wave up a gradual linear slope using both the Serre and the shallow water wave equations.

The results of these simulations demonstrated that the Serre and shallow water

wave equations produce similar results for shoaling waves when the wavelength is large compared to the water depth. This is not surprising as this is the regime under which the shallow water wave equations are derived. However, outside this regime the shallow water wave equations are a poor model for wave shoaling and propagation, poorly approximating the shape and maximum height of waves. Furthermore we demonstrate that for steep waves generated by shoaling, the shallow water wave equations can underestimate the arrival time and amplitude of an incoming wave. These results suggest that for a tsunami it is sufficient to use the shallow water wave equations in stages two and some of stage three, even for large changes in bathymetry. Although dispersive equations such as the Serre equations are required to accurately capture fluid behaviour in stages three and four nearer to the coastline, particularly when wavelengths are short or waves are steep. Since the Serre equations represent only a moderate increase in run-times this suggests that our inundation models should be based on them.

My Contribution:

This paper was primarily produced by me in collaboration with my coauthors, based on research that I primarily undertook.

Relevance to Thesis:

The results of the second-order finite difference volume method for the experiments of Beji and Battjes [17] studying periodic waves over a submerged bar are reproduced in Chapter 6.

Behaviour of the Serre Equations in the Presence of Steep Gradients Revisited

Wave Motion (2018)

J.P.A. Pitt, C. Zoppou and S.G. Roberts

Abstract:

We use numerical methods to study the short term behaviour of the Serre equations in the presence of steep gradients because there are no known analytical solutions for these problems. In keeping with the literature we study a class of initial condition problems that are a smooth approximation to the initial conditions of the dam-break problem. This class of initial condition problems allow us to observe the behaviour of the Serre equations with varying steepness of the initial conditions. The numerical solutions of the Serre equations are justified by demonstrating that as the resolution increases they converge to a solution with little error in conservation of mass, momentum and energy independent of the numerical method. We observe and justify four different structures of the converged numerical solutions depending on the steepness of the initial conditions. Two of these structures were observed in the literature, with the other two not being commonly found in the literature. The numerical solutions are then used to assess how well the analytical solution of the shallow water wave equations captures the mean behaviour of the solution of the Serre equations for the dam-break problem in the short term. Lastly the numerical solutions are compared to asymptotic results in the literature to approximate the depth and location of the front of an undular bore.

My Contribution:

This paper was primarily produced by me in collaboration with my coauthors, based on research that I primarily undertook.

Relevance to Thesis:

The results of this paper are summarised in Chapter 2. This paper demonstrates the utility of using a finite volume based method to solve the Serre equations in the presence of steep gradients. Hence, the further development of these methods in this thesis.

Bibliography

- [1] L. Euler. Principes généraux du mouvement des fluides. *Mémoires de l'académie des sciences de Berlin*, 11:274–315, 1757.
- [2] C.L.M.H. Navier. Mémoire sur les lois du mouvement des fluides. *Mémoires de l'Académie Royale des Sciences de l'Institut de France*, 6:389–440, 1823.
- [3] G.G. Stokes. *Transactions of the Cambridge Philosophical Society*, 8:287, 1845.
- [4] A.J. Chorin. The numerical solution of the Navier-Stokes equations for an incompressible fluid. *Bulletin of the American Mathematical Society*, 73(6):928–931, 1967.
- [5] C. Taylor and P. Hood. Numerical solution of the Navier-Stokes equations using the finite element technique. *Computers & Fluids*, 1:73–100, 1973.
- [6] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *Journal of Computational Physics*, 131(2):267–279, 1997.
- [7] D. Lannes and P. Bonneton. Derivation of asymptotic two-dimensional time-dependent equations for surface water wave propagation. *Physics of Fluids*, 21(1):16601, 2009.
- [8] The Clawpack Development Team. Clawpack documentation, 2018. URL <http://www.clawpack.org/>.
- [9] X. Wang. Comcot, 2009. URL <http://223.4.213.26/archive/tsunami/cornell/comcot.htm>.
- [10] S.G. Roberts. ANUGA, 2018. URL <https://anuga.anu.edu.au/>.

