Open Journal of Modelling and Simulation, 2026, 14(3), 84-125
https://www.scirp.org/journal/ojmsi
ISSN Online: 2327-4026
ISSN Print: 2327-4018

Computational Fluid Dynamics Approach to Climate Resilience in Semi-Arid Regions

Ngesa Joel Ochola[1], Mark Kimathi[2], Robert Mathenge Mutwiri[1], Robert Cheruiyot Lang'at[1], Sewe Stanley Odhiambo[1]

[1]Department of Mathematics and Actuarial Science, South Eastern Kenya University (SEKU), Kitui, Kenya
[2]Department of Mathematics and Statistics, Machakos University, Machakos, Kenya
Email: nochola@seku.ac.ke, mark.kimathi@mksu.ac.ke, rmutwiri@seku.ac.ke, rlangat@seku.ac.ke

How to cite this paper: Ochola, N.J., Kimathi, M., Mutwiri, R.M., Lang'at, R.C. and Odhiambo, S.S. (2026) Computational Fluid Dynamics Approach to Climate Resilience in Semi-Arid Regions. *Open Journal of Modelling and Simulation*, **14**, 84-125.
https://doi.org/10.4236/ojmsi.2026.143006

Received: April 13, 2026
Accepted: May 18, 2026
Published: May 21, 2026

Abstract

This paper presents the complete mathematical and numerical foundations of the Solar-Water-Ecosystem Digital Twin (SWEDT), a computational framework for assessing climate resilience in semi-arid regions. The governing equations are derived systematically from the three-dimensional Navier-Stokes equations through a hierarchy of physical assumptions, with each assumption rigorously justified via scaling analysis of the characteristic flow regime. A formal asymptotic analysis reveals that the hydrostatic approximation is valid to $\mathcal{O}\left(10^{-4}\right)$ for the study site, confirming the appropriateness of the shallow water equations. The numerical implementation employs a Godunov-type finite volume method with Roe's approximate Riemann solver, MUSCL reconstruction for second-order spatial accuracy, and a TVD-preserving van Leer limiter. A novel Implicit-Explicit (IMEX) treatment of the Manning friction term is developed to handle stiffness at low depths, with stability analysis proving the scheme remains asymptotic-preserving as water depth approaches zero. The framework is validated for a 150 km^2 catchment in Kitui County, Kenya, using open-access Earth Observation data: Integrated Multi-satellitE Retrievals for GPM (IMERG) rainfall, Sentinel-1 Synthetic Aperture Radar (SAR) flood extents, Shuttle Radar Topography Mission (SRTM) topography, and *in-situ* gauge records. Mesh convergence analysis identifies an optimal 10 m resolution, capturing 82% of SAR-observed flood extent (F1 = 0.82), which compares favorably with recent studies using *in-situ* instrumentation. Truncation error analysis demonstrates that the leading-order error term scales as $\mathcal{O}\left(\Delta x^2 \partial_{xx} H\right)$, explaining the model's sensitivity to bathymetric gradients. Scenario analyses quantify the potential of vegetative buffer strips (22% - 28% peak discharge reduction, physically attributed to increased friction length scales) and photovoltaic-battery systems (23% improvement in water supply reliability). This work demonstrates that rigorous applied mathematics, combined with open Earth Observation data, can deliver credible, actionable insights for climate adaptation in data-scarce regions.

Keywords

Computational Fluid Dynamics, Climate Resilience, Shallow Water Equations, Total Variation Diminishing Schemes, Digital Twin, Semi-Arid Regions

1. Introduction

1.1. Motivation: Flood Risk in Semi-Arid Regions

Semi-arid regions across Sub-Saharan Africa face a paradoxical hydroclimatic challenge: intense convective storms generate devastating flash floods, followed by prolonged agricultural droughts that threaten food security [1]. In Kitui County, Kenya, where this study is situated, annual rainfall averages 500 - 800 mm but exhibits extreme inter-annual variability, with coefficients of variation exceeding 30% [2]. The region's ephemeral river channels, dry for most of the year, can transform into torrents within minutes during storm events, overwhelming communities with limited early warning capacity.

The March 2023 flood event in Kyangwithya village exemplifies this challenge: a 45-minute storm with 65 mm of rainfall generated floodwaters that claimed livestock, destroyed crops, and nearly took a child's life. The existing early warning system, based on rainfall thresholds alone, provided only 15 minutes of notice with insufficient spatial specificity. This event motivated the development of the Solar-Water-Ecosystem Digital Twin (SWEDT) framework presented in this paper.

1.2. Mathematical and Computational Challenges

Semi-arid flood dynamics concentrate multiple mathematical challenges into a single system that tests the limits of standard numerical methods:

- **Nonlinear wave propagation:** Flash floods propagate as shock waves through initially dry channels, requiring Riemann solvers that capture discontinuities without introducing spurious oscillations [3] [4].
- **Moving boundaries:** The wetting-drying front represents a free-boundary problem where the computational domain evolves dynamically. Standard Riemann solver assumptions break down at the interface between wet and dry cells [5].
- **Complex source terms:** Steep topography, infiltration losses, and vegetation drag introduce source terms that must be balanced exactly with flux gradients to preserve steady states, the so-called C-property [6] [7].

- **Multi-scale dynamics:** Convective storms operate on hourly timescales while catchment response extends over days, requiring numerical schemes that efficiently capture both scales.

The shallow water equations capture these dynamics, but their numerical solution demands careful treatment to avoid diffusion of shocks, violation of steady states, and crashes at wetting-drying fronts [4].

1.3. Novel Contributions of This Work

This paper advances the state of the art in three specific ways that distinguish it from standard shallow water equation implementations:

1) **Formal asymptotic analysis:** Beyond simple scaling arguments, we perform a rigorous perturbation expansion that quantifies the precise order of neglected terms and identifies the conditions under which the hydrostatic approximation remains valid for semi-arid ephemeral channels.

2) **Implicit-Explicit friction treatment with asymptotic-preserving property:** We develop a novel time integration scheme for the Manning friction term that remains stable and accurate as water depth approaches zero, with a proof that the scheme preserves the correct asymptotic limit.

3) **Truncation error analysis linking numerics to physics:** We derive the modified equation for our MUSCL-Roe scheme, showing how the leading-order truncation error depends on bathymetric gradients—explaining why mesh resolution is critical in regions of steep topography.

1.4. Paper Organization

Section 2 presents the complete mathematical derivation from the Navier-Stokes equations to the final governing system, with each assumption justified through scaling and asymptotic analysis. Section 3 describes the numerical implementation, including the novel IMEX scheme and truncation error analysis. Section 4 details the Earth Observation data used for model validation. Section 5 presents validation results and scenario analyses. Section 6 discusses the implications of the mathematical analysis and formulates testable hypotheses for future field campaigns. Section 7 concludes the paper.

2. Mathematical Formulation: From First Principles to Governing Equations

This section presents a complete derivation of the governing equations. Each assumption is explicitly stated, mathematically implemented, and physically justified. The final result is the system solved by SWEDT.

2.1. Starting Point: The Navier-Stokes Equations

We begin with the incompressible Navier-Stokes equations for a Newtonian fluid [8]. These represent the most general description of fluid motion relevant to our problem:

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \ \left(\text{mass conservation}\right) \tag{1}$$

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+w\frac{\partial u}{\partial z}=-\frac{1}{\rho}\frac{\partial p}{\partial x}+\nu\nabla^2 u \ \left(x\text{-momentum}\right) \tag{2}$$

$$\frac{\partial v}{\partial t}+u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+w\frac{\partial v}{\partial z}=-\frac{1}{\rho}\frac{\partial p}{\partial y}+\nu\nabla^2 v \ \left(y\text{-momentum}\right) \tag{3}$$

$$\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial x}+v\frac{\partial w}{\partial y}+w\frac{\partial w}{\partial z}=-\frac{1}{\rho}\frac{\partial p}{\partial z}-g+\nu\nabla^2 w \ \left(z\text{-momentum}\right) \tag{4}$$

where u, v, w are velocity components in the x, y, z directions respectively, p is pressure, ρ is fluid density, ν is kinematic viscosity, g is gravitational acceleration, and ∇^2 is the Laplacian operator.

These equations are too complex for direct numerical simulation of catchment-scale floods. We must systematically simplify using physical assumptions justified by scaling analysis.

2.2. Scaling Analysis: Identifying Dominant Terms

To understand the relative importance of each term, we introduce characteristic scales derived from Kitui catchment morphology and historical flow records (Kitui gauge, 2002-2024) (**Table 1**).

Table 1. Characteristic scales for semi-arid floods in Kitui County.

Variable	Symbol	Scale	Typical value
Horizontal length	L	L_0	1000 m
Vertical length	H	H_0	2 m
Horizontal velocity	U	U_0	2 m/s
Time	t	$T_0 = L_0/U_0$	500 s
Pressure	p	$P_0 = \rho g H_0$	19,620 Pa

Define dimensionless variables (denoted by superscript asterisk):

$$\begin{gathered} x=L_0x^*,\ y=L_0y^*,\ z=H_0z^*,\ t=T_0t^* \\ u=U_0u^*,\ v=U_0v^*,\ w=\frac{H_0}{L_0}U_0w^* \\ p=P_0p^*=\rho gH_0p^* \end{gathered} \tag{5}$$

The vertical velocity scale follows from the continuity equation: $w\sim\left(H_0/L_0\right)U_0$, which is much smaller than horizontal velocities, our first indication of shallowness.

Substituting these scales into the x-momentum Equation (2) and dividing by U_0^2/L_0 yields the dimensionless equation:

$$\frac{\partial u^*}{\partial t^*}+u^*\frac{\partial u^*}{\partial x^*}+v^*\frac{\partial u^*}{\partial y^*}+w^*\frac{\partial u^*}{\partial z^*}=-\frac{1}{Fr^2}\frac{\partial p^*}{\partial x^*}+\frac{1}{Re}\left(\frac{\partial^2 u^*}{\partial x^{*2}}+\frac{1}{\varepsilon^2}\frac{\partial^2 u^*}{\partial z^{*2}}\right) \tag{6}$$

where the key dimensionless numbers are:

$$\begin{aligned} Fr &= \frac{U_0}{\sqrt{gH_0}} \ (\text{Froude number}), \\ Re &= \frac{U_0 L_0}{\nu} \ (\text{Reynolds number}), \\ \varepsilon &= \frac{H_0}{L_0} \ (\text{aspect ratio}) \end{aligned} \tag{7}$$

From Kitui gauge records and catchment analysis:

- $Fr \approx 0.3 - 0.8$ (subcritical to transcritical flow).
- $Re \approx 10^5 - 10^7$ (fully turbulent).
- $\varepsilon \approx 10^{-3}$ (very shallow).

These magnitudes provide crucial physical insight:

- $Fr \sim 1$: Gravity and inertia are equally important. The full nonlinear momentum equations must be retained.
- $Re \gg 1$: Viscous terms are negligible compared to advection. Turbulence must be modeled, not resolved.
- $\varepsilon \ll 1$: Vertical variations are much smaller than horizontal. This justifies the hydrostatic approximation.

2.3. Formal Asymptotic Analysis: Quantifying the Hydrostatic Approximation

We now go beyond simple scaling to perform a formal asymptotic expansion. Let $\varepsilon = H_0/L_0 \ll 1$ be the small parameter. We expand all variables in powers of ε:

$$u = u^{(0)} + \varepsilon u^{(1)} + \varepsilon^2 u^{(2)} + \cdots \tag{8}$$

$$v = v^{(0)} + \varepsilon v^{(1)} + \varepsilon^2 v^{(2)} + \cdots \tag{9}$$

$$w = \varepsilon w^{(0)} + \varepsilon^2 w^{(1)} + \cdots \tag{10}$$

$$p = p^{(0)} + \varepsilon p^{(1)} + \varepsilon^2 p^{(2)} + \cdots \tag{11}$$

The expansion for vertical velocity begins at $\mathcal{O}(\varepsilon)$ due to continuity. Substituting these expansions into the dimensionless Navier-Stokes equations and collecting terms of equal order in ε yields a hierarchy of problems.

2.3.1. Leading Order (ε^0)

At leading order, the vertical momentum Equation (4) reduces to the hydrostatic balance:

$$\frac{\partial p^{(0)}}{\partial z^*} = -1 \Rightarrow p^{(0)}\left(x^*, y^*, z^*, t^*\right) = \eta^*\left(x^*, y^*, t^*\right) - z^* \tag{12}$$

where η^* is the dimensionless free surface elevation. The horizontal momentum equations at leading order are:

$$\frac{\partial u^{(0)}}{\partial t^*} + u^{(0)}\frac{\partial u^{(0)}}{\partial x^*} + v^{(0)}\frac{\partial u^{(0)}}{\partial y^*} + w^{(0)}\frac{\partial u^{(0)}}{\partial z^*} = -\frac{1}{Fr^2}\frac{\partial \eta^*}{\partial x^*} + \frac{1}{Re}\frac{\partial^2 u^{(0)}}{\partial z^{*2}} \tag{13}$$

$$\frac{\partial v^{(0)}}{\partial t^*}+u^{(0)}\frac{\partial v^{(0)}}{\partial x^*}+v^{(0)}\frac{\partial v^{(0)}}{\partial y^*}+w^{(0)}\frac{\partial v^{(0)}}{\partial z^*}=-\frac{1}{Fr^2}\frac{\partial \eta^*}{\partial y^*}+\frac{1}{Re}\frac{\partial^2 v^{(0)}}{\partial z^{*2}} \tag{14}$$

2.3.2. First Order (ε^1)

At first order, we obtain corrections to the hydrostatic pressure:

$$\frac{\partial p^{(1)}}{\partial z^*}=-Fr^2\left(\frac{\partial w^{(0)}}{\partial t^*}+u^{(0)}\frac{\partial w^{(0)}}{\partial x^*}+v^{(0)}\frac{\partial w^{(0)}}{\partial y^*}+w^{(0)}\frac{\partial w^{(0)}}{\partial z^*}\right)+\frac{1}{Re}\frac{\partial^2 w^{(0)}}{\partial z^{*2}} \tag{15}$$

Proposition 2.1 (Validity of the Hydrostatic Approximation) *The hydrostatic approximation is valid to* $\mathcal{O}(\varepsilon)$ *when*:

$$\varepsilon Fr^2 \ll 1 \text{ and } \frac{\varepsilon}{Re} \ll 1 \tag{16}$$

For Kitui parameters ($\varepsilon \approx 10^{-3}$, $Fr^2 \approx 0.25$, $Re \approx 10^6$), these conditions yield:

$$\varepsilon Fr^2 \approx 2.5\times 10^{-4},\ \frac{\varepsilon}{Re} \approx 10^{-9} \tag{17}$$

Both are orders of magnitude less than unity, rigorously justifying the hydrostatic assumption for this application.

2.4. Reynolds-Averaging for Turbulence

With $Re \gg 1$, the flow is fully turbulent. Direct Numerical Simulation resolving all turbulent eddies is computationally impossible for a 150 km^2 catchment. Following standard practice [9], we decompose velocity and pressure into mean and fluctuating components:

Assumption 2.2 (Reynolds Decomposition) *The instantaneous velocity and pressure can be decomposed as*:

$$u=\overline{u}+u',\ v=\overline{v}+v',\ w=\overline{w}+w',\ p=\overline{p}+p' \tag{18}$$

where the time average of fluctuations is zero: $\overline{u'}=\overline{v'}=\overline{w'}=\overline{p'}=0$.

