HydroPol2D Phase 1 Formula and Dynamics Validation Report

HydroPol2D Development Team

July 16, 2026

Abstract

This report documents Phase 1 of HydroPol2D validation. Phase 1 is limited to formula and dynamics validation using analytical solutions, independent reference calculations, exact bookkeeping tests, and simple synthetic parameter-recovery cases. The shared spatial case-study domain is a v-tilted synthetic catchment, used to embed module-scale analytical tests in a consistent HydroPol2D grid. This version reports completed tests for canopy interception, snow, infiltration, evapotranspiration, groundwater and recharge, hydrodynamics, routing alternatives, Neal (2012) channel-subgrid hydraulics, reservoirs, boundary conditions, spatial rainfall, water-quality buildup and washoff, and human risk. The MATLAB routines are compared against analytical solutions, independent calculations, exact storage ledgers, or documented fine-grid numerical references.

1 Report Status

This document is now scoped to Phase 1 only. Phase 1 tests do not claim field validation. They test whether HydroPol2D equations, storage updates, boundary bookkeeping, and simplified dynamics reproduce known or independently computed reference behavior. Numerical results are included only when current-code outputs have been generated and compared against a documented reference.

2 Phase 1 Scope

Phase 1 validates formulas and dynamics using analytical solutions, independent reference calculations, manufactured or simplified benchmark solutions, exact bookkeeping, and simple synthetic inverse-modeling cases. These tests are best described as computational verification and benchmark validation of individual modules. Synthetic inverse modeling is used only as parameter recovery: known synthetic observations are generated from known parameters, and HydroPol2D or its diagnostic routines must recover those parameters within documented tolerances.

The Phase 1 execution registry is maintained in `HydroPol2D_Model/Validation/Phase1_Cases.csv`. Analytical and bookkeeping reference outputs are stored under `HydroPol2D_Model/Validation/Reference_Output`

The shared Phase 1 spatial case-study domain is the v-tilted synthetic catchment under `HydroPol2D_Model/Validation/Phase1_VTilted_Catchment`. The analytical or independent reference solution remains the truth source; the v-tilted catchment provides a consistent DEM, LAI, soil, land-cover, groundwater, channel, and boundary-condition context for HydroPol2D runs. Ritter dam-break and non-breaking-wave benchmarks retain their own geometries because their analytical comparisons require specialized domains.

3 Case-study Overview

The Phase 1 case-study domain is a synthetic v-tilted catchment. It has a central channel strip and two lateral hillslopes that drain toward the channel, with a longitudinal slope toward the outlet. The domain provides spatial heterogeneity without introducing field-data uncertainty.

The generated raster grid has 50 rows by 81 columns at 20 m resolution. Three zones are used in the first canopy test: a left hillslope with sparse vegetation, a channel strip with no canopy, and a right hillslope with denser vegetation. Table 1 summarizes the domain zones used in the completed canopy case.

Table 1: V-tilted Phase 1 catchment zones used in the canopy interception test.

Zone	Cells	LAI	Initial soil moisture (mm)
Left hillslope	2000	0.5	50
Channel strip	100	0.0	30
Right hillslope	1950	2.0	70

4 Validation Methods

Each validation case must define: model flags, input data, expected behavior, diagnostic outputs, metrics, acceptance thresholds, and report figure/table identifiers. The standard diagnostic bundle is:

- mass-balance table;
- time-series and profile comparisons where applicable;
- raster maps or one-dimensional profiles for spatial behavior;
- metric summary table;
- pass/fail table with threshold references.

Storage and algebraic unit tests use exact mass closure as the primary criterion. The default tolerance is $1 \times 10^{-6}\,\mathrm{m}^3$ for unit tests or 0.01 % of input volume when the volume scale makes an absolute tolerance inappropriate. Full-domain cases target final water-balance error below 0.1 %. Hydrograph and profile benchmarks report root-mean-square error, mean absolute error, maximum error, relative L_2 error, peak timing error, peak magnitude error, and Nash-Sutcliffe efficiency where applicable.

No case is report-ready when the diagnostic summary contains unexplained `NaN` values, undocumented residuals, or missing pass/fail criteria.