- [11] C. Eskilsson and S. J. Sherwin. A triangular spectral/hp discontinuous galerkin method for modelling 2d shallow water equations. *International Journal for Numerical Methods in Fluids*, 45(6):605–623, 2004.
- [12] J. Grue, E.N. Pelinovsky, D. Fructus, T. Talipova, and C. Kharif. Formation of undular bores and solitary waves in the strait of Malacca caused by the 26 December 2004 Indian Ocean tsunami. *Journal of Geophysical Research: Oceans*, 113(C5):008, 2008.
- [13] J.T. Kirby, F. Shi, B. Tehranirad, J.C. Harris, and S.T. Grilli. Dispersive tsunami waves in the ocean: Model equations and sensitivity to dispersion and coriolis effects. *Ocean Modelling*, 62:39–55, 2013.
- [14] C. Zoppou. *Numerical Solution of the One-dimensional and Cylindrical Serre Equations for Rapidly Varying Free Surface Flows*. PhD thesis, Australian National University, Mathematical Sciences Institute, College of Physical and Mathematical Sciences, Australian National University, Canberra, ACT 2600, Australia, 2014.
- [15] C. Zoppou, J. Pitt, and S. Roberts. Numerical solution of the fully non-linear weakly dispersive Serre equations for steep gradient flows. *Applied Mathematical Modelling*, 48:70–95, 2017.
- [16] J.L. Hammack and H. Segur. The Korteweg-de Vries equation and water waves. Part 3. Oscillatory waves. *Journal of Fluid Mechanics*, 84(2):337–358, 1978.
- [17] S. Beji and J.A. Battjes. Numerical simulation of nonlinear wave propagation over a bar. *Coastal Engineering*, 23(1):1–16, 1994.
- [18] J.P.A. Pitt, C. Zoppou, and S.G. Roberts. Behaviour of the Serre equations in the presence of steep gradients revisited. *Wave Motion*, 76(1):61–77, 2018.
- [19] R.M. Sorensen. *Basic coastal engineering*. New York, NY Springer, 3rd edition, 2006.
- [20] F. Serre. Contribution à l'étude des écoulements permanents et variables dans les canaux. *La Houille Blanche*, 6:830–872, 1953.
- [21] C.H. Su and C.S. Gardner. Korteweg-de Vries equation and generalisations. III. Derivation of the Korteweg-de Vries equation and Burgers equation. *Journal of Mathematical Physics*, 10(3):536–539, 1969.

- [22] A.E. Green and P.M. Naghdi. A derivation of equations for wave propagation in water of variable depth. *Journal of Fluid Mechanics*, 78(2):237–246, 1976.
- [23] F.J. Seabra-Santos, D.P. Renouard, and A.M. Tempeville. Numerical and experimental study of the transformation of a solitary wave over a shelf or isolated obstacle. *Journal of Fluid Mechanics*, 176:117–134, 1981.
- [24] E. Barthélémy. Nonlinear shallow water theories for coastal waves. *Surveys in Geophysics*, 25(3):315–337, 2004.
- [25] P. Bonneton, F. Chazel, D. Lannes, F. Marche, and M. Tissier. A splitting approach for the fully nonlinear and weakly dispersive Green-Naghdi model. *Journal of Computational Physics*, 230(4):1479–1498, 2011.
- [26] J.A. Liggett. *Fluid Mechanics*. McGraw-Hill civil engineering series. McGraw-Hill Inc., New York, 1994.
- [27] O. Le Métayer, S. Gavrilyuk, and S. Hank. A numerical scheme for the Green-Naghdi model. *Journal of Computational Physics*, 229(6):2034–2045, 2010.
- [28] M. Li, P. Guyenne, F. Li, and L. Xu. High order well-balanced CDG-FE methods for shallow water waves by a Green-Naghdi model. *Journal of Computational Physics*, 257(1):169–192, 2014.
- [29] W. Choi and R. Camassa. Fully nonlinear internal waves in a two-fluid system. *Journal of Fluid Mechanics*, 396:1–36, 1999.
- [30] J.D. Carter and R. Cienfuegos. The kinematics and stability of solitary and cnoidal wave solutions of the Serre equations. *European Journal of Mechanics-B/Fluids*, 30(3):259–268, 2011.
- [31] Y.A. Li. Hamiltonian structure and linear stability of solitary waves of the Green-Naghdi equations. *Journal of Nonlinear Mathematical Physics*, 9:99–105, 2002.
- [32] G.A. El, R.H.J. Grimshaw, and N.F. Smyth. Unsteady undular bores in fully nonlinear shallow-water theory. *Physics of Fluids*, 18(2):027104, 2006.
- [33] D. Dutykh, D. Clamond, P. Milewski, and D. Mitsotakis. Finite volume and pseudo-spectral schemes for the fully nonlinear 1D Serre equations. *European Journal of Applied Mathematics*, 24(5):761–787, 2013.