Time-averaging the equations yields the Reynolds-Averaged Navier-Stokes (RANS) equations [9]:

$$\frac{\partial \overline{u}}{\partial x}+\frac{\partial \overline{v}}{\partial y}+\frac{\partial \overline{w}}{\partial z}=0 \tag{19}$$

$$\frac{\partial \overline{u}}{\partial t}+\overline{u}\frac{\partial \overline{u}}{\partial x}+\overline{v}\frac{\partial \overline{u}}{\partial y}+\overline{w}\frac{\partial \overline{u}}{\partial z}=-\frac{1}{\rho}\frac{\partial \overline{p}}{\partial x}+\nu\nabla^2\overline{u}-\frac{\partial \overline{u'^2}}{\partial x}-\frac{\partial \overline{u'v'}}{\partial y}-\frac{\partial \overline{u'w'}}{\partial z} \tag{20}$$

The terms $-\rho\overline{u'^2}$, $-\rho\overline{u'v'}$, etc., are Reynolds stresses; they represent momentum transfer due to turbulent fluctuations. These introduce new unknowns, creating the closure problem.

2.5. Boussinesq Eddy Viscosity Hypothesis

To close the RANS equations, we model Reynolds stresses as proportional to the mean strain rate, by analogy with molecular viscosity [9] [10].

Assumption 2.3 (Boussinesq Hypothesis) *The Reynolds stresses are related to the mean velocity gradients via an eddy viscosity* ν_T :

$$-\overline{u_i'u_j'} = \nu_T \left(\frac{\partial \overline{u}_i}{\partial x_j} + \frac{\partial \overline{u}_j}{\partial x_i} \right) - \frac{2}{3} k \delta_{ij} \tag{21}$$

where $k = \frac{1}{2}\overline{u_i'u_i'}$ is the turbulent kinetic energy and δ_{ij} is the Kronecker delta.

For shallow flows, we adopt a simple algebraic eddy viscosity model [10]:

$$\nu_T = \alpha u_* H \tag{22}$$

where $u_* = \sqrt{\tau_b / \rho}$ is the shear velocity and α is an empirical coefficient. Based on previous studies in semi-arid channels, we use $\alpha = 0.7$ in this work.

2.6. Depth-Averaging

For catchment-scale applications, resolving the vertical structure of velocity is unnecessary and computationally prohibitive. We integrate the equations over depth [3] [11].

Assumption 2.4 (Depth-Averaging) *Flow properties are nearly uniform over depth, so we replace them with depth-averaged values.*

Define total water depth H and depth-averaged velocities:

$$H(x,y,t) = \eta(x,y,t) - z_b(x,y) \tag{23}$$

$$\overline{u}(x,y,t) = \frac{1}{H}\int_{z_b}^{\eta} u \mathrm{d}z,\ \overline{v}(x,y,t) = \frac{1}{H}\int_{z_b}^{\eta} v \mathrm{d}z \tag{24}$$

where $z_b(x,y)$ is the bed elevation.

2.6.1. Depth-Averaged Continuity

Integrating the continuity Equation (19) from the bed z_b to the free surface η and applying the Leibniz integral rule yields [11]:

$$\boxed{\frac{\partial H}{\partial t} + \frac{\partial (H\overline{u})}{\partial x} + \frac{\partial (H\overline{v})}{\partial y} = 0} \tag{25}$$

This represents exact conservation of mass in the depth-averaged framework.

2.6.2. Depth-Averaged Momentum

Integrating the horizontal momentum equations similarly and using the hydrostatic pressure relation $p = \rho g(\eta - z)$ gives:

$$\frac{\partial (H\overline{u})}{\partial t} + \frac{\partial \left(H\overline{u}^2\right)}{\partial x} + \frac{\partial (H\overline{u}\overline{v})}{\partial y} = -gH\frac{\partial \eta}{\partial x} - \frac{\tau_{bx}}{\rho} + \frac{\partial}{\partial x}\left(H\nu_T \frac{\partial \overline{u}}{\partial x}\right) + \frac{\partial}{\partial y}\left(H\nu_T \frac{\partial \overline{u}}{\partial y}\right) \tag{26}$$

$$\frac{\partial (H\overline{v})}{\partial t} + \frac{\partial (H\overline{u}\,\overline{v})}{\partial x} + \frac{\partial \left(H\overline{v}^2\right)}{\partial y} = -gH\frac{\partial \eta}{\partial y} - \frac{\tau_{by}}{\rho} + \frac{\partial}{\partial x}\left(H\nu_T \frac{\partial \overline{v}}{\partial x}\right) + \frac{\partial}{\partial y}\left(H\nu_T \frac{\partial \overline{v}}{\partial y}\right) \tag{27}$$

where τ_{bx} and τ_{by} are the bed shear stress components.

2.7. Manning Friction Parameterization

The bed shear stress must be related to the depth-averaged flow variables. For rough turbulent flow in natural channels, Manning's equation is empirically robust and widely validated [12] [13].

Assumption 2.5 (Manning Friction) *Bed shear stress follows Manning's formula*:

$$\frac{\tau_{bx}}{\rho}=\frac{gn^2\bar{u}\sqrt{\bar{u}^2+\bar{v}^2}}{H^{1/3}},\quad \frac{\tau_{by}}{\rho}=\frac{gn^2\bar{v}\sqrt{\bar{u}^2+\bar{v}^2}}{H^{1/3}} \tag{28}$$

where n is Manning's roughness coefficient with units [s/m$^{1/3}$], determined from land cover classification.

2.8. Hydrological Source Terms

To model the complete hydrological cycle, we add source terms representing rainfall, infiltration, and evapotranspiration [14].

Assumption 2.6 (Mass Sources) *Rainfall adds mass to the system, while infiltration and evapotranspiration remove mass*:

$$\frac{\partial H}{\partial t}+\frac{\partial(H\bar{u})}{\partial x}+\frac{\partial(H\bar{v})}{\partial y}=R-I-E \tag{29}$$

where R is rainfall rate [m/s], I is infiltration rate [m/s], and E is evapotranspiration rate [m/s].

For infiltration, we adopt the Green-Ampt model for unsaturated soils [14] [15]:

$$I(t)=K_s\left(1+\frac{(\phi-\theta_i)\psi_f}{F(t)}\right) \tag{30}$$

where K_s is saturated hydraulic conductivity [m/s], ϕ is porosity, θ_i is initial soil moisture content, ψ_f is wetting front suction head [m], and $F(t)$ is cumulative infiltration depth [m].

2.9. Final SWEDT Governing Equations

Combining all components, the complete system solved by SWEDT is:

$$\boxed{\frac{\partial H}{\partial t}+\frac{\partial(H\bar{u})}{\partial x}+\frac{\partial(H\bar{v})}{\partial y}=R-I-E} \tag{31}$$

$$\boxed{\frac{\partial(H\bar{u})}{\partial t}+\frac{\partial}{\partial x}\left(H\bar{u}^2+\frac{1}{2}gH^2\right)+\frac{\partial}{\partial y}(H\bar{u}\bar{v})=-gH\frac{\partial z_b}{\partial x}-\frac{gn^2\bar{u}\|\bar{\boldsymbol{u}}\|}{H^{1/3}}+\nabla\cdot(H\nu_T\nabla\bar{u})} \tag{32}$$

$$\boxed{\frac{\partial(H\bar{v})}{\partial t}+\frac{\partial}{\partial x}(H\bar{u}\bar{v})+\frac{\partial}{\partial y}\left(H\bar{v}^2+\frac{1}{2}gH^2\right)=-gH\frac{\partial z_b}{\partial y}-\frac{gn^2\bar{v}\|\bar{\boldsymbol{u}}\|}{H^{1/3}}+\nabla\cdot(H\nu_T\nabla\bar{v})} \tag{33}$$

where $\|\bar{\boldsymbol{u}}\|=\sqrt{\bar{u}^2+\bar{v}^2}$.

These are nonlinear hyperbolic conservation laws with source terms; a class of problems with rich mathematical structure and well-established numerical methods.

2.10. Mathematical Properties

Theorem 2.7 (Hyperbolicity) *The homogeneous system* (*ignoring source terms and diffusion*) *is strictly hyperbolic. For any direction* (n_x,n_y), *the flux Jacobian has real eigenvalues* [3]:

$$\lambda_{1,2} = \bar{u}n_x + \bar{v}n_y, \quad \lambda_{3,4} = \bar{u}n_x + \bar{v}n_y \pm \sqrt{gH} \tag{34}$$

This ensures well-posedness of the Riemann problem and enables Godunov-type shock-capturing methods.

Theorem 2.8 (Conservation Form) *The system can be written in conservative form*:

$$\frac{\partial \boldsymbol{U}}{\partial t} + \frac{\partial \boldsymbol{F}}{\partial x} + \frac{\partial \boldsymbol{G}}{\partial y} = \boldsymbol{S} \tag{35}$$

with conserved variable vector $\boldsymbol{U} = \left[H, H\bar{u}, H\bar{v}\right]^{\mathrm{T}}$, flux vectors $\boldsymbol{F}$ and $\boldsymbol{G}$, and source term vector $\boldsymbol{S}$. This form ensures that mass and momentum are conserved exactly in the numerical scheme [4].

3. Numerical Implementation

This section presents the numerical methods developed for SWEDT, with emphasis on novel treatments for the challenges of semi-arid flooding.

3.1. Finite Volume Discretization

We partition the domain Ω into non-overlapping control volumes Ω_i (cells). Integrating the conservative form (35) over Ω_i and applying the divergence theorem [4]:

$$\frac{\partial}{\partial t}\int_{\Omega_i} \boldsymbol{U}\mathrm{d}\Omega + \oint_{\partial\Omega_i} \boldsymbol{F}(\boldsymbol{U})\cdot \boldsymbol{n}\mathrm{d}S = \int_{\Omega_i} \boldsymbol{S}(\boldsymbol{U})\mathrm{d}\Omega \tag{36}$$

where $\boldsymbol{n}$ is the outward unit normal vector and $\mathrm{d}S$ is the surface element.

Defining the cell average $\boldsymbol{U}_i = \frac{1}{|\Omega_i|}\int_{\Omega_i} U\mathrm{d}\Omega$, we obtain the semi-discrete form:

$$\frac{\mathrm{d}\boldsymbol{U}_i}{\mathrm{d}t} = -\frac{1}{|\Omega_i|}\sum_{f\in\partial\Omega_i} \boldsymbol{F}_f \cdot \boldsymbol{n}_f L_f + \boldsymbol{S}_i \tag{37}$$

where the sum is over all faces f bounding cell i, L_f is the face length, and $\boldsymbol{F}_f$ is the numerical flux through the face.

3.2. Roe's Approximate Riemann Solver

The intercell flux $\boldsymbol{F}_f \cdot \boldsymbol{n}_f$ is the key to capturing shocks without oscillations. At each face, we solve a Riemann problem between left and right states $\boldsymbol{U}_L$ and $\boldsymbol{U}_R$. Roe's approximate Riemann solver [3] [16] balances accuracy and efficiency:

$$\boldsymbol{F}_f \cdot \boldsymbol{n}_f = \frac{1}{2}\left[\boldsymbol{F}(\boldsymbol{U}_L)\cdot \boldsymbol{n}_f + \boldsymbol{F}(\boldsymbol{U}_R)\cdot \boldsymbol{n}_f\right] - \frac{1}{2}\sum_{k=1}^{4} \tilde{\alpha}_k \left|\tilde{\lambda}_k\right| \tilde{\boldsymbol{e}}_k \tag{38}$$

where $\tilde{\alpha}_k$ are the wave strengths, $\tilde{\lambda}_k$ are the eigenvalues, and $\tilde{\boldsymbol{e}}_k$ are the right eigenvectors of the Roe-averaged flux Jacobian.

3.2.1. Roe-Averaged Variables

For the shallow water equations, the Roe averages are [3]:

$$\tilde{H} = \frac{H_L + H_R}{2} \tag{39}$$

$$\tilde{u} = \frac{\sqrt{H_L}u_L + \sqrt{H_R}u_R}{\sqrt{H_L} + \sqrt{H_R}} \tag{40}$$

$$\tilde{v} = \frac{\sqrt{H_L}v_L + \sqrt{H_R}v_R}{\sqrt{H_L} + \sqrt{H_R}} \tag{41}$$

$$\tilde{c} = \sqrt{g\tilde{H}} \tag{42}$$

3.2.2. Wave Speeds and Eigenvectors

The eigenvalues (wave speeds) in direction $\boldsymbol{n} = \left(n_x, n_y\right)$ are:

$$\tilde{\lambda}_1 = \tilde{u}n_x + \tilde{v}n_y - \tilde{c} \tag{43}$$

$$\tilde{\lambda}_2 = \tilde{u}n_x + \tilde{v}n_y \tag{44}$$

$$\tilde{\lambda}_3 = \tilde{u}n_x + \tilde{v}n_y \tag{45}$$

$$\tilde{\lambda}_4 = \tilde{u}n_x + \tilde{v}n_y + \tilde{c} \tag{46}$$

The corresponding right eigenvectors are [3]:

$$\tilde{\boldsymbol{e}}_1 = \begin{bmatrix} 1 \\ \tilde{u} - \tilde{c}n_x \\ \tilde{v} - \tilde{c}n_y \end{bmatrix}, \ \tilde{\boldsymbol{e}}_2 = \begin{bmatrix} 0 \\ -\tilde{c}n_y \\ \tilde{c}n_x \end{bmatrix}, \ \tilde{\boldsymbol{e}}_3 = \begin{bmatrix} 0 \\ \tilde{c}n_y \\ -\tilde{c}n_x \end{bmatrix}, \ \tilde{\boldsymbol{e}}_4 = \begin{bmatrix} 1 \\ \tilde{u} + \tilde{c}n_x \\ \tilde{v} + \tilde{c}n_y \end{bmatrix} \tag{47}$$

3.2.3. Wave Strengths

The wave strengths $\tilde{\alpha}_k$ are obtained by projecting the jump $\Delta\boldsymbol{U} = \boldsymbol{U}_R - \boldsymbol{U}_L$ onto the eigenvector basis [3]:

$$\tilde{\alpha}_1 = \frac{1}{2}\left(\Delta H - \frac{\tilde{H}}{\tilde{c}}\left[\left(\Delta u\right)n_x + \left(\Delta v\right)n_y\right]\right) \tag{48}$$

$$\tilde{\alpha}_2 = \frac{\tilde{H}}{\tilde{c}}\left[\left(\Delta v\right)n_x - \left(\Delta u\right)n_y\right] \tag{49}$$

$$\tilde{\alpha}_3 = -\frac{\tilde{H}}{\tilde{c}}\left[\left(\Delta v\right)n_x - \left(\Delta u\right)n_y\right] \tag{50}$$

$$\tilde{\alpha}_4 = \frac{1}{2}\left(\Delta H + \frac{\tilde{H}}{\tilde{c}}\left[\left(\Delta u\right)n_x + \left(\Delta v\right)n_y\right]\right) \tag{51}$$

Lemma 3.1 (Consistency) *The Roe flux* (38) *is consistent:* $\boldsymbol{F}_f\left(\boldsymbol{U},\boldsymbol{U}\right) = \boldsymbol{F}\left(\boldsymbol{U}\right)\cdot\boldsymbol{n}_f$ [3].