4.1 Phase 1 Evidence Hierarchy

For Phase 1, the preferred evidence hierarchy is:

1. analytical or semi-analytical solution;
2. independently computed reference formula;

3. manufactured or simplified benchmark solution;
4. synthetic inverse-modeling parameter recovery;
5. cross-model comparison against a simple trusted solver.

The first three are the main evidence sources. Synthetic inverse modeling is useful for checking identifiability and parameter sensitivity, but it must not be presented as field validation.

4.2 Phase 1 Reference Cases

Phase 1 now includes explicit analytical/reference cases for the main computational modules. Priority cases include Green-Ampt and Philip infiltration, FAO-style ET with availability limits, degree-day snow storage, canopy bucket storage, linear-reservoir groundwater recession, Ritter and non-breaking-wave hydrodynamics, corrected kinematic/diffusive routing, Manning/channel conveyance, reservoir rating-curve depletion, exact inflow/stage volume bookkeeping, tiny raster rainfall totals, analytical wash-off, and deterministic human-risk thresholds.

For HydroPol2D runs, these cases should be embedded in the v-tilted synthetic catchment wherever possible. The report must still compare model outputs against the analytical/reference solution, not against the v-tilted domain itself.

5 Phase 1 Validation Matrix

The Phase 1 registry is maintained in `HydroPol2D_Model/Validation/Phase1_Cases.csv`. Table 2 summarizes the current Phase 1 evidence chain.

Table 2: HydroPol2D Phase 1 validation matrix.

Component	Reference evidence	Case-study domain	Status
Canopy interception	Independent canopy-bucket reference	V-tilted zones	Report-ready
Snow	Independent snow formula/storage reference	V-tilted zones	Report-ready
Infiltration	Independent layered storage and Darcy–van Genuchten regime references	V-tilted hillslope cells	Report-ready

Component	Reference evidence	Case-study domain	Status
Evapotranspiration	Independent reference ET and extraction limits	V-tilted zones	Report-ready
Groundwater/recharge	Analytical recharge/capillary references, Dupuit hillslope, linearized Boussinesq hillslope, and fine-timestep Boussinesq reference	V-tilted and idealized hillslope profiles	Report-ready
Hydrodynamics	Full-momentum analytical benchmarks; local-inertial, CA, kinematic, and diffusive diagnostics	V-tilted and specialized benchmark domains	Report-ready
Routing alternatives	Corrected kinematic/diffusive solvers pass plane, V-tilted, and stage checks; Ritter remains diagnostic	V-tilted slopes	Report-ready

Component	Reference evidence	Case-study domain	Status
Neal (2012) channel-subgrid hydraulics	Fine-grid local-inertial references for in-bank, composite-channel, and staged-overbank dynamics	Synthetic straight and compound-channel reaches	Report-ready
Reservoir/control structures	Rating-curve storage ODE	Idealized reservoir/control cell	Report-ready
Inflow hydrograph BC	Exact imposed-volume bookkeeping	Rectangular full-momentum domain	Report-ready
Stage hydrograph BC	Prescribed-stage enforcement and volume bookkeeping	Rectangular full-momentum domain	Report-ready
Spatial rainfall	Exact raster rainfall totals	Tiny GeoTIFF raster stack	Report-ready
Water quality	Analytical SWMM-style build-up/wash-off ODE	Single- and two-cell module drivers	Report-ready
Human risk	Independent drag/friction and force-balance references	Synthetic depth/velocity grids	Report-ready

6 Canopy Interception

The canopy interception case isolates the HydroPol2D canopy storage routine before coupled infiltration, runoff, or groundwater processes are considered. The test uses the v-tilted catchment zones described in Table 1. The left hillslope has LAI = 0.5, the channel strip has LAI = 0, and the right hillslope has LAI = 2.0. A representative 20 m by 20 m cell is used for volume conversion.