- [34] D. Mitsotakis, B. Ilan, and D. Dutykh. On the Galerkin/finite-element method for the Serre equations. *Journal of Scientific Computing*, 61(1):166–195, 2014.
- [35] D. Mitsotakis, D. Dutykh, and D. Carter. On the nonlinear dynamics of the traveling-wave solutions of the Serre system. *Wave Motion*, 70(1):166–182, 2017.
- [36] J.S.A. do Carmo, J.A. Ferreira, L. Pinto, and G. Romanazzi. An improved Serre model: Efficient simulation and comparative evaluation. *Applied Mathematical Modelling*, 56:404–423, 2018.
- [37] V.A. Dougalis, A. Duran, M.A. Lopez-Marcos, and D.E. Mitsotakis. Numerical study of the stability of solitary waves of the Bona-Smith family of Boussinesq systems. *Journal of Nonlinear Science*, 17(6):569–607, 2007.
- [38] R. Cienfuegos, E. Barthélemy, and P. Bonneton. A fourth-order compact finite volume scheme for fully nonlinear and weakly dispersive Boussinesq-type equations. Part I: Model development and analysis. *International Journal for Numerical Methods in Fluids*, 51(11):1217–1253, 2006.
- [39] S.F. Bradford and B.F. Sanders. Finite volume schemes for the Boussinesq equations. In *Ocean Wave Measurement and Analysis (2001)*, pages 953–962. American Society of Civil Engineers, 2002.
- [40] A. G. Filippini, M. Kazolea, and M. Ricchiuto. A flexible genuinely nonlinear approach for nonlinear wave propagation, breaking and run-up. *Journal of Computational Physics*, 310:381–417, 2016.
- [41] J. Pitt. *A Second Order Well Balanced Hybrid Finite Volume and Finite Difference Method for the Serre Equations*. Honour’s thesis, Australian National University, Mathematical Sciences Institute, College of Physical and Mathematical Sciences, Australian National University, Canberra, ACT 2600, Australia, 2014.
- [42] P.M. Prenter. *Splines and Variational Methods*. A Wiley-Interscience publication. Wiley, New York, 1975.
- [43] K.E. Atkinson. *An introduction to numerical analysis*. John Wiley & Sons, Canada, 2nd edition, 1989.