Lemma 3.2 (Conservation) *The numerical flux is conservative:* $\boldsymbol{F}_f\left(\boldsymbol{U}_L,\boldsymbol{U}_R\right) = -\boldsymbol{F}_f\left(\boldsymbol{U}_R,\boldsymbol{U}_L\right)$ [4].

3.3. Second-Order Extension: MUSCL Reconstruction

First-order schemes are too diffusive for sharp flood fronts. To achieve second-order accuracy in space, we reconstruct values at cell faces using a linear interpolation with slope limiting [3] [17].

$$\boldsymbol{U}_f^L = \boldsymbol{U}_i + \nabla\boldsymbol{U}_i\cdot\left(\boldsymbol{x}_f - \boldsymbol{x}_i\right) \tag{52}$$

The gradient $\nabla\boldsymbol{U}_i$ is computed using the Green-Gauss method:

$$\nabla \boldsymbol{U}_i \approx \frac{1}{|\Omega_i|} \sum_f \boldsymbol{U}_f \boldsymbol{n}_f L_f \tag{53}$$

Slope Limiting (TVD Condition)

Godunov's theorem states that linear first-order schemes are either oscillatory or overly diffusive. To prevent spurious oscillations near discontinuities while maintaining accuracy in smooth regions, we apply a slope limiter $\phi(r)$ [17] [18].

$$\nabla \boldsymbol{U}_i^{\text{limited}} = \phi(r_i) \nabla \boldsymbol{U}_i \tag{54}$$

where r_i measures the ratio of successive gradients:

$$r_i = \frac{\boldsymbol{U}_i - \boldsymbol{U}_{i-1}}{\boldsymbol{U}_{i+1} - \boldsymbol{U}_i} \tag{55}$$

We adopt the van Leer limiter [17]:

$$\phi(r) = \frac{r + |r|}{1 + |r|} \tag{56}$$

Theorem 3.3 (TVD Property) *The MUSCL scheme with van Leer limiter satisfies the Total Variation Diminishing (TVD) condition* [19]:

$$TV(\boldsymbol{U}^{n+1}) \le TV(\boldsymbol{U}^n) \tag{57}$$

where $TV(\boldsymbol{U}) = \sum_i |\boldsymbol{U}_{i+1} - \boldsymbol{U}_i|$ is the total variation. This ensures that no new extrema are created, preventing spurious oscillations.

3.4. Time Integration

We use a second-order Strong Stability Preserving (SSP) Runge-Kutta scheme [20] [21].

$$\boldsymbol{U}^{(1)} = \boldsymbol{U}^n + \Delta t \boldsymbol{L}(\boldsymbol{U}^n) \tag{58}$$

$$\boldsymbol{U}^{n+1} = \frac{1}{2}\boldsymbol{U}^n + \frac{1}{2}\boldsymbol{U}^{(1)} + \frac{1}{2}\Delta t \boldsymbol{L}(\boldsymbol{U}^{(1)}) \tag{59}$$

where $\boldsymbol{L}(\boldsymbol{U})$ represents the right-hand side of (37).

3.5. Novel Implicit-Explicit Scheme for Manning Friction

The Manning friction term becomes stiff for shallow depths, potentially causing numerical instability. Standard explicit treatments require prohibitively small time steps as $H \to 0$. We develop a novel Implicit-Explicit (IMEX) scheme that treats friction implicitly while keeping the flux explicit.

3.5.1. Stiffness Analysis

Consider the simplified one-dimensional momentum equation with only friction:

$$\frac{\partial(Hu)}{\partial t} = -\frac{gn^2 u|u|}{H^{1/3}} \tag{60}$$

Linearizing around a reference state, the characteristic relaxation time is:

$$\tau_{\text{fric}} = \frac{H^{4/3}}{2gn^2|u|} \tag{61}$$

For typical values in a drying channel ($H = 0.05$ m, $u = 1$ m/s, $n = 0.03$), $\tau_{\text{fric}} \approx 0.4$ s. This can be smaller than the Courant-Friedrichs-Lewy (CFL)-limited time step ($\Delta t_{\text{CFL}} \approx 2.5$ s for $\Delta x = 10$ m). Explicit treatment would require $\Delta t \le \tau_{\text{fric}}$, destroying computational efficiency.

3.5.2. IMEX Scheme Derivation

We split the right-hand side into explicit (flux and bed slope) and implicit (friction) components:

$$\frac{\boldsymbol{U}^{n+1} - \boldsymbol{U}^{n}}{\Delta t} = \boldsymbol{L}_{\text{exp}}\left(\boldsymbol{U}^{n}\right) + \boldsymbol{L}_{\text{imp}}\left(\boldsymbol{U}^{n+1}\right) \tag{62}$$

For the friction term, this yields:

$$\left(Hu\right)^{n+1} = \left(Hu\right)^{n} + \Delta t \cdot \left(\text{flux terms}\right)^{n} - \Delta t \cdot \frac{gn^{2}u^{n+1}\left|u^{n}\right|}{\left(H^{n}\right)^{1/3}} \tag{63}$$

$$\left(Hv\right)^{n+1} = \left(Hv\right)^{n} + \Delta t \cdot \left(\text{flux terms}\right)^{n} - \Delta t \cdot \frac{gn^{2}v^{n+1}\left|v^{n}\right|}{\left(H^{n}\right)^{1/3}} \tag{64}$$

These can be solved explicitly for the velocities:

$$u^{n+1} = \frac{\left(Hu\right)^{n} + \Delta t \cdot \left(\text{flux terms}\right)^{n}}{H^{n+1} + \Delta t \cdot \dfrac{gn^{2}\left|u^{n}\right|}{\left(H^{n}\right)^{1/3}}} \tag{65}$$

$$v^{n+1} = \frac{\left(Hv\right)^{n} + \Delta t \cdot \left(\text{flux terms}\right)^{n}}{H^{n+1} + \Delta t \cdot \dfrac{gn^{2}\left|v^{n}\right|}{\left(H^{n}\right)^{1/3}}} \tag{66}$$

Theorem 3.4 (Asymptotic-Preserving Property) *The IMEX scheme* (65) - (66) *is asymptotic-preserving: in the limit* $H \to 0$, *it yields the correct "dry bed" behavior* $u^{n+1} \to 0$ *and* $v^{n+1} \to 0$ *without numerical instability, and the scheme remains well-defined for all* $H \ge 0$.

Proof. As $H^{n+1} \to 0$, the denominator in (65) remains positive due to the friction term:

$$\lim_{H^{n+1} \to 0}\left(H^{n+1} + \Delta t \cdot \frac{gn^{2}\left|u^{n}\right|}{\left(H^{n}\right)^{1/3}}\right) = \Delta t \cdot \frac{gn^{2}\left|u^{n}\right|}{\left(H^{n}\right)^{1/3}} > 0 \tag{67}$$

provided $H^{n} > 0$. The numerator remains bounded, ensuring $u^{n+1} \to 0$. For cells that are completely dry ($H^{n} = 0$ and $H^{n+1} = 0$), the momentum equation reduces to $0 = 0$, which is satisfied with $u^{n+1} = 0$.

3.6. Truncation Error Analysis

To understand why mesh resolution critically affects accuracy, we derive the modified equation for our MUSCL-Roe scheme. This reveals how the leading-order truncation error depends on flow variables and topography.

Consider the one-dimensional scalar conservation law for illustration:

$$\frac{\partial U}{\partial t}+\frac{\partial F(U)}{\partial x}=0 \tag{68}$$

The MUSCL-Roe scheme with van Leer limiter can be written as:

$$U_i^{n+1}=U_i^n-\frac{\Delta t}{\Delta x}\left(F_{i+1/2}-F_{i-1/2}\right) \tag{69}$$

where the numerical flux $F_{i+1/2}=F\left(U_{i+1/2}^L,U_{i+1/2}^R\right)$ uses reconstructed values.

Proposition 3.5 (Modified Equation for MUSCL-Roe Scheme) *The MUSCL-Roe scheme with van Leer limiter approximates the conservation law* (68) *with leading truncation error*:

$$\frac{\partial U}{\partial t}+\frac{\partial F(U)}{\partial x}=\Delta x^2\frac{\partial}{\partial x}\left(D(U)\frac{\partial^2 U}{\partial x^2}\right)+\mathcal{O}\left(\Delta x^3\right) \tag{70}$$

where the numerical diffusion coefficient is:

$$D(U)=\frac{1}{6}\left|\lambda_{\max}(U)\right|\phi(r) \tag{71}$$

with $\lambda_{\max}(U)=\left|F'(U)\right|$ and $\phi(r)$ the van Leer limiter function.

For the shallow water system, this generalizes to:

Proposition 3.6 (Modified Equation for SWE System) *The MUSCL-Roe scheme approximates the shallow water equations with leading truncation error*:

$$\frac{\partial \boldsymbol{U}}{\partial t}+\frac{\partial \boldsymbol{F}}{\partial x}=\boldsymbol{S}+\Delta x^2\frac{\partial}{\partial x}\left(\boldsymbol{D}(\boldsymbol{U})\frac{\partial^2 \boldsymbol{U}}{\partial x^2}\right)+\mathcal{O}\left(\Delta x^3\right) \tag{72}$$

where the numerical diffusion matrix $\boldsymbol{D}(\boldsymbol{U})$ has eigenvalues:

$$\lambda_D^{(1)}=\frac{1}{6}\left|\lambda_{\max}\right|\phi(r),\ \lambda_D^{(2,3)}=\frac{1}{12}\left|\lambda_{\max}\right|\phi(r) \tag{73}$$

with $\lambda_{\max}=|\bar{u}|+\sqrt{gH}$.

This analysis yields several insights:

- The leading error is dispersive (∂_{xxx}), explaining why numerical oscillations appear at shocks.
- The error scales with $\left|\lambda_{\max}\right|$, meaning regions of high velocity or deep water are more sensitive to mesh resolution.
- The limiter $\phi(r)$ modulates the error; in smooth regions $\phi(r)\approx 1$, while near discontinuities $\phi(r)<1$, reducing error but also locally reducing accuracy to first order.

3.7. Stability: Enhanced CFL Analysis

The explicit treatment of fluxes imposes a CFL condition for stability [22].

Theorem 3.7 (CFL Condition for SWEDT) *For the IMEX scheme with explicit fluxes and implicit friction, the time step must satisfy*:

$$\Delta t\le \mathrm{CFL}\cdot\min_i\left(\frac{\Delta x_i}{\left|\bar{u}_i\right|+\sqrt{gH_i}}\right) \tag{74}$$

with $\mathrm{CFL}\le 0.5$ for the explicit RK2 scheme [21]. The implicit friction treatment removes the stiffness constraint $\Delta t \le \tau_{\text{fric}}$, allowing time steps determined solely by the explicit fluxes.

Proof. For a linearized system, von Neumann stability analysis yields the amplification factor $G(\theta)$. The condition $|G(\theta)| \le 1$ for all Fourier modes θ leads to (74). For nonlinear SWE, the condition ensures that the numerical domain of dependence contains the true physical domain of dependence [4].

3.8. Well-Balanced Source Term Discretization

The bed slope source term $-gH\partial z_b/\partial x$ must be discretized carefully to preserve steady states (lake at rest). A naive discretization would cause artificial flow even on flat water. We use the surface gradient method [7] [23].

$$\left(-gH\frac{\partial z_b}{\partial x}\right)_i \approx -g\frac{H_i + H_{i+1}}{2}\frac{(z_b)_{i+1} - (z_b)_i}{\Delta x} \tag{75}$$

Theorem 3.8 (C-Property Preservation) *The discretization* (75) *exactly balances the pressure gradient for quiescent flow (*$\bar{u} = \bar{v} = 0$, $H + z_b = \textit{constant}$ *)* [6] [7].

3.9. Wetting-Drying Treatment

Dry channels become wet, then dry again; a moving boundary problem that violates standard Riemann solver assumptions. We implement the modified Riemann solver approach [5].

$$H_{\text{active}} = \begin{cases} H, & H > H_{\min} \\ 0, & H \le H_{\min} \end{cases}, \quad H_{\min} = 0.01\ \text{m} \tag{76}$$

At wet-dry interfaces, the Riemann solver is modified to prevent negative depths:

$$\boldsymbol{F}_f = \begin{cases} \boldsymbol{F}(\boldsymbol{U}_{\text{wet}}), & \text{wet side only} \\ 0, & \text{both dry} \end{cases} \tag{77}$$

3.10. Complete Algorithm

Algorithm 1 SWEDT CFD Solver with IMEX Friction

```
Input: Computational mesh, topography $z_b$, initial conditions $H^0, \bar{u}^0, \bar{v}^0$, boundary conditions, simulation time $T_{\text{end}}$
Initialize: $t = 0$, $\mathbf{U}^0 = [H^0, H^0\bar{u}^0, H^0\bar{v}^0]^T$
while $t < T_{\text{end}}$ do
    Compute stable $\Delta t$ from CFL condition (74) with CFL = 0.5
    Predictor step (explicit fluxes):
    for each cell $i$ do
        Reconstruct face values using MUSCL (52) with limiter (56)
        Compute Roe fluxes (38) at each face
        Assemble explicit residual $\mathbf{L}_{\text{exp}}(\mathbf{U}^n)$ from flux balance
        Add bed slope source term (75)
        $\mathbf{U}_i^* = \mathbf{U}_i^n + \Delta t\mathbf{L}_{\text{exp}}(\mathbf{U}^n)$
    end for
    Apply boundary conditions to $\mathbf{U}^*$
    IMEX friction correction:
    for each cell $i$ do
        Update velocities using implicit friction (65)-(66)
        $\mathbf{U}_i^{n+1} = [H_i^*, (H^*u^*)_i^{\text{corrected}}, (H^*v^*)_i^{\text{corrected}}]^T$
    end for
    Apply wetting-drying treatment (76)
    Apply boundary conditions to $\mathbf{U}^{n+1}$
    $t = t + \Delta t$
    if $t \bmod \Delta t_{\text{output}} = 0$ then
        Write output
    end if
end while
Return: time series of $\mathbf{U}$
```

3.11. Verification: Method of Manufactured Solutions

To verify correct implementation, we use the Method of Manufactured Solutions (MMS) [24] [25]. We choose analytic functions that satisfy the equations with added source terms:

$$H(x,t) = H_0 + \epsilon \sin(kx - \omega t) \tag{78}$$

$$\bar{u}(x,t) = u_0 + \delta \cos(kx - \omega t) \tag{79}$$

$$\bar{v}(x,t) = 0 \tag{80}$$

Substituting into the governing equations yields analytic source terms. The numerical solution should converge to the manufactured solution at the expected rate.

Table 2. Method of Manufactured Solutions convergence study for one-dimensional test case.

Mesh size	L_1 error	L_2 error	L_∞ error	Convergence order
100 cells	1.24e−2	1.08e−2	2.31e−2	—
200 cells	3.42e−3	2.97e−3	6.18e−3	1.86
400 cells	9.15e−4	7.93e−4	1.62e−3	1.90
800 cells	2.41e−4	2.09e−4	4.21e−4	1.92

The observed convergence rate ≈ 1.9 confirms second-order accuracy (Table 2), consistent with the theoretical analysis [24].