6.1 Canopy Method

The MATLAB routine `interceptionModel.m` was run directly through the case driver `run_canopy_interception_model.m`. The independent reference follows the same intended canopy-bucket bookkeeping. For each zone and time step, maximum canopy storage is

$$S_{\max} = C\,LAI, \tag{1}$$

where $C = 0.2$ mm per LAI unit. Potential canopy evaporation is reduced by the previous fractional canopy storage:

$$E_t = \min\left(\frac{S_{t-1}}{S_{\max}} E_{p,t},\ P_t + S_{t-1}\right), \tag{2}$$

with $E_t = 0$ when $S_{\max} = 0$. Temporary storage before throughfall is

$$S_t^* = S_{t-1} + P_t - E_t. \tag{3}$$

Throughfall and final canopy storage are then

$$T_t = \max(S_t^* - S_{\max}, 0), \qquad S_t = \min(S_t^*, S_{\max}). \tag{4}$$

Stemflow is not active in the current HydroPol2D routine for this test. The diagnostic residual is

$$r_t = (S_t - S_{t-1}) - (P_t - E_t - T_t). \tag{5}$$

The forcing sequence includes canopy filling, throughfall after capacity is exceeded, evaporation drawdown, and an LAI=0 bypass condition in the channel strip. The reference output is stored in `Reference_Outputs/Phase1/canopy_bucket/P1-CANOPY-001_reference.csv`; the MATLAB HydroPol2D output is stored in `Canopy_Interception/SingleCell_Storage_Balance/Outputs/Validation/HydroP`

6.2 Canopy Results

The HydroPol2D MATLAB canopy output matched the independent reference to round-off precision. Table 3 summarizes the diagnostic metrics. The case is marked report-ready because the benchmark comparison, mass-balance check, and NaN check all passed.

Table 3: Canopy interception validation metrics for case P1-CANOPY-001.

Metric	Value	Threshold	Pass
Canopy storage RMSE (mm)	1.50×10^{-17}	1.0×10^{-9}	yes
Throughfall RMSE (mm)	1.31×10^{-17}	1.0×10^{-9}	yes
Evaporation RMSE (mm)	7.31×10^{-18}	1.0×10^{-9}	yes
Maximum storage error (mm)	5.55×10^{-17}	1.0×10^{-9}	yes
Maximum throughfall error (mm)	5.55×10^{-17}	1.0×10^{-9}	yes
Maximum evaporation error (mm)	2.78×10^{-17}	1.0×10^{-9}	yes
Maximum mass residual (m^3)	2.22×10^{-17}	1.0×10^{-6}	yes

The pass/fail summary reports `status = pass`, `mass_balance_pass = true`, `benchmark_pass = true`, `nan_check_pass = true`, and `report_ready = true`. This validates the canopy interception formula and storage bookkeeping for the tested v-tilted-zone conditions. It does not validate field canopy interception behavior, because no field throughfall or stemflow observations are used in Phase 1.

7 Snow Module

The snow validation applies prescribed cold, transitional, and warm forcings to representative v-tilted zones. The case calls `Snow_Model_Function.m` directly and compares the MATLAB output against an independent Python reference that mirrors the active HydroPol2D formulas. The test checks rain/snow partitioning, snow water equivalent (SWE), snow depth, snow density, melt, sublimation, and snowpack mass closure.

7.1 Snow Method

The active precipitation partition in `Snow_Model_Function.m` uses a linear snow fraction between 4 °C and 7 °C:

$$f_s = \min\left(\max\left(\frac{7 - T_a}{7 - 4}, 0\right), 1\right), \qquad P_s = f_s P, \qquad P_r = P - P_s. \tag{6}$$

This is an implementation detail: the routine accepts a threshold argument, but the active partition currently uses the hard-coded 4–7 °C transition. Phase 1 therefore validates the implemented formula, not a configurable threshold formulation.

After snowfall is added to SWE, the routine computes net radiation, updates density and diagnostic snow depth from previous SWE, and computes melt as the sum of degree-day and radiation terms bounded by available SWE:

$$M_t = \min\left(SWE_{t-1} + P_s,\ \max\left(DDF\,T_a + \frac{(1-\alpha)Q_{net}}{334}, 0\right)\right). \tag{7}$$