- [44] P.L. Roe. Characteristic-based schemes for the Euler equations. *Annual Review of Fluid Mechanics*, 18(1):337–365, 1986.
- [45] B. Van Leer. Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection. *Journal of Computational Physics*, 23(3):276–299, 1977.
- [46] A. Harten. High resolution schemes for hyperbolic conservation laws. *Journal of Computational Physics*, 49(3):357–3935, 1983.
- [47] P.J. Davis and P. Rabinowitz. *Methods of Numerical Integration*. Computer science and applied mathematics. Academic Press, London, 2nd edition, 1984.
- [48] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C*. Cambridge University Press, London, 2nd edition, 2002.
- [49] A. Kurganov, S. Noelle, and G. Petrova. Semidiscrete central-upwind schemes for hyperbolic conservation laws and Hamilton-Jacobi equations. *Journal of Scientific Computing, Society for Industrial and Applied Mathematics*, 23(3):707–740, 2002.
- [50] E. Audusse, F. Bouchut, M. Bristeau, R. Klein, and B. Perthame. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *Journal of Scientific Computing, Society for Industrial and Applied Mathematics*, 25(6):2050–2065, 2004.
- [51] S. Gottlieb, C. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *Review, Society for Industrial and Applied Mathematics*, 43(1):89–112, 2001.
- [52] R. Courant, K. Friedrichs, and H. Lewy. On the partial difference equations of mathematical physics. *IBM Journal of Research and Development*, 11(2):215–234, 1967.
- [53] B. Gustafsson. The convergence rate for difference approximations to mixed initial boundary value problems. *Mathematics of Computation*, 29(130):396–406, 1975.
- [54] A. Kurganov and G. Petrova. A second-order well-balanced positivity preserving central-upwind scheme for the Saint-Venant system. *Communications in Mathematical Sciences*, 5(1):133–160, 2007.

- [55] S.D. Conte and C. De Boor. *Elementary numerical analysis: An algorithmic approach*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, 3rd edition, 1980.
- [56] P. Lax and R. Richtmyer. Survey of the stability of linear finite difference equations. *Communications on Pure and Applied Mathematics*, 9(2):267–293, 1956.
- [57] B. Koren. A robust upwind discretization method for advection, diffusion and source term. In C.B. Vreugdenhil and B. Koren, editors, *Numerical Methods for Advection-Diffusion Problems*, chapter 5, pages 117–138. Vieweg, Braunschweig, 1993.
- [58] A.T. Ippen and G. Kulin. The shoaling and breaking of the solitary wave. *Coastal Engineering Proceedings*, 1(5):4, 1954.
- [59] N.J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, 1996.
- [60] S. Beji and J.A. Battjes. Experimental investigation of wave propagation over a bar. *Coastal Engineering*, 19(1):151–162, 1993.
- [61] D. Lannes. *The Water Waves Problem: Mathematical Analysis and Asymptotics*, volume 188 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, 2013.
- [62] J. Pitt, C. Zoppou, and S.G Roberts. Importance of dispersion for shoaling waves. *Modelling and Simulation Society of Australia and New Zealand*, 22(1):1725–1730, 2017.
- [63] Y. Zhang, A.B. Kennedy, N. Panda, C. Dawson, and J.J. Westerink. Boussinesq-Green-Naghdi rotational water wave theory. *Coastal Engineering*, 73(1):13–27, 2013.
- [64] V. Roeber. *Boussinesq-type mode for nearshore wave processes in fringing reef environment*. PhD thesis, Department of Ocean and Resource Engineering, University of Hawaii, Manoa, Honolulu, HI, U.S.A, 2010.
- [65] C.E. Synolakis. The runup of solitary waves. *Journal of Fluid Mechanics*, 185:523–545, 1987.

- [66] A. Bollermann, S. Noelle, and M. Lukáčová-Medvidová. Finite volume evolution Galerkin methods for the shallow water equations with dry beds. *Communications in Computational Physics*, 10(2):371–404, 2011.
- [67] M. Tissier, P. Bonneton, F. Marche, F. Chazel, and D. Lannes. Serre Green-Naghdi modelling of wave transformation breaking and run-up using a high-order finite-volume finite-difference scheme. *Coastal Engineering Proceedings*, 1(32):1–13, 2011.
- [68] J.S.A. do Carmo, J.A. Ferreira, and L. Pinto. On the accurate simulation of nearshore and dam break problems involving dispersive breaking waves. *Wave Motion*, 85:125 – 143, 2019.
- [69] V. T. Chow. *Open-channel hydraulics*. Caldwell, New Jersey : Blackburn Press, 2008.
- [70] R.J. LeVeque. *Finite volume methods for hyperbolic problems*, volume 54 of *Cambridge Texts in Applied Mathematics*. Cambridge University Press, New York, 2002.