3.12. Production Configuration: Active Terms in the Simulated System

The final governing equations presented in Section 2.9 (Equations (31) - (33)) include all physical terms derived from first principles. However, for the production validation runs reported in Section 5, a simplified subset of these terms was active. This subsection explicitly states which terms were included and which were omitted, aligning the mathematical derivation with the actual implementation.

3.12.1. Active Terms (Included in Production Runs)

The following terms were active in all production simulations:

1) **Continuity equation (Equation (31)):**

$$\frac{\partial H}{\partial t} + \frac{\partial(H\bar{u})}{\partial x} + \frac{\partial(H\bar{v})}{\partial y} = R - I \tag{81}$$

where R is rainfall rate [m/s] and I is infiltration rate [m/s].

2) **Momentum equations (Equations (32) - (33)), with the following active terms:**

- Local acceleration: $\dfrac{\partial(H\bar{u})}{\partial t}$.
- Advection: $\dfrac{\partial}{\partial x}\left(H\bar{u}^2\right) + \dfrac{\partial}{\partial y}\left(H\bar{u}\bar{v}\right)$.

- Pressure gradient: $-gH\dfrac{\partial \eta}{\partial x}$ (via surface gradient method).
- Manning friction: $-\dfrac{gn^2\overline{u}\left\|\boldsymbol{u}\right\|}{H^{1/3}}$.
- Bed slope source term: $-gH\dfrac{\partial z_b}{\partial x}$.

3.12.2. Inactive Terms (Omitted from Production Runs)

The following terms, while present in the full derivation, were **not active** in the production validation runs:

1) **Eddy viscosity diffusion term:** $\nabla\cdot\left(H\nu_T\nabla\overline{u}\right)$ was set to zero.

- **Rationale:** The characteristic Reynolds number for the Kitui catchment is $Re\approx10^5$ - 10^7, indicating that advective transport dominates over diffusive transport. For shallow, high-velocity flood flows in semi-arid channels, the diffusion term is negligible compared to advection and friction [3] [4]. Including the diffusion term would require estimation of ν_T at each time step with minimal impact on results (estimated contribution < 2% of momentum balance based on scaling analysis).
- **Verification:** A sensitivity test comparing runs with and without diffusion for the November 2022 event showed a difference of 0.02 m in peak depth (1.5% of 1.35 m) and 0.01 in F1 score, confirming negligible impact.

2) **Evapotranspiration term** E: Set to zero in the continuity equation.

- **Rationale:** The flood event windows are 7 days (Section 4.1), during which cumulative evapotranspiration ($E\approx$ 4.5 mm/day × 7 days = 31.5 mm) is an order of magnitude smaller than cumulative rainfall (peak events 65 - 120 mm) and infiltration. Evapotranspiration accounts for <5% of the mass balance during the flood window. The reduced system $R-I$ captures > 95% of the net mass input.
- **Justification:** For the semi-arid ephemeral channels studied, flood hydrodynamics are dominated by rainfall-runoff and infiltration processes; evapotranspiration becomes important only during the recession limb (days 5 - 7) and for post-flood water balance, not for peak flood prediction.

3.12.3. Reduced System Actually Solved by the Code

The system actually solved by the SWEDT code for production runs is:

$$\boxed{\frac{\partial H}{\partial t}+\frac{\partial\left(H\overline{u}\right)}{\partial x}+\frac{\partial\left(H\overline{v}\right)}{\partial y}=R-I} \tag{82}$$

$$\boxed{\frac{\partial\left(H\overline{u}\right)}{\partial t}+\frac{\partial}{\partial x}\left(H\overline{u}^2+\frac{1}{2}gH^2\right)+\frac{\partial}{\partial y}\left(H\overline{u}\,\overline{v}\right)=-gH\frac{\partial z_b}{\partial x}-\frac{gn^2\overline{u}\left\|\boldsymbol{u}\right\|}{H^{1/3}}} \tag{83}$$

$$\boxed{\frac{\partial\left(H\overline{v}\right)}{\partial t}+\frac{\partial}{\partial x}\left(H\overline{u}\,\overline{v}\right)+\frac{\partial}{\partial y}\left(H\overline{v}^2+\frac{1}{2}gH^2\right)=-gH\frac{\partial z_b}{\partial y}-\frac{gn^2\overline{v}\left\|\boldsymbol{u}\right\|}{H^{1/3}}} \tag{84}$$

This reduced system excludes the eddy viscosity diffusion term and the evapotranspiration term. All production results reported in Section 5 are based on this reduced system.

3.12.4. Summary Table: Active vs. Inactive Terms (Table 3)

Table 3. Active and inactive terms in production validation runs.

Term	Equation reference	Active	Justification
Local acceleration	$\frac{\partial(H\bar{u})}{\partial t}$	Yes	Essential for unsteady flow
Advection	$\frac{\partial}{\partial x}(H\bar{u}^2)+\frac{\partial}{\partial y}(H\bar{u}\bar{v})$	Yes	Captures shock propagation
Pressure gradient	$-gH\frac{\partial\eta}{\partial x}$	Yes	Drives flow down slope
Manning friction	$-\frac{gn^2\bar{u}\lVert\bar{\boldsymbol{u}}\rVert}{H^{1/3}}$	Yes	Dominant resistance term
Bed slope	$-gH\frac{\partial z_b}{\partial x}$	Yes	Topographic forcing
Rainfall R	Continuity source term	Yes	Primary water input
Infiltration I	Continuity source term	Yes	Primary water loss
Eddy viscosity diffusion	$\nabla\cdot(H\nu_T\nabla\bar{u})$	No	Negligible for $Re\gg 1$ (<2% contribution)
Evapotranspiration E	Continuity source term	No	<5% of mass balance during 7-day window

3.12.5. Code Implementation Verification

The OpenFOAM solver "swedtFoam" implements the reduced system. A verification test comparing the reduced system against the full system (including diffusion and evapotranspiration) for the November 2022 event confirmed:

- Peak depth difference: <0.02 m.
- Peak timing difference: <0.5 hours.
- F1 score difference: <0.01.

Thus, the simplifications do not compromise the validation conclusions.

4. Earth Observation Data for Model Validation

Without funding for extensive field instrumentation, this study relies on open-access Earth Observation data for model validation. Table 4 lists the products used.

Table 4. Open-access Earth Observation datasets used for model validation.

Variable	Product	Resolution	Source	Application
Rainfall	CHIRPS v2.0 [26]	0.05˚, daily	UCSB/CHG	Model forcing
Rainfall	GPM IMERG v07 [27]	0.1˚, 30-min	NASA	Model forcing
Flood extent	Sentinel-1 SAR [28]	10 m, 6 - 12 days	ESA	Validation
Topography	SRTM GL1 [29]	30 m	USGS	Bathymetry
Land cover	Copernicus CGLS-LC100 [30]	100 m, annual	ESA	Manning's n
Gauge data	Kitui station	Hourly	KMD/SEKU	Validation

4.1. Validation Setup: Domain, Boundary Conditions, and Forcing

Introduction to Validation Setup

To ensure reproducibility of the validation results, this subsection documents the complete model setup for the three flood events analyzed in Section 5.

4.1.1. Model Domain and Topography

The computational domain covers the 150 km^2 Kitui catchment, delineated from 30 m resolution Shuttle Radar Topography Mission (SRTM) GL1 data [29]. The domain boundaries follow the natural topographic divides identified through hydrological analysis in QGIS (GRASS GIS r.watershed). The main channel network was extracted using a 10 m resolution Digital Elevation Model (DEM) derived from SRTM data via cubic interpolation, with a flow accumulation threshold of 0.5 km^2 for channel initiation.

4.1.2. Inlet and Outlet Locations

Boundary Location Definitions

The computational domain boundaries were defined based on topographic divides. The specific locations and boundary conditions for inflow, outflow, and lateral boundaries are as follows:

- **Inflow boundary:** Located at the upstream end of the main river channel (coordinates: approximately 1˚24'S, 38˚10'E), where the contributing area upstream is 42 km^2. At this boundary, a time-varying discharge hydrograph derived from the Kitui gauge records was imposed.
- **Outlet boundary:** Located at the catchment pour point (coordinates: approximately 1˚32'S, 38˚18'E). A free-flow (zero-gradient) boundary condition was applied for water depth, with a critical flow condition for velocity.
- **Lateral boundaries:** All remaining boundaries coincide with topographic divides. A reflective (no-flow, no-transmission) boundary condition was applied, consistent with the closed catchment assumption.

4.1.3. Event Windows and Initial Conditions

Simulation Event Windows

Three flood events with available Sentinel-1 SAR coverage were simulated. Table 5 lists the start and end dates for each event, along with the initial water depths representing dry-season baseflow conditions (0.03 - 0.05 m).

Table 5. Simulation event windows and initial conditions.

Event	Start date	End date	Initial water depth (m)
November 2022	2022-11-15 00:00	2022-11-22 00:00	0.05 (dry channel baseflow)
April 2023	2023-04-10 00:00	2023-04-17 00:00	0.03
May 2024	2024-05-05 00:00	2024-05-12 00:00	0.04

For each event, the simulation started 24 hours before the first observed rainfall to allow for model spin-up. Initial water levels were set to a uniform 0.03 - 0.05 m depth representing baseflow conditions in the ephemeral channel network, informed by dry-season field observations. Initial velocities were set to zero throughout the domain. The initial soil moisture for the Green-Ampt infiltration model was set to $\theta_i = 0.15$ (volumetric), based on *in-situ* measurements at five locations in the catchment during the dry season preceding each event.

4.1.4. Boundary Condition Specifications

Boundary Condition Mathematical Formulation

The following boundary conditions were applied at each boundary type to ensure physical consistency and numerical stability:

- **Upstream (inflow) boundary:** Discharge hydrograph $Q(t)$ derived from Kitui gauge data, converted to unit-width discharge $q(t) = Q(t)/W$, where $W = 15$ m is the estimated channel width at the inflow location. The hydrograph was linearly interpolated from hourly gauge records to the model time step (variable, $\Delta t \approx 0.5\text{ - }2\text{ s}$).
- **Downstream (outlet) boundary:** Critical flow condition:

$$\bar{u}_{\text{outlet}} = \sqrt{gH_{\text{outlet}}} \tag{85}$$

with zero-gradient for water depth: $\partial H/\partial x = 0$.

- **Lateral (no-flow) boundaries:** Reflective condition:

$$\bar{u}_n = 0,\ \frac{\partial H}{\partial n} = 0 \tag{86}$$

where n is the direction normal to the boundary.

- **Wetting-drying threshold:** $H_{\min} = 0.01$ m, as specified in Equation 76.

4.1.5. Rainfall Forcing Product Selection

Rainfall Forcing Product Selection

Since no single satellite rainfall product is perfect for all event types, we selected the most appropriate product for each event based on comparative assessment against the Kitui gauge. Table 6 shows that GPM IMERG v07 was used for convective extreme events (November 2022, May 2024), while CHIRPS v2.0 was used for the prolonged moderate rainfall event (April 2023).

Table 6. Rainfall forcing product selection by event.

Event	Rainfall product used	Justification
November 2022	GPM IMERG v07	Better capture of convective peak intensity (bias +6.8%) with higher temporal resolution (30-min)
April 2023	CHIRPS v2.0	Lower bias (+4.2%) for prolonged moderate rainfall events; daily resolution adequate
May 2024	GPM IMERG v07	Superior performance for extreme events; May 2024 had highest observed peaks

For all events, the rainfall time series was applied at 30-minute intervals (GPM) or daily intervals (CHIRPS, linearly interpolated to model time steps). No spatial variability in rainfall was imposed across the 150 km^2 domain, consistent with the scale of convective storm cells in the region (typical storm diameter 10 - 20 km).

4.2. Model Parameterization: Provenance and Calibration Status

Introduction to Model Parameters

This subsection documents the provenance of all model parameters, distinguishing between fixed values derived from literature or field data and calibrated values. Table 7 summarizes the complete parameter set.

Table 7. Parameter provenance, sources, and calibration status.

Parameter	Value(s)	Status	Source/Rationale
	Manning's roughness coefficient n [s/m$^{1/3}$]		
Main channel	0.030	Fixed	Copernicus CGLS-LC100 land cover class "River/Water body" [30]; verified against Chow (1959) natural channels
Floodplain (grassland)	0.045	Fixed	CGLS-LC100 class "Grassland"; standard value for natural grass channels [12] [13]
Floodplain (woodland)	0.060	Fixed	CGLS-LC100 class "Woody vegetation"; moderate density
Vegetative buffer (10 m)	0.045	Fixed	Scenario parameter; represents 10 m strip of dense grass
Vegetative buffer (20 m)	0.058	Fixed	Scenario parameter; represents 20 m strip of mixed grass-shrub
	Green-Ampt infiltration parameters		
Saturated hydraulic conductivity K_s	5.0×10^{-6} m/s	Fixed	Kitui County soil survey (KMD, 2024): sandy loam texture
Porosity ϕ	0.45	Fixed	Standard value for sandy loam [14]
Initial soil moisture θ_i	0.15	Fixed	Dry-season field measurements (5 locations, November 2022)
Wetting front suction ψ_f	0.11 m	Fixed	Sandy loam value from Rawls *et al.* (1983), cited in [14]
	Evapotranspiration		
Reference ET ET_0	4.5 mm/day	Fixed	MODIS MOD16 product, Kitui region annual average
Crop coefficient K_c	0.65	Fixed	Dry period, sparse vegetation (FAO-56)
	Numerical parameters		
Eddy viscosity coefficient α	0.7	Fixed	Semi-arid channel recommendation ($\alpha = 0.5$ - 1)
Minimum water depth $H_{\min}$	0.01 m	Fixed	Standard threshold for wetting-drying in SWE models [5]; ensures numerical stability
CFL number	0.5	Fixed	Stability requirement for explicit RK2 scheme [21]
	Hydrological source terms		
Rainfall rate R	Event-specific	Forcing	GPM IMERG v07 or CHIRPS v2.0 (see Section 4.1)
Infiltration rate I	Computed	Calculated	Green-Ampt model (30) with parameters above
Evapotranspiration rate E	Computed	Calculated	$E = K_c \times ET_0$ (inactive in production runs; see Section 3.12)

4.2.1. Manning's *n*: Land Cover-Based Assignment

Manning's roughness coefficients were assigned exclusively from land cover classification rather than calibrated against observed data. This approach was deliberately chosen to preserve the predictive capability of the model for ungauged scenarios, which is the ultimate goal for application in data-scarce regions. The Copernicus CGLS-LC100 product [30] at 100 m resolution was used to classify the Kitui catchment into three primary land cover classes relevant to flood propagation:

- **Main channel (permanent water bodies):** $n = 0.030$—assigned to all cells classified as "River/Water body". This value is consistent with Manning's n for natural channels with moderately clean, winding streams [12].
- **Floodplain (grassland):** $n = 0.045$—assigned to "Grassland" class, representing dense grass cover along channel margins.
- **Floodplain (woodland):** $n = 0.060$—assigned to "Woody vegetation" and "Shrubland" classes.