Sublimation is computed from wind speed and a vapor-pressure-gradient term and is also bounded by available SWE after melt. The storage update is

$$SWE_t = SWE_{t-1} + P_s - M_t - E_s, \tag{8}$$

with diagnostic residual

$$r_t = SWE_{t-1} + P_s - M_t - E_s - SWE_t. \tag{9}$$

The reference forcing uses three v-tilted scenarios: left hillslope, channel strip, and right hillslope. The temperatures span fully snow-dominated, transitional mixed rain/snow, and rain-only conditions. The reference output is stored in `Reference_Outputs/Phase1/snow_degree_day/P1-SNOW-001_refere`
the MATLAB output is stored in `Snow/ColdWarmPartition_Melt/Outputs/Validation/HydroPol2D_Snow_Model`

7.2 Snow Results

The HydroPol2D MATLAB snow output matched the independent reference to round-off precision. Table 4 summarizes the diagnostic metrics. The case is marked report-ready because the benchmark comparison, mass-balance check, and NaN check all passed.

The pass/fail summary reports `status = pass`, `mass_balance_pass = true`, `benchmark_pass = true`, `nan_check_pass = true`, and `report_ready = true`. This validates the implemented snow partition, melt, sublimation, density, depth, and SWE bookkeeping formulas for the tested v-tilted-zone conditions. It does not validate field snowpack behavior because no snow-pillow, snow-course, or remotely sensed SWE observations are used in Phase 1.

Table 4: Snow-module validation metrics for case P1-SNOW-001.

Metric	Value	Threshold	Pass
SWE RMSE (mm)	1.59×10^{-14}	1.0×10^{-8}	yes
Melt RMSE (mm)	1.27×10^{-14}	1.0×10^{-8}	yes
Snowfall RMSE (mm)	7.54×10^{-15}	1.0×10^{-8}	yes
Rainfall RMSE (mm)	4.19×10^{-16}	1.0×10^{-8}	yes
Sublimation RMSE (mm)	1.86×10^{-20}	1.0×10^{-8}	yes
Snow density RMSE ($\mathrm{kg\,m^{-3}}$)	1.78×10^{-13}	1.0×10^{-8}	yes
Maximum mass residual ($\mathrm{m^3}$)	0.0	1.0×10^{-6}	yes

8 Infiltration

The infiltration validation uses the v-tilted catchment as the shared Phase 1 spatial context, but the truth source is an independent MATLAB reference driver that evaluates the intended infiltration and layered-storage bookkeeping outside the full hydrodynamic solver. This is deliberate: Phase 1 asks whether the equations and storage dynamics behave correctly before interpreting routed hydrographs or field-scale calibration.

8.1 Infiltration Method

Seven regimes were tested with the current code using `run_vtilted_infiltration_validation_suite.m`. The test grid and land-cover/soil zones come from `Phase1_VTilted_Catchment`; the infiltration cases are listed in `Infiltration_Case_Registry.csv`. The suite evaluates: no-infiltration bypass, supply-limited rainfall, capacity-limited intake, finite-storage saturation excess, delayed layered percolation and recharge, shallow-groundwater saturation excess, and spatial soil contrast. The hydraulic parameters were reset to physically interpretable USDA-style sandy-loam, loam, and clay-loam values using the Mualem–van Genuchten framework [3, 10, 2]. Layer multipliers are used only to represent plausible structured topsoil, compacted soil, or moderately slower subsoil conductivity.

For each time step, precipitation is partitioned into infiltration, surface excess, layer storage, and recharge. The independent reference enforces the same intended top-down layer order used by the implemented model: near-surface layer, root-zone layer, transmission layer, and groundwater recharge when the vadose storage is full and recharge is enabled. The diagnostic closure residual is

$$r_t = P_t - F_t - R_t - G_t - \Delta S_t, \tag{10}$$

where P_t is precipitation, F_t is infiltrated water retained in the vadose layers, R_t is surface excess/runoff supply, G_t is recharge, and ΔS_t is the change in vadose storage. The primary acceptance checks are physically distinct rather than one aggregate score: supply-limited cases must infiltrate all rainfall, capacity-limited cases must generate excess when rainfall exceeds intake, storage-limited cases must saturate the vadose store, recharge cases must show delayed transmission and recharge, and spatial-contrast cases must generate more excess from the low-conductivity zone.

A coupled HydroPol2D full-domain runoff-control case was also executed with adaptive hydrodynamic stepping. It completed stably after allowing the inertial solver to reduce the internal time step from the registry forcing interval. That coupled run is retained as diagnostic only because the full routed water-balance benchmark has not yet been defined for this infiltration section.