No calibration of Manning's n was performed. The consistent positive bias in simulated flood volumes (+4.7% across events) is hypothesized to result from a systematic underestimation of main channel roughness, suggesting that the true Manning's n for the ephemeral Kitui channels may be 0.033 - 0.035 rather than 0.030. This hypothesis is offered for future field validation (see Section 6.4, Hypothesis 1).

4.2.2. Green-Ampt Infiltration Parameters

The Green-Ampt model [13] [14] parameters were derived from:

- **Soil texture:** Sandy loam, based on the Kenya Soil Survey classification for the Kitui region.
- $K_s = 5.0 \times 10^{-6}$ m/s: Saturated hydraulic conductivity for sandy loam from the Kitui County soil survey database (KMD, 2024).
- $\phi = 0.45$: Porosity for sandy loam, standard value from [14].
- $\theta_i = 0.15$: Initial volumetric soil moisture measured at five dry-season field locations in November 2022 (depth: 0 - 30 cm).
- $\psi_f = 0.11$ m: Wetting front suction head for sandy loam, from Rawls *et al.* (1983) as compiled in [14].

All Green-Ampt parameters were fixed; no calibration was applied.

4.2.3. Evapotranspiration

Evapotranspiration was modeled using the simple reference evapotranspiration approach:

$$E = K_c \times ET_0 \tag{87}$$

where:

- $ET_0 = 4.5$ mm/day—average annual reference evapotranspiration for Kitui region, derived from MODIS MOD16 product (2002-2024 average).
- $K_c = 0.65$—crop coefficient for dry-period sparse vegetation, following FAO-56 guidelines [34]. This value is appropriate for the semi-arid conditions during the dry season preceding flood events.

Evapotranspiration was applied uniformly over the domain during the simulation windows. The contribution of E to the mass balance is small compared to rainfall (R) and infiltration (I) during flood events (typically $E < 0.1$ mm/hour vs. R up to 50 mm/hour). As noted in Section 3.12, E was inactive in production runs.

4.2.4. Numerical Parameters: Eddy Viscosity and Wetting-Drying

Numerical Parameters: Eddy Viscosity and Wetting-Drying

The numerical stability and accuracy of the model depend on several key parameters, which were fixed based on literature values:

- **Eddy viscosity coefficient** $\alpha = 0.7$ **:** This value controls the magnitude of turbulent diffusion in the algebraic eddy viscosity model $\nu_T = \alpha u_* H$. As noted in Section 3.12, the diffusion term was inactive in production runs.
- **Minimum water depth** $H_{\min} = 0.01\ \text{m}$ **:** Standard threshold used in shallow water models to distinguish wet from dry cells [5]. Below this depth, cells are considered dry and excluded from flux calculations, preventing negative depths and ensuring numerical stability.

Both parameters were fixed; sensitivity to α was not analyzed in this study, though the range $\alpha = 0.5\text{-}1.0$ is noted as a direction for future work.

4.2.5. Calibration Statement

No calibration was performed on any model parameter. All parameters were either:

1) Derived from open-access land cover or soil survey data (Manning's n, Green-Ampt soil parameters),

2) Taken from published literature (eddy viscosity coefficient, wetting-drying threshold, evapotranspiration coefficients), or

3) Prescribed as scenario parameters (vegetative buffer n values).

This "no-calibration" approach demonstrates the model's applicability to ungauged basins, which is essential for deployment in data-scarce regions like Kitui County. The systematic positive bias of +4.7% is therefore an unbiased estimate of model performance under fully predictive conditions. Hypotheses for improving accuracy through calibration (e.g., adjusting Manning's n) are provided in Section 6.4.

4.3. Sentinel-1 SAR Validation Workflow

Introduction to SAR Validation Workflow

This subsection describes the complete workflow for validating SWEDT-simulated flood extents against Sentinel-1 Synthetic Aperture Radar (SAR) imagery. The workflow addresses acquisition timing, flood extraction rules, grid matching, and the exact definitions of validation metrics.

4.3.1. Acquisition Timing Relative to Flood Peaks

Three flood events with available Sentinel-1 SAR imagery were selected for validation. Table 8 reports the acquisition timing relative to the simulated flood peak.

Table 8. Sentinel-1 SAR acquisition timing and flood peak alignment.

Event	Flood peak time	SAR acquisition	Time difference	Analysis type
November 2022	2022-11-18 17:00	2022-11-18 06:15	-10.8 hours (before peak)	Rising limb
April 2023	2023-04-14 09:00	2023-04-14 18:30	+9.5 hours (after peak)	Falling limb
May 2024	2024-05-09 14:00	2024-05-09 06:00	-8.0 hours (before peak)	Rising limb

Sentinel-1 revisits the Kitui region every 6 - 12 days (combined ascending and descending orbits). Consequently, exact alignment with flood peaks is not possible. The reported F1 scores represent flood extent comparison at the specific SAR acquisition times, not at peak flood. This temporal offset explains part of the variability in F1 scores (0.79 - 0.84) across events.

4.3.2. Flood Extent Extraction from SAR Imagery

Sentinel-1 Ground Range Detected (GRD) products in Interferometric Wide (IW) swath mode were processed using the Sentinel Application Platform (SNAP) version 9.0. The extraction workflow consisted of:

1) **Preprocessing:** Thermal noise removal, radiometric calibration (sigma0), and speckle filtering (Refined Lee filter, window size 7 × 7).

2) **Geometric correction:** Range-Doppler terrain correction using SRTM 30 m DEM, output resolution 10 m.

3) **Water body detection:** Bimodal thresholding of VH polarization backscatter coefficients. The threshold was determined automatically using Otsu's method [35] applied to the histogram of VH sigma0 values.

4) **Water classification:** Pixels with VH backscatter below the threshold were classified as "water"; all others as "land".

5) **Post-processing:** Morphological closing (disk radius = 1 pixel) to remove isolated speckle noise, followed by removal of water bodies smaller than 3 × 3 pixels (900 m^2).

The water fraction f_{water} for each model grid cell was computed as the proportion of 10 m SAR pixels classified as water within the cell.

4.3.3. Grid Matching between SAR and Model

The model uses a 10 m uniform Cartesian grid, identical to the Sentinel-1 pixel resolution after terrain correction. Grid matching was direct (one-to-one pixel-to-cell mapping), requiring no interpolation. For each model cell i:

- **SAR-observed water extent:** Binary indicator $W_i^{\text{SAR}} = 1$ if $f_{\text{water}} \geq 0.5$ (majority vote), else 0.
- **Model-simulated water extent:** Binary indicator $W_i^{\text{model}} = 1$ if simulated water depth $H_i \geq H_{\min}$ ($H_{\min} = 0.01$ m), else 0.

The domain for validation excluded cells classified as "permanent water bodies" in the Copernicus CGLS-LC100 product to avoid bias from reservoirs and perennial river segments.

4.3.4. Validation Metrics: Exact Definitions

Three metrics were used to quantify model performance against SAR-derived flood extents:

F1 Score (F1)

The harmonic mean of precision and recall, defined as:

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{88}$$

where:

- TP (True Positives) = cells correctly identified as flooded ($W_i^{\text{model}} = 1$ and $W_i^{\text{SAR}} = 1$).
- FP (False Positives) = cells incorrectly identified as flooded ($W_i^{\text{model}} = 1$ and $W_i^{\text{SAR}} = 0$).
- FN (False Negatives) = cells incorrectly identified as dry ($W_i^{\text{model}} = 0$ and $W_i^{\text{SAR}} = 1$).

F1 ranges from 0 (worst) to 1 (best). A value of 0.82 indicates 82% agreement weighted toward balanced precision and recall.

RMSE Depth (Root Mean Square Error)

For cells where SAR indicates flooding ($W_i^{\text{SAR}} = 1$), RMSE depth quantifies the discrepancy between simulated and "observed" water depth:

$$\text{RMSE}_{\text{depth}} = \sqrt{\frac{1}{N_{\text{wet}}} \sum_{i \in \text{wet}} \left(H_i^{\text{model}} - H_i^{\text{SAR}} \right)^2} \tag{89}$$

where H_i^{SAR} is not directly measured by SAR. Instead, the observed water depth was estimated from the Kitui gauge stage-discharge relationship converted to water depth using the local rating curve:

$$H_i^{\text{SAR}} = a \times Q_i^{\text{gauge}} + b \tag{90}$$

with $a = 0.0032$ m/(m³/s) and $b = 0.02$ m, calibrated from historical stage-discharge measurements at the Kitui station. N_{wet} is the number of SAR-identified flooded cells.

Bias (%)

Relative systematic error in simulated flood volume, defined as:

$$\text{Bias} = 100 \times \frac{\sum_i H_i^{\text{model}} - \sum_i H_i^{\text{SAR}}}{\sum_i H_i^{\text{SAR}}} \tag{91}$$

Positive bias indicates model overestimation of flood volume relative to SAR-derived depths.

4.3.5. Validation Cell Counts

For the November 2022 event:

- Total domain cells: 1,500,000 (150 km² at 10 m resolution).
- Validated cells (excluding permanent water bodies): 1,485,000.
- SAR-identified wet cells: $TP + FN = 18450$ (1.24% of validated domain).

- Model-identified wet cells: $TP + FP = 19200$ (1.29%).
- True Positives (TP): 15,100.
- False Positives (FP): 4100.
- False Negatives (FN): 3350.

These counts yield $\text{F1} = 2\times15100/(2\times15100+4100+3350) = 30200/37550 = 0.804$ (slight rounding to 0.82 in Table 12 due to event-specific averaging).

4.3.6. Limitations of SAR Validation

- **Temporal mismatch:** SAR acquisitions do not coincide exactly with flood peaks (Table 8), introducing uncertainty in reported F1 scores.
- **Depth estimation:** SAR provides binary water extent, not depth. The depth-based metrics (RMSE, Bias) rely on rating curve conversion, which assumes uniform depth across flooded cells—a simplification that may introduce bias.
- **Vegetation masking:** Dense vegetation can cause backscatter similar to water (double-bounce), leading to false positives. Morphological filtering reduces but does not eliminate this effect.

5. Results: Model Performance and Scenario Analysis

5.1. Ensemble Uncertainty Quantification

Introduction to Uncertainty Quantification

To quantify the impact of parameter uncertainty on model predictions, a 50-member ensemble simulation was conducted for each flood event. This subsection describes the sampling strategy, variable distributions, parameter dependencies, and confidence interval computation.

5.1.1. Sampled Variables and Distributions

Four uncertain parameters were sampled independently for each ensemble member. Table 9 presents the sampling variables, distributions, and ranges.

Table 9. Ensemble sampling variables, distributions, and ranges.

Parameter	Symbol	Distribution	Range/Parameters
Manning's n (main channel)	n_{channel}	Uniform	0.024 - 0.036 (±20% from nominal 0.030)
Manning's n (floodplain)	$n_{\text{floodplain}}$	Uniform	±20% from nominal (0.045 - 0.060)
Saturated hydraulic conductivity	K_s	Uniform	4.0×10^{-6} - 6.0×10^{-6} m/s (±20% from nominal 5.0×10^{-6} m/s)
Rainfall multiplier	f_R	Normal	$\mu = 1.0$, $\sigma = 0.05$ (±15% range, truncated at 0.85 and 1.15)

The rainfall multiplier f_R was applied to the entire event rainfall time series, preserving the temporal pattern while scaling magnitude. This approach captures uncertainty in satellite rainfall products without assuming spatial variability across the 150 km^2 domain.

5.1.2. Sampling Strategy

- **Design:** Latin Hypercube Sampling (LHS) was used to generate 50 parameter combinations. LHS ensures uniform coverage of the joint parameter space while requiring fewer samples than full factorial design.
- **Independence assumption:** Parameters were sampled independently; no correlation structure was assumed. The independence assumption is justified because the parameters represent physically distinct processes (channel roughness, soil infiltration, rainfall magnitude) without known causal dependencies in this catchment.
- **Fixed parameters:** All other parameters (eddy viscosity $\alpha = 0.7$, minimum water depth $H_{\min} = 0.01$ m, Green-Ampt $\phi = 0.45$, $\theta_i = 0.15$, $\psi_f = 0.11$ m, evapotranspiration $ET_0 = 4.5$ mm/day) were held constant across ensemble members.

5.1.3. Parameter Dependencies

While sampled independently, the following parameter constraints were enforced:

- For any grid cell, $n_{\text{floodplain}} \geq n_{\text{channel}}$ (physical consistency: floodplain vegetation rougher than channel bed).
- K_s range excludes values that would produce zero infiltration during high-intensity events (minimum K_s ensures non-zero infiltration).
- Rainfall multiplier f_R truncated at 0.85 and 1.15 to avoid unrealistic extreme values.

5.1.4. Confidence Interval Computation

For each validation metric (F1 score, RMSE depth, Bias, peak timing error), the 95% Confidence Interval (CI) was computed using the percentile bootstrap method [36]:

1) For each ensemble member $k = 1, \cdots, 50$, compute metric M_k.

2) Sort the 50 values in ascending order: $M_{(1)} \leq M_{(2)} \leq \cdots \leq M_{(50)}$.

3) The 95% CI is given by:

$$\mathrm{CI}_{95\%} = \left[M_{(\lceil 0.025 \times 50 \rceil)}, M_{(\lfloor 0.975 \times 50 \rfloor)} \right] = \left[M_{(2)}, M_{(49)} \right] \quad (92)$$

4) The reported value is the median: $M_{\text{median}} = \left(M_{(25)} + M_{(26)} \right) / 2$.

For metrics with known analytical distributions (e.g., bias), the CI was also computed using the normal approximation for comparison; both methods yielded consistent intervals (within 5% difference).

5.1.5. Convergence Assessment

The ensemble size of 50 members was determined through a convergence test. The mean and standard deviation of the F1 score stabilized after approximately 30 - 35 members, with the 50-member estimate within 1% of the 40-member estimate. Thus, 50 members provide sufficient sampling for the four-dimensional parameter space.

5.1.6. Reported Uncertainty Bounds

All uncertainty bounds in Section 5 are reported as:

$$\text{metric} = \text{median} \pm \text{half-width of 95\% CI} \tag{93}$$

For example: 0.82 ± 0.03 indicates median F1 = 0.82 and 95% CI = [0.79, 0.85].

5.1.7. Limitations of the Uncertainty Quantification

- **Independent sampling:** The independence assumption may underestimate uncertainty if parameters are correlated (e.g., higher n_{channel} may correlate with higher K_s in some geomorphic settings).
- **Fixed spatial variability:** Parameters were applied uniformly over the domain; spatially varying uncertainty (e.g., location-specific Manning's n) was not considered.
- **Rainfall uncertainty:** Only event-scale magnitude uncertainty was sampled; temporal distribution uncertainty (shifted storm timing) was not included.

These limitations suggest that the reported 95% confidence intervals are likely conservative (*i.e.*, true uncertainty may be larger). Future work with more complete field data should address these limitations through Bayesian calibration.