8.2 Infiltration Results

All seven independent infiltration-regime cases passed after the soil properties were replaced with physically consistent texture-based parameters. Table 5 summarizes the main diagnostics. Closure errors were zero or round-off scale, with the largest residual equal to 1.23×10^{-11} mm in the layered percolation case. The pass/fail file reports `status = pass` and `report_ready = true` for every independent-regime case.

Table 5: Infiltration validation diagnostics for the v-tilted Phase 1 cases.

Case	Regime	Rain (mm)	Infil. (mm)	Runoff (mm)	Recharge (mm)	Closure (mm)
P1-INFIL-VT-000	No infiltration	60.0	0.0	60.0	0.0	0.0
P1-INFIL-VT-001	Supply limited	10.0	10.0	0.0	0.0	0.0
P1-INFIL-VT-002	Capacity limited	120.0	34.157	85.843	0.0	1.71×10^{-13}
P1-INFIL-VT-003	Storage limited	180.0	72.5	107.5	0.0	0.0
P1-INFIL-VT-004	Layered percolation	384.0	384.0	0.0	197.614	1.23×10^{-11}
P1-INFIL-VT-005	Shallow groundwater	150.0	75.0	75.0	22.683	-2.84×10^{-14}
P1-INFIL-VT-006	Spatial soil contrast	120.0	77.880	42.120	0.0	7.11×10^{-15}

The results show the expected behavior under recognizable soil textures. In the supply-limited sandy-loam case, the full 10 mm rainfall input infiltrated. In the capacity-limited clay-loam case, only 34.157 mm infiltrated and 85.843 mm became excess, consistent with rainfall exceeding hydraulic intake capacity. In the shallow sandy-loam storage case, the vadose store filled to 86.25 mm and subsequent rainfall became saturation-excess runoff. In the moist sandy-loam percolation case, all 384 mm of rainfall infiltrated; 197.614 mm drained to recharge and 301.386 mm remained in vadose storage, producing a gradual recharge rise rather than an immediate pulse. In the shallow-groundwater loam case, limited vadose capacity produced 75 mm of excess and 22.683 mm of recharge. The spatial-contrast case produced more excess from the compacted low-conductivity sandy-loam realization than from the high-conductivity realization.

Figures 1 and 2 show the rate diagnostics used to visually inspect the same cases. Individual per-case plots with hydrograph, infiltration-rate, capacity, and layer-storage panels are stored in `VTilted_Infiltration/Outputs/Validation/Figures`. This validates the implemented infiltration bookkeeping and layered-soil dynamics for the tested Phase 1 regimes. It does not validate field infiltration parameters, macropore flow, preferential flow, or a calibrated catchment hydrograph.

9 Evapotranspiration

The evapotranspiration validation uses the Phase 1 v-tilted grid to test both potential ET calculation and actual ET extraction. The first case compares HydroPol2D's FAO-style potential/reference ET calculation against an independent implementation of the same Penman–Monteith-style equations [1]. The remaining cases call `Evaporation_Evapotranspiration_Module.m` directly and compare actual soil ET, open-water evaporation, final storage, invalid-cell masking, and mass residuals against an independent algebraic reference.

9.1 Evapotranspiration Method

Seven v-tilted ET regimes were tested with `run_vtilted_et_validation_suite.m`: potential ET formula reproduction, open-water storage limitation, demand-limited soil extraction,