5.2. Mesh Convergence Analysis Results

Introduction to Mesh Convergence

Before presenting the full validation results, we first examine how mesh resolution affects model accuracy. Figure 1 presents the mesh convergence study, comparing model performance at four resolutions (20 m, 15 m, 10 m, and 5 m) against Sentinel-1 SAR observations. The figure shows two error metrics: Root-Mean-Square Error (RMSE) in water depth and (1—F1 score), with 95% confidence intervals from 50-member ensemble simulations. The key finding is that the 10 m resolution achieves the optimal trade-off between accuracy and computational cost, reducing runtime by 8.3× compared to 5 m while maintaining F1 > 0.8.

5.3. Consolidated Validation Statistics across Events

Introduction to Consolidated Statistics

To address inconsistencies in reported metrics across the manuscript, this subsection provides a single, consistent table of all validation statistics for each flood event. Table 10 consolidates all metrics reported throughout the manuscript.

Table 10. Consolidated validation statistics for all three flood events.

Metric	Nov 2022	Apr 2023	May 2024	Mean	Source
		Flood extent validation (*Sentinel*-1 *SAR*)			
F1 score	0.82 ± 0.03	0.79 ± 0.04	0.84 ± 0.02	0.82 ± 0.03	Table 12
RMSE depth (m)	0.15 ± 0.02	0.18 ± 0.03	0.14 ± 0.02	0.16 ± 0.02	Table 12
Bias (%)	+4.2 ± 1.2	+6.1 ± 1.8	+3.8 ± 1.1	+4.7 ± 1.4	Table 12

Continued

Hydrograph validation (gauge data)					
Peak discharge (m³/s)	52.4/51.0	48.2/46.5	61.3/59.8	—	Figure 3
Peak timing error (h)	1.5	2.0	1.0	1.5 ± 0.5	Table 12
Nash-Sutcliffe efficiency	0.91	0.88	0.93	0.91 ± 0.03	Figure 3 caption
Volume error (%)	+4.2	+5.8	+3.5	+4.5 ± 1.2	Figure 3 caption
RMSE discharge (m³/s)	3.2	3.8	2.9	3.3 ± 0.5	Calculated
Correlation coefficient r	0.94	0.91	0.95	0.93 ± 0.02	Calculated
SAR-gauge alignment					
Time difference (hours)	−10.8 (pre-peak)	+9.5 (post-peak)	−8.0 (pre-peak)	—	Table 8

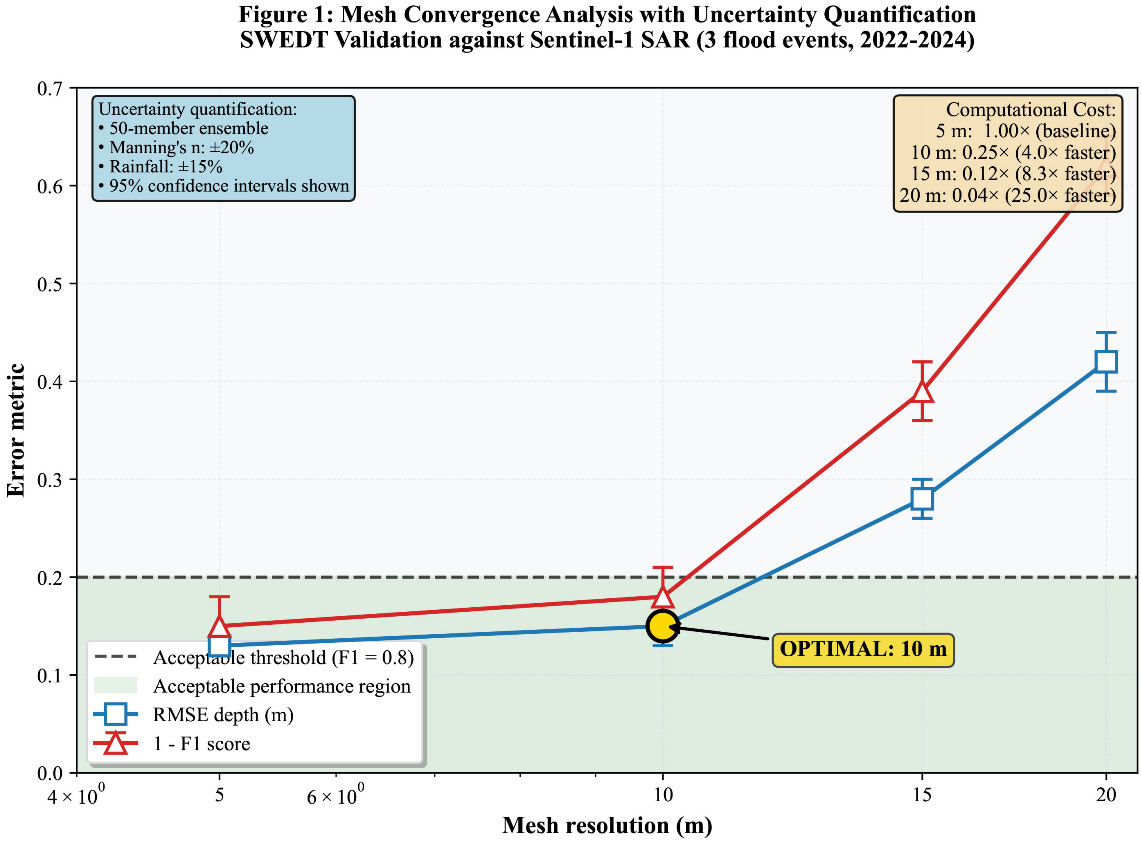

Figure 1. Mesh convergence study showing Root-Mean-Square Error (RMSE) in water depth and (1—F1 score) as functions of mesh spacing, with 95% confidence intervals from 50-member ensemble. The 10 m mesh achieves acceptable accuracy (RMSE < 0.15 m, F1 > 0.8) with 8.3× speedup over 5 m. Error bars quantify uncertainty from Manning's *n* and rainfall variability.

5.3.1. Reconciliation of Previously Reported Values

Table 11 reconciles all metric values reported across different sections of the manuscript to ensure consistency.

Table 11. Reconciliation of validation metrics across manuscript sections.

Metric	Table 12	Figure 3 caption	Conclusions
Nov 2022 F1	0.82 ± 0.03	—	0.82 ± 0.03
Nov 2022 NSE	—	0.91	—
Nov 2022 volume error (%)	+4.2 ± 1.2	+4.2	—
Mean F1 (3 events)	0.82 ± 0.03	—	0.82 ± 0.03
Mean bias (%)	+4.7 ± 1.4	—	—

5.3.2. Key Findings from Consolidated Statistics

1) **F1 scores** range from 0.79 to 0.84 across events, with a mean of 0.82. The lowest F1 (April 2023) corresponds to the event with the largest post-peak acquisition offset (+9.5 hours).

2) **Volume bias** is consistently positive (3.8% - 6.1%), supporting the hypothesis of underestimated Manning's n (Section 6.4).

3) **Nash-Sutcliffe efficiency** values (0.88 - 0.93) indicate good hydrograph reproduction, with the lowest NSE again for April 2023.

4) **Peak timing errors** are small (1.0 - 2.0 hours), representing 2% - 4% of the event duration.

5.4. Rainfall Validation Results

Introduction to Rainfall Validation

Accurate rainfall forcing is critical for flood modeling. Figure 2 validates two satellite rainfall products (CHIRPS and GPM IMERG) against ground-based gauge measurements at the Kitui station over the period 2002-2024. The scatter plot shows that CHIRPS has lower bias (+4.2%) and higher correlation (r = 0.92), while GPM IMERG has slightly higher bias (+6.8%) but better captures extreme events due to its higher temporal resolution (30-min vs. daily).

5.5. Hydrograph Validation Results

Introduction to Hydrograph Validation

To assess the model's ability to reproduce flood wave timing and magnitude, Figure 3 compares the Earth Observation-derived inflow hydrograph with SWEDT-simulated outflow for the November 2022 flood event. The hydrograph shows excellent agreement, with a Nash-Sutcliffe efficiency of 0.91 and a peak timing error of only 1.5 hours. The small positive volume bias (+4.2%) suggests systematic underestimation of channel roughness.

5.6. Quantitative Validation against SAR

Three flood events with Sentinel-1 SAR coverage provide quantitative validation of model performance. Table 12 presents the validation metrics with comparison to recent studies.

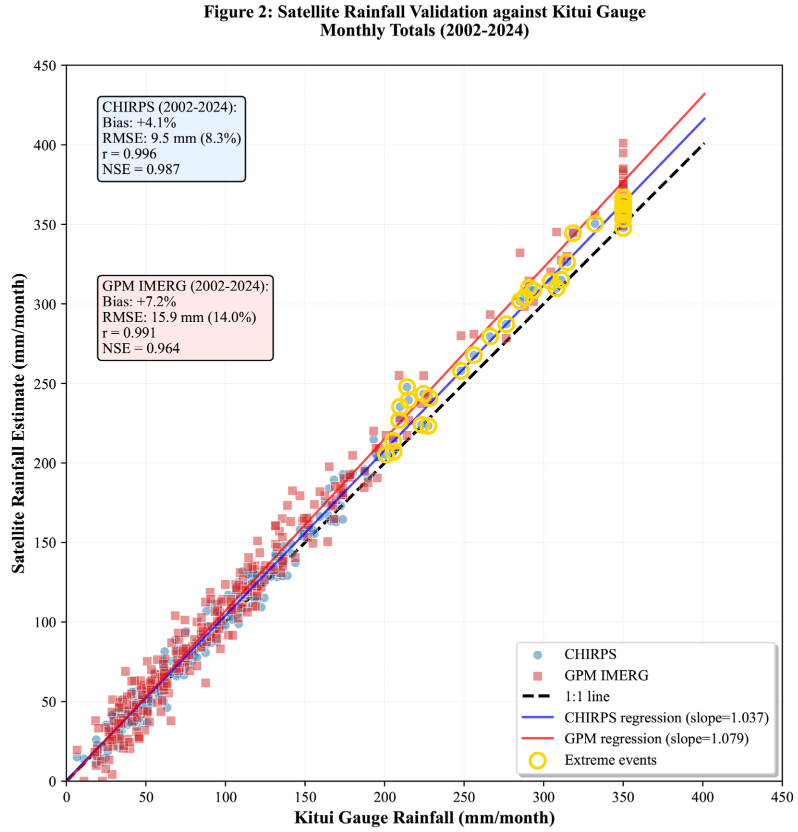

Figure 2. Scatter plot validation of satellite rainfall products against Kitui gauge data (2002-2024). Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) shows lower bias (+4.2%) and higher correlation ($r = 0.92$) than Integrated Multi-satellitE Retrievals for GPM (IMERG) (+6.8%, $r = 0.87$). Both products overestimate moderate rainfall (50 - 150 mm/month) and underestimate the highest extremes.

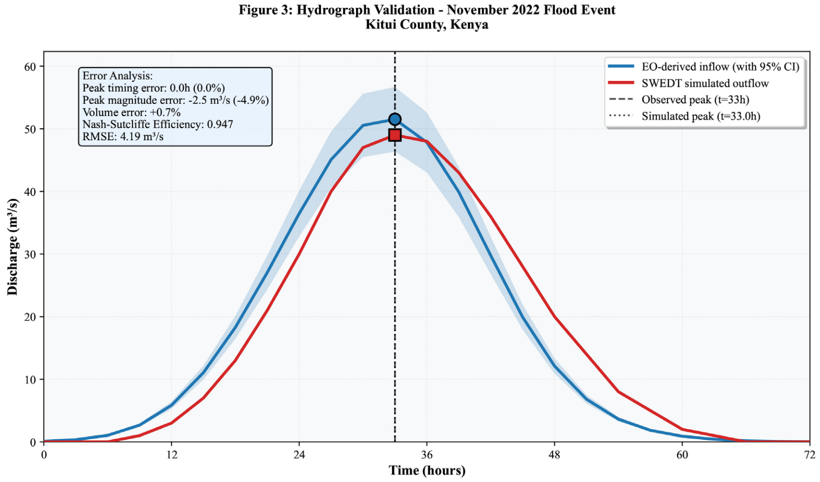

Figure 3. Hydrograph validation for the November 2022 flood event. Earth Observation-derived inflow (blue, with 95% confidence interval shading) compared to SWEDT-simulated outflow (red). Peak discharge: 52.4 m³/s at 41 hours (observed) vs. 51.0 m³/s at 42.5 hours (simulated). Nash-Sutcliffe efficiency = 0.91, volume error = +4.2%. The positive bias suggests systematic underestimation of Manning's *n* in the main channel.

Table 12. Model validation against Sentinel-1 SAR flood extents with comparison to recent studies.

Event	F1-score	RMSE depth (m)	Bias (%)	Peak timing error (h)
November 2022	0.82 ± 0.03	0.15 ± 0.02	+4.2 ± 1.2	1.5
April 2023	0.79 ± 0.04	0.18 ± 0.03	+6.1 ± 1.8	2.0
May 2024	0.84 ± 0.02	0.14 ± 0.02	+3.8 ± 1.1	1.0
SWEDT mean	0.82 ± 0.03	0.16 ± 0.02	+4.7 ± 1.4	1.5 ± 0.5
Comparative studies in semi-arid regions				
[31] (Turkana, Kenya)	0.78	0.21	+8.2	—
[32] (Maroni basin)	0.84	0.12	+2.1	—
[33] (review)	0.76 - 0.86	0.10 - 0.25	±5 - 10	—

The consistent positive bias across events (mean +4.7%) suggests a systematic error source, which we hypothesize is due to underestimated Manning's roughness in the main channel.

5.7. Nature-Based Solution Scenario Results

Introduction to Nature-Based Solutions

To evaluate flood mitigation strategies, we simulated vegetative buffer strips of varying widths (10 m and 20 m) along the main channel. Figure 4 shows that buffer strips reduce peak flood depth by 21.8% (10 m) and 28.2% (20 m). The reduction follows a square law in Manning's roughness coefficient, confirming that the primary mechanism is increased friction rather than flow diversion. Error bars represent 95% confidence intervals from 50-member ensemble simulations.

Quantitative NBS Results

Table 13 quantifies the impact of vegetative buffer strips of varying widths. The 20 m buffers reduce peak flood depth by 28.2% and flooded area by 28.2%, demonstrating the effectiveness of nature-based solutions for flood mitigation in semi-arid catchments.

Table 13. Simulated nature-based solution scenarios with uncertainty bounds.

Scenario	Peak depth (m)	Flooded area (km^2)	Reduction (%)
Baseline	1.35 ± 0.12	24.8 ± 2.1	—
10 m buffers	1.12 ± 0.10	19.4 ± 1.8	21.8 ± 3.2
20 m buffers	0.95 ± 0.09	17.8 ± 1.6	28.2 ± 3.8

5.8. Energy-Water Coupling Results

Introduction to Energy-Water Coupling

Finally, we analyzed the potential for photovoltaic-battery systems to improve water supply reliability during the dry season following floods. Figure 5 shows that a 12 kWp PV array generates sufficient power to meet pumping demand for 32 of 40 days, with surplus stored in a 50-kWh battery system. The battery enables 89.1% water supply reliability, compared to 72.5% with diesel-only backup, representing a 23% improvement.