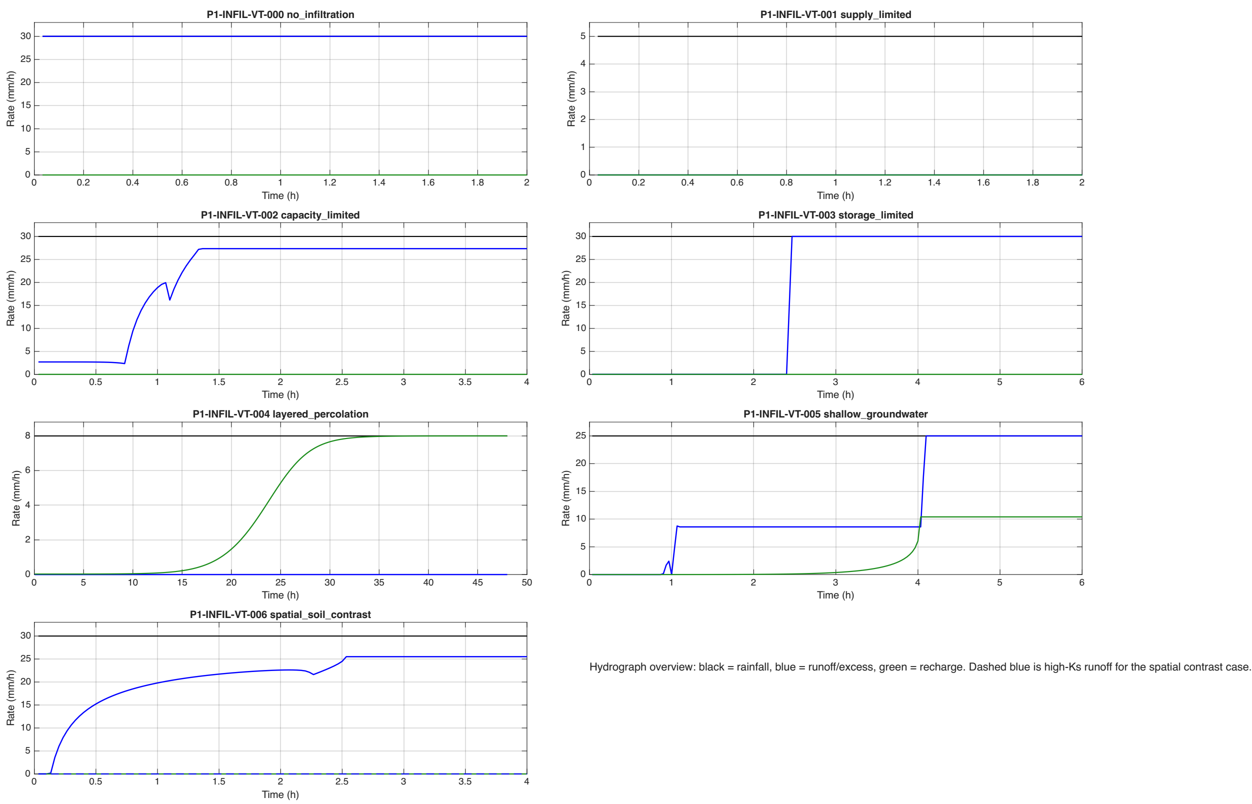

Figure 1: V-tilted infiltration hydrograph overview. Rates are shown in $\mathrm{mm\,h^{-1}}$; black lines are rainfall, blue lines are runoff or surface-excess generation, and green lines are recharge. The spatial-contrast case also shows the high-conductivity response with a dashed blue line.