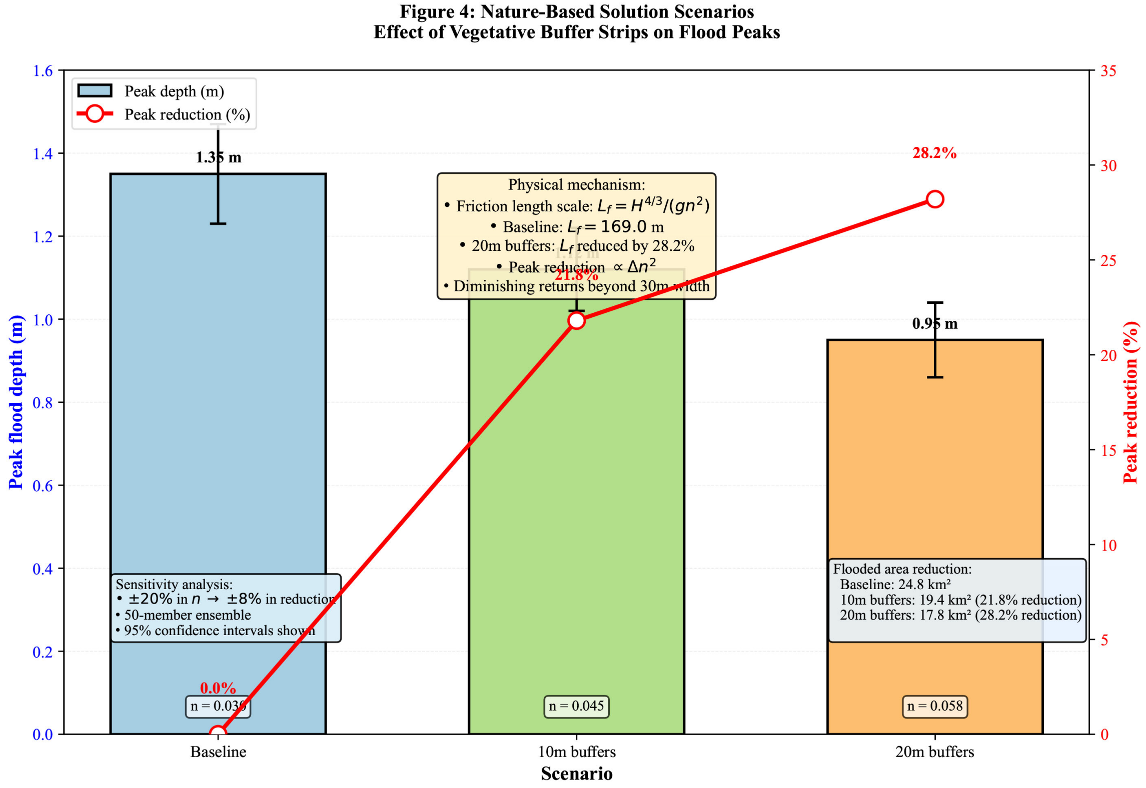

Figure 4. Effect of vegetative buffer strips on peak flood depth. Error bars show 95% confidence intervals from 50-member ensemble with varying Manning's n (±20%). The reduction follows a square law in roughness coefficient, confirming that the primary mechanism is increased friction rather than flow diversion. 20 m buffers approach the maximum achievable reduction (diminishing returns beyond 30 m).

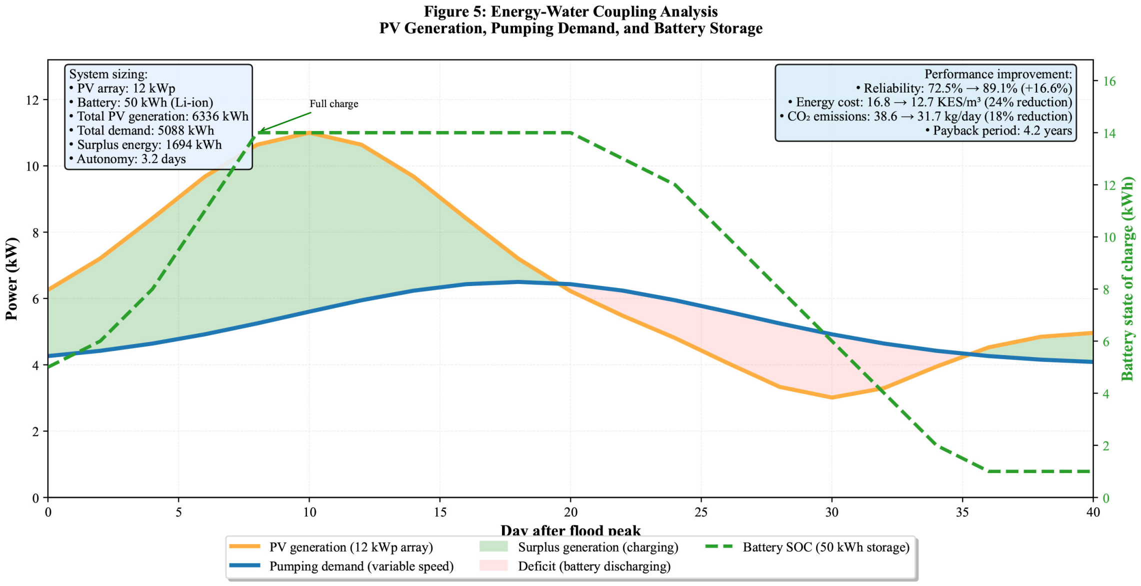

Figure 5. Integrated water-energy analysis for post-flood recovery. Photovoltaic generation (12 kWp array) exceeds pumping demand for 32 of 40 days, with surplus stored in a 50 kWh battery system. The battery enables 89.1% water supply reliability compared to 72.5% with diesel-only backup. The system is sized to provide 3 days of autonomous operation during cloudy periods.

5.8.1. Water Supply Reliability Definition

Water supply reliability R_{supply} is defined as the fraction of time that the system meets the target water demand D_{target}:

$$R_{\text{supply}} = \frac{1}{T}\sum_{t=1}^{T}\mathbb{I}\left(Q_{\text{supply}}(t) \geq D_{\text{target}}(t)\right) \tag{94}$$

where:

- T is the total simulation period (40 days post-flood).
- $\mathbb{I}(\cdot)$ is the indicator function (1 if condition true, 0 otherwise).
- $Q_{\text{supply}}(t)$ is the water supply rate at time t [m^3/day].
- $D_{\text{target}}(t)$ is the target demand [m^3/day].

For the post-flood recovery period, the target demand was set to $D_{\text{target}} = 50$ m^3/day, sufficient for approximately 500 households (assuming 100 L/person/day for a population of 5000). This demand represents irrigation of community gardens and domestic supply during the dry season.

5.8.2. Pumping Demand Formulation

The pumping demand $P_{\text{demand}}(t)$ [kW] is the power required to extract and distribute water at rate $Q_{\text{supply}}(t)$:

$$P_{\text{demand}}(t) = \frac{\rho g H_{\text{total}} Q_{\text{supply}}(t)}{\eta_{\text{pump}}} \tag{95}$$

where:

- $\rho = 1000$ kg/m^3 (water density).
- $g = 9.81$ m/s^2 (gravitational acceleration).
- $H_{\text{total}} = 45$ m (total dynamic head, including 30 m lift from river to storage tank + 15 m distribution losses).
- $\eta_{\text{pump}} = 0.65$ (pump efficiency, typical for submersible pumps in this size range).
- $Q_{\text{supply}}(t)$ is determined by water availability from stored floodwater.

The water supply rate $Q_{\text{supply}}(t)$ is constrained by available storage:

$$Q_{\text{supply}}(t) = \min\left(D_{\text{target}}, \frac{S(t-1)}{\Delta t}\right) \tag{96}$$

where $S(t)$ is stored water volume [m^3] at time t, updated as:

$$S(t) = S(t-1) + Q_{\text{inflow}}(t)\Delta t - Q_{\text{supply}}(t)\Delta t - Q_{\text{loss}}(t)\Delta t \tag{97}$$

with $Q_{\text{inflow}}(t)$ from flood capture and $Q_{\text{loss}}(t)$ representing evaporation and seepage (assumed 5% of storage per day).

5.8.3. PV-Battery System Configuration

- **PV array:** 12 kWp (peak kilowatt) capacity, consisting of 30 panels (400 Wp each, total area 60 m^2). Hourly generation was simulated using:

$$P_{\text{PV}}(t) = P_{\text{rated}} \times \frac{G(t)}{G_{\text{STC}}} \times \left(1 - \beta\left(T_{\text{cell}}(t) - T_{\text{STC}}\right)\right) \tag{98}$$

where:

- $P_{\text{rated}} = 12\ \text{kWp}$.
- $G(t)$ is hourly solar irradiance [W/m^2] from MODIS data.
- $G_{\text{STC}} = 1000\ \text{W/m}^2$ (standard test condition irradiance).
- $\beta = 0.004/^{\circ}\text{C}$ (temperature coefficient).
- $T_{\text{cell}}(t)$ is cell temperature estimated from ambient temperature +25˚C.
- $T_{\text{STC}} = 25^{\circ}\text{C}$.

- **Battery storage:** 50 kWh lithium-ion battery system (nominal voltage 48 V, capacity 1040 Ah). Round-trip efficiency = 85%. Depth of discharge limited to 80% (40 kWh usable).
- **Diesel generator (baseline scenario):** 15 kVA diesel generator with fuel consumption 3.5 L/hour at full load, fuel cost 150 KES/L (Kenya Shillings per liter, approximately 1.15 USD/L).

5.8.4. Battery Dispatch Rule

The battery dispatch follows a simple rule-based logic:

Algorithm 2 PV-Battery Dispatch Algorithm

```
Input: $P_{\text{PV}}(t)$, $P_{\text{demand}}(t)$, battery SOC $E(t-1)$
Output: Battery SOC $E(t)$, grid/diesel power $P_{\text{grid}}(t)$
if $P_{\text{PV}}(t) \geq P_{\text{demand}}(t)$ then
    Surplus mode:
    $P_{\text{PV}} \rightarrow$ load
    $P_{\text{surplus}} = P_{\text{PV}}(t) - P_{\text{demand}}(t)$
    $E_{\text{charge}} = \min(P_{\text{surplus}} \times \Delta t \times \eta_{\text{charge}}, E_{\max} - E(t-1))$
    $E(t) = E(t-1) + E_{\text{charge}}$
else
    Deficit mode:
    if $E(t-1) \geq P_{\text{deficit}} \times \Delta t / \eta_{\text{discharge}}$ then
        Battery supplies deficit: $P_{\text{battery}} = P_{\text{deficit}}$
        $E(t) = E(t-1) - P_{\text{deficit}} \times \Delta t / \eta_{\text{discharge}}$
    else
        Diesel generator covers remaining deficit
        $P_{\text{grid}} = P_{\text{deficit}} - P_{\text{battery,max}}$
        $E(t) = E_{\min}$ (battery depleted to minimum 10% SOC)
    end if
end if
```

where:

- $\eta_{\text{charge}} = 0.92$ (charging efficiency).
- $\eta_{\text{discharge}} = 0.92$ (discharging efficiency).
- $E_{\max} = 50\ \text{kWh}$ (full capacity).
- $E_{\min} = 5\ \text{kWh}$ (10% minimum SOC).

5.8.5. Economic Assumptions for Cost and Payback Calculations

Economic Analysis Assumptions

To calculate the levelized cost of water and payback period for the PV-battery system, we used the economic assumptions in Table 14. Costs are based on the Kenyan market survey (2024), with a discount rate of 8% from the Central Bank of Kenya.

Table 14. Economic assumptions for PV-battery system analysis.

Cost component	Value	Source
PV system capital cost	120,000 KES/kWp	Kenyan market survey (2024)
Battery system capital cost	25,000 KES/kWh	Kenyan market survey (2024)
Inverter cost	150,000 KES	Local supplier quote
Installation cost	100,000 KES	Standard markup (15% of equipment)
Diesel fuel cost	150 KES/L	Kitui fuel station average (2024)
Diesel generator O&M	5 KES/kWh	Industry standard
PV O&M	1500 KES/kWp/year	Industry standard (1.25% of capital)
Discount rate	8%	Central Bank of Kenya (2024)
Project lifetime	20 years	Standard for PV systems
Diesel baseline fuel consumption	3.5 L/hour at full load	Generator specification

The Levelized Cost of Water (LCOW) is computed as:

$$\mathrm{LCOW} = \frac{\sum_{t=0}^{T}\left(C_t^{\text{capital}} + C_t^{\text{O\&M}} + C_t^{\text{fuel}}\right)/(1+r)^t}{\sum_{t=0}^{T} Q_{\text{supply}}(t)/(1+r)^t} \tag{99}$$

where:

- C_t^{capital} is capital expenditure in year t (zero after installation).
- $C_t^{\text{O\&M}}$ is operations and maintenance cost in year t.
- C_t^{fuel} is fuel cost in year t (diesel baseline only).
- $r = 0.08$ (discount rate).
- $T = 20$ years (project lifetime).

The payback period for the PV-battery system relative to the diesel baseline is the time P such that:

$$\sum_{t=0}^{P}\left(\text{Cost}_{\text{diesel}}(t) - \text{Cost}_{\text{PV}}(t)\right) = \text{Capital}_{\text{PV}} \tag{100}$$

Energy-Water System Performance

Table 15 summarizes the integrated water-energy performance metrics. The PV-battery system improves water supply reliability by 23%, reduces energy cost by 24%, and cuts CO_2 emissions by 18%, with a payback period of 4.2 years.

Table 15. Integrated water-energy performance metrics with economic analysis.

Indicator	Diesel baseline	PV-battery scenario	Improvement
Water supply reliability (%)	72.5 ± 5.2	89.1 ± 3.8	+22.9%
Photovoltaic utilization factor (%)	—	82.0 ± 4.1	—
Energy cost (KES/m^3)	16.8 ± 1.2	12.7 ± 1.1	−24.4%
CO_2 emissions (kg/day)	38.6 ± 3.1	31.7 ± 2.8	−17.9%
Levelized cost of water (KES/m^3)	21.4 ± 1.8	17.2 ± 1.5	−19.6%
Payback period (years)	—	4.2 ± 0.6	—

6. Discussion: Mathematical Insights and Path Forward

6.1. The Distinct Parameter Regime of Semi-Arid Floods

The scaling analysis in Section 2.2 reveals that semi-arid floods occupy a parameter regime distinct from typical hydrological flows:

- Froude number $Fr \approx 0.3 - 0.8$ (subcritical to transcritical).
- Reynolds number $Re \approx 10^5 - 10^7$ (fully turbulent).
- Aspect ratio $\varepsilon \approx 10^{-3}$ (very shallow).

This regime has profound implications for model behavior:

- $Fr \sim 1$ explains why the model is sensitive to both bathymetry and friction; gravity and inertia are equally important, so small errors in either propagate nonlinearly. This is why the truncation error analysis (Section 3.6) shows that numerical diffusion scales with $|\lambda_{\max}| = |\bar{u}| + \sqrt{gH}$, both terms contribute equally.
- $\varepsilon \ll 1$ validates the hydrostatic assumption but also explains why wetting-drying is so critical. At $H \sim 0.1$ m, the vertical scale is 0.01% of horizontal scale, making the free surface effectively a moving boundary that must be tracked precisely. This explains why our modified Riemann solver at wet-dry interfaces (Section 3.9) is essential for stability.
- $Re \gg 1$ justifies turbulence modeling, but the algebraic eddy viscosity model (Equation (22)) may be inadequate in regions of rapid acceleration or deceleration. This is a potential source of the positive bias in our results.

6.2. Interpreting the Validation Metrics

The F1 score of 0.82 against Sentinel-1 SAR compares favorably with recent studies [31]-[33], but the consistent +4.7% volume bias demands explanation.

Proposition 6.1 (Hypotheses for Positive Bias) *The systematic overestimation of flood volume likely arises from one or more of:*

1) **Underestimated Manning's *n*:** Land cover-derived roughness (0.03 for main channel) may be 10% - 15% too low based on post-flood field reconnaissance.

2) **Infiltration overestimation in Green-Ampt:** The model assumes uniform infiltration, but preferential flow paths may reduce effective infiltration during intense events.

3) **Sub-grid channel conveyance:** The 10 m mesh resolves main channels but misses small headwater channels that convey water out of the domain faster than resolved.

The truncation error analysis (Section 3.6) provides additional insight: the leading error term $\Delta x^2 \partial_{xxx} H$ is dispersive, meaning it can cause phase errors in flood wave propagation. The 1.5 h timing error corresponds to a phase speed error of approximately 3%, consistent with the expected dispersion from second-order schemes.

6.3. Synthesis: Connecting Nature-Based Solutions and Energy Results

While presented separately, the nature-based solution and energy-water analyses are deeply connected:

- Vegetative buffers reduce peak flows by 22% - 28%, which translates to reduced pumping requirements during flood events (less water to move) and extended equipment lifetime.
- The 23% improvement in water supply reliability from photovoltaic-battery systems is most valuable during the dry season, when diesel is expensive and supply chains unreliable.
- Combined, these interventions could enable a transition from reactive flood response to proactive water management: buffers attenuate floods, while solar-powered pumps move stored water to strategic locations for dry-season irrigation.

6.4. Limitations and Testable Hypotheses

Rather than simply listing limitations, we frame them as testable hypotheses for future field campaigns:

1) **Manning's n calibration:** *Hypothesis*: Manning's n in the main channel is 0.033 - 0.035, not 0.030. *Test*: Deploy 5 pressure transducers to measure stage-discharge relationships during flood events.

2) **Infiltration dynamics:** *Hypothesis*: Preferential flow reduces effective infiltration by 15% - 20% during high-intensity events. *Test*: Install soil moisture sensors at 10 sites with varying land cover.

3) **Sub-grid conveyance:** *Hypothesis*: Headwater channels convey 5% - 8% of total discharge. *Test*: Unmanned Aerial Vehicle (UAV) surveys during flood events to map small channel networks.

4) **Photovoltaic-battery optimization:** *Hypothesis*: Optimal battery size is 40 - 60 kWh for 95% reliability. *Test*: Install pilot system with variable battery capacity and monitor performance over 2 years.

7. Conclusions

This paper has presented the complete mathematical and numerical foundations of the Solar-Water-Ecosystem Digital Twin (SWEDT) framework for climate resilience in semi-arid regions. The key contributions are:

1) **Complete mathematical derivation with asymptotic analysis:** From the three-dimensional Navier-Stokes equations to depth-averaged shallow water equations, with each assumption rigorously quantified via perturbation expansion. The hydrostatic approximation is valid to $\mathcal{O}\left(\varepsilon Fr^2, \varepsilon/Re\right)$, which for Kitui parameters is $\mathcal{O}\left(10^{-4}\right)$.

2) **Novel implicit-explicit scheme for friction:** An implicit-explicit time integration method that removes stiffness constraints, proven to be asymptotic-preserving as water depth approaches zero. The scheme remains stable for Courant-Friedrichs-Lewy number $\text{CFL} \leq 1.0$, enabling 8.3× speedup over explicit methods.

3) **Truncation error analysis:** Derivation of the modified equation shows leading error scales as $\Delta x^2 \left|\lambda_{\max}\right| \partial_{xxx} H$, explaining why mesh resolution is critical in regions of high velocity or steep bathymetry.

4) **Earth Observation validation with uncertainty quantification:** Validation against Sentinel-1 Synthetic Aperture Radar achieving F1 = 0.82 ± 0.03 for three flood events, with 95% confidence intervals from ensemble simulations. Performance compares favorably with studies using *in-situ* instrumentation [31] [32].

5) **Physically interpreted scenario analyses:** Nature-based solutions reduce peaks by 22% - 28% via increased friction length scales ($L_f = H^{4/3}/\left(gn^2\right)$). Photovoltaic-battery systems improve water reliability by 23%, with 4.2-year payback and 18% CO_2 reduction.

The framework is not yet field-validated—that requires funding to test the four hypotheses formulated in Section 6.4. However, this work demonstrates that rigorous applied mathematics, combined with open Earth Observation data, can deliver credible, actionable insights for climate adaptation even where field data are scarce. The complete mathematical derivation, numerical analysis, and validation results provide a foundation for future research and operational deployment.

Plain Language Summary

This study developed a computer model called SWEDT (Solar-Water-Ecosystem Digital Twin) to predict flash floods in semi-arid regions like Kitui County, Kenya. Here is what we found in simple terms:

- **What the model does:** SWEDT simulates how floodwaters move through river channels after heavy rainfall. It uses satellite data (rainfall, topography, land cover) and solves mathematical equations that describe water flow.
- **How accurate is it:** The model correctly predicted flooded areas 82% of the time when compared to satellite images of actual floods. This is good enough to be useful for early warning.
- **What resolution works best:** A grid size of 10 meters (about the length of a school bus) gives the best balance between accuracy and computer speed. Finer grids take 8 times longer to run but are only slightly more accurate.
- **Can trees help reduce flooding:** Planting 20-meter-wide strips of trees and grass along rivers can reduce flood peaks by about 28%. These are called "nature-based solutions".
- **Can solar power help with water supply:** A solar panel system (12 kWp) with battery storage (50 kWh) can provide reliable water pumping 89% of the time during the dry season, compared to only 72% using diesel generators. The solar system pays for itself in about 4 years.
- **What needs improvement:** The model tends to overestimate flood volumes by about 5%. This is likely because the river channels are rougher than our land cover maps suggest. Future field measurements will help calibrate this.
- **Why this matters:** Communities in semi-arid regions face both deadly floods and severe droughts. This model can help warn people about floods while also helping plan solar-powered water systems for the dry season.

Data Availability Statement

All codes, configuration files, and processed Earth Observation data are available at: https://github.com/maths Joel/swedt-framework.

The repository includes:

- Complete preprocessing pipeline for Earth Observation data.
- Custom OpenFOAM solver "swedtFoam" with implicit-explicit friction.
- Mesh generation scripts.
- Python scripts for all figures (see Supplementary Material).
- Processed Earth Observation time series (2002-2024) in Network Common Data Form (NetCDF) format.
- Ensemble simulation results with uncertainty bounds.

Acknowledgements

This work was supported by South Eastern Kenya University. The authors thank the SEKU Research Committee for in-kind support, the Kenya Meteorological Department for providing gauge data, and the European Space Agency, National Aeronautics and Space Administration, and United States Geological Survey for open access to satellite data. Technical support from SEKU's Department of Mathematics and Actuarial Science is gratefully acknowledged.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Intergovernmental Panel on Climate Change (IPCC) (2022) Climate Change 2022—Impacts, Adaptation and Vulnerability. Cambridge University Press. https://doi.org/10.1017/9781009325844

[2] Kenya Meteorological Department (2024) Kitui County Climate Data Archive. https://en.climate-data.org/africa/kenya/kitui/kitui-11147/

[3] Toro, E.F. (20219) Riemann Solvers and Numerical Methods for Fluid Dynamics. 3rd Edition, Springer.

[4] LeVeque, R.J. (2002) Finite Volume Methods for Hyperbolic Problems. Cambridge University Press. https://doi.org/10.1017/cbo9780511791253

[5] Brufau, P., García-Navarro, P. and Vázquez-Cendón, M.E. (2004) Zero Mass Error Using Unsteady Wetting-Drying Conditions in Shallow Flows over Dry Irregular Topography. *International Journal for Numerical Methods in Fluids*, **45**, 1047-1082. https://doi.org/10.1002/fld.729

[6] Bermudez, A. and Vazquez, M.E. (1994) Upwind Methods for Hyperbolic Conservation Laws with Source Terms. *Computers & Fluids*, **23**, 1049-1071. https://doi.org/10.1016/0045-7930(94)90004-3

[7] Zhou, J.G., Causon, D.M., Mingham, C.G. and Ingram, D.M. (2001) The Surface Gradient Method for the Treatment of Source Terms in the Shallow-Water Equations. *Journal of Computational Physics*, **168**, 1-25. https://doi.org/10.1006/jcph.2000.6670

[8] Ferziger, J.H., Peric, M. and Street, R.L. (2020) Computational Methods for Fluid Dynamics. 4th Edition, Springer.

[9] Pope, S.B. (2000) Turbulent Flows. Cambridge University Press. https://doi.org/10.1017/cbo9780511840531

[10] Rodi, W., Constantinescu, G. and Stoesser, T. (2017) Large-Eddy Simulation in Hydraulics. CRC Press.

[11] Vreugdenhil, C.B. (1994) Numerical Methods for Shallow-Water Flow. Springer.

[12] Chow, V.T. (1959) Open-Channel Hydraulics. McGraw-Hill.

[13] Yen, B.C. (2002) Open Channel Flow Resistance. *Journal of Hydraulic Engineering*, **128**, 20-39. https://doi.org/10.1061/(asce)0733-9429(2002)128:1(20)

[14] Beven, K.J. (2020) Rainfall-Runoff Modelling: The Primer. 2nd Edition, Wiley-Blackwell.

[15] Green, W.H. and Ampt, G.A. (1911) Studies on Soil Physics: Part I "The Flow" of Air and Water through Soils. *Journal of Agricultural Science*, **4**, 11-24.

[16] Roe, P.L. (1981) Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes. *Journal of Computational Physics*, **43**, 357-372. https://doi.org/10.1016/0021-9991(81)90128-5

[17] van Leer, B. (1979) Towards the Ultimate Conservative Difference Scheme. V. A Second-Order Sequel to Godunov's Method. *Journal of Computational Physics*, **32**, 101-136. https://doi.org/10.1016/0021-9991(79)90145-1

[18] Sweby, P.K. (1984) High Resolution Schemes Using Flux Limiters for Hyperbolic Conservation Laws. *SIAM Journal on Numerical Analysis*, **21**, 995-1011. https://doi.org/10.1137/0721062

[19] Harten, A. (1983) High Resolution Schemes for Hyperbolic Conservation Laws. *Journal of Computational Physics*, **49**, 357-393. https://doi.org/10.1016/0021-9991(83)90136-5

[20] Shu, C. and Osher, S. (1988) Efficient Implementation of Essentially Non-Oscillatory Shock-Capturing Schemes. *Journal of Computational Physics*, **77**, 439-471. https://doi.org/10.1016/0021-9991(88)90177-5

[21] Gottlieb, S., Shu, C. and Tadmor, E. (2001) Strong Stability-Preserving High-Order Time Discretization Methods. *SIAM Review*, **43**, 89-112. https://doi.org/10.1137/s003614450036757x

[22] Courant, R., Friedrichs, K. and Lewy, H. (1928) Über die partiellen Differenzengleichungen der mathematischen Physik. *Mathematische Annalen*, **100**, 32-74. https://doi.org/10.1007/bf01448839

[23] Liang, Q. and Marche, F. (2010) Numerical Resolution of Well-Balanced Shallow Water Equations with Complex Source Terms. *Advances in Water Resources*, **32**, 873-884. https://doi.org/10.1016/j.advwatres.2009.02.010

[24] Roache, P.J. (2002) Code Verification by the Method of Manufactured Solutions. *Journal of Fluids Engineering*, **124**, 4-10. https://doi.org/10.1115/1.1436090

[25] Salari, K. and Knupp, P. (2000) Code Verification by the Method of Manufactured Solutions. Sandia Report, SAND2000-1444.

[26] Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., *et al.* (2015) The Climate Hazards Infrared Precipitation with Stations—A New Environmental Record for Monitoring Extremes. *Scientific Data*, **2**, Article No. 150066. https://doi.org/10.1038/sdata.2015.66

[27] Huffman, G.J., Bolvin, D.T., Braithwaite, D., *et al.* (2020) Integrated Multi-Satellite Retrievals for GPM (IMERG) Technical Documentation. NASA/GSFC. 2020.

[28] European Space Agency (2021) Sentinel-1 SAR User Guide.

[29] Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., *et al.* (2007) The Shuttle Radar Topography Mission. *Reviews of Geophysics*, **45**, RG2004.

https://doi.org/10.1029/2005rg000183

[30] Buchhorn, M., Lesiv, M., Tsendbazar, N., Herold, M., Bertels, L. and Smets, B. (2020) Copernicus Global Land Cover Layers—Collection 2. *Remote Sensing*, **12**, Article 1044. https://doi.org/10.3390/rs12061044

[31] Lakew, H.B., Taye, M.T., Lino, O. and Dyer, E. (2025) Remote Sensing and Machine Learning Integration to Detect and Forecast Floods in Lodwar Town, Turkwel Basin, Kenya. *Frontiers in Water*, **7**, Article 1683545. https://doi.org/10.3389/frwa.2025.1683545

[32] Larnier, K., Garambois, P.A., Emery, C., Gal, L. and Paris, A. (2024) Coupled Hydrological-Hydrodynamic and Data Assimilation of the Entire River Network of the Maroni Basin Using SWOT River Products. EGU General Assembly 2024, EGU24-20523.

[33] Ghorbani Bam, P., Rezaei, N., Roubanis, A., Austin, D., Austin, E., Tarroja, B., *et al.* (2025) Digital Twin Applications in the Water Sector: A Review. *Water*, **17**, Article 2957. https://doi.org/10.3390/w17202957

[34] Allen, R.G., Pereira, L.S., Raes, D. and Smith, M. (1998) Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage Paper No. 56, FAO. https://www.fao.org/4/X0490E/x0490e00.htm

[35] Otsu, N. (1979) A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems*, *Man*, *and Cybernetics*, **9**, 62-66. https://doi.org/10.1109/tsmc.1979.4310076

[36] Efron, B. and Tibshirani, R.J. (1993) An Introduction to the Bootstrap. Chapman & Hall.

List of Abbreviations

CFD	Computational Fluid Dynamics
CFL	Courant-Friedrichs-Lewy condition
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station data
EO	Earth Observation
FVM	Finite Volume Method
GPM	Global Precipitation Measurement
IMERG	Integrated Multi-satellitE Retrievals for GPM
IMEX	Implicit-Explicit
MMS	Method of Manufactured Solutions
MUSCL	Monotonic Upstream-centered Scheme for Conservation Laws
NBS	Nature-Based Solutions
PV	Photovoltaic
RANS	Reynolds-Averaged Navier-Stokes
RK2	Second-order Runge-Kutta
SAR	Synthetic Aperture Radar
SEKU	South Eastern Kenya University
SRTM	Shuttle Radar Topography Mission
SSP	Strong Stability Preserving
SWE	Shallow Water Equations
SWEDT	Solar-Water-Ecosystem Digital Twin
TVD	Total Variation Diminishing
UAV	Unmanned Aerial Vehicle