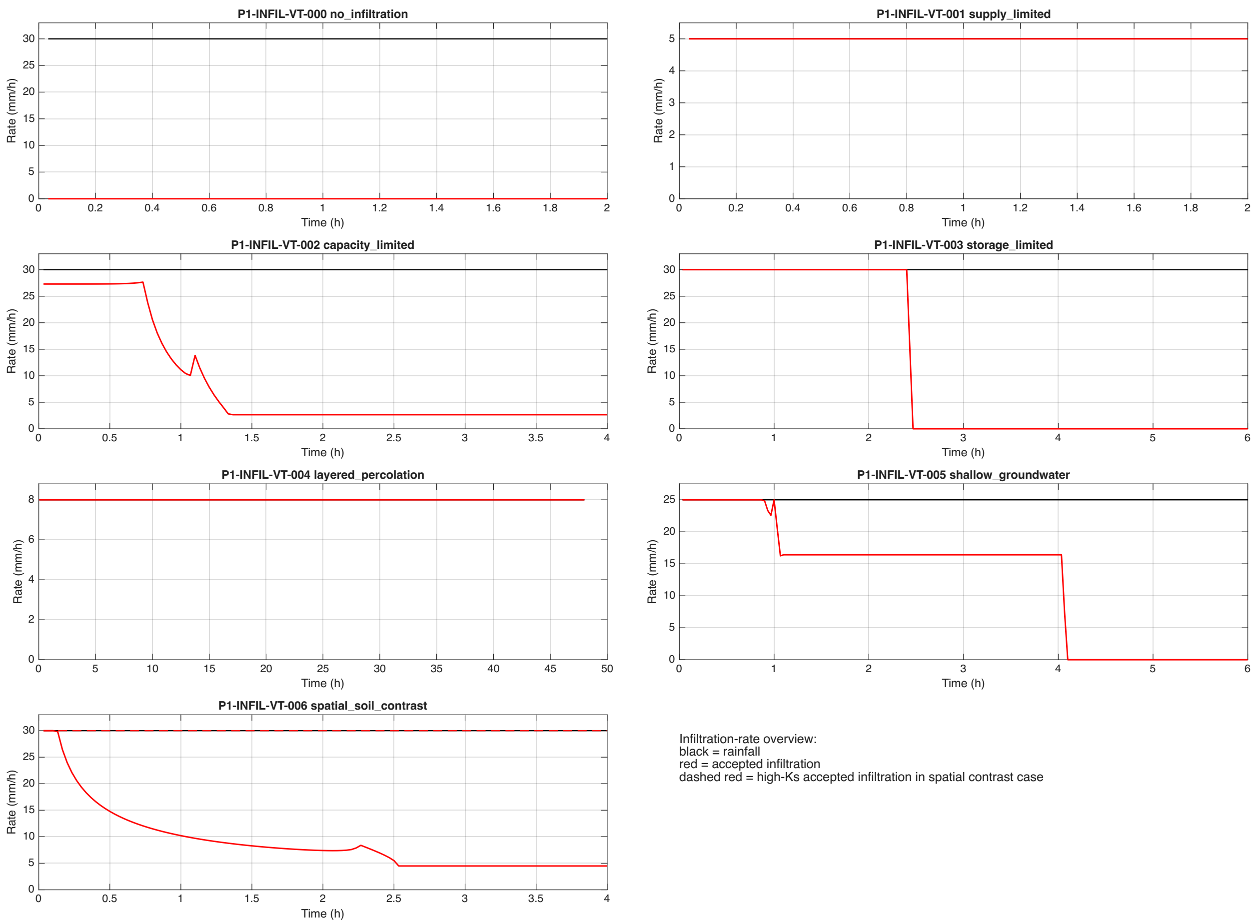

Figure 2: V-tilted infiltration-rate overview. Rates are shown in $\mathrm{mm\,h^{-1}}$; black lines are rainfall and red lines are accepted infiltration. The spatial-contrast case also shows the high-conductivity accepted-infiltration response with a dashed red line.

storage-limited soil extraction, partial near-surface root access, internal ETP mode, and dry/impervious/invalid-cell masking. The extraction reference uses the same physical bookkeeping expected from the model: ponded cells evaporate only from surface water and cannot evaporate more than the stored depth; unponded pervious soil cells extract from root-accessible near-surface storage first and then from the root-zone layer; transmission-zone storage is not root-accessible; impervious, dry, and invalid cells do not extract soil ET.

For extraction cases, the storage residual is

$$r_{\mathrm{ET}} = S_0 - S_t - E_s - E_w, \tag{11}$$

where S_0 and S_t are initial and final combined soil-plus-surface storage, E_s is soil evapotranspiration, and E_w is open-water evaporation over the time step. The report-ready threshold is $|r_{\mathrm{ET}}| < 10^{-6}$ m^3 with ET and storage errors at numerical precision.

9.2 Evapotranspiration Results

All seven v-tilted ET cases passed. Table 6 summarizes the primary diagnostics. The potential ET formula case produced identical mean reference and HydroPol2D ETP values, 5.7669 mm d^{-1}. For the extraction cases, reference and HydroPol2D soil ET, surface evaporation, and final storage matched exactly to machine precision. The largest absolute mass residual was 9.24×10^{-14} m^3.

Table 6: Evapotranspiration validation diagnostics for the v-tilted Phase 1 cases.

Case	Regime	Mean ETP (mm d^{-1})	Soil ET (mm sum)	Surface E (mm sum)	Residual (m^3)
P1-ET-VT-000	Potential ET formula	5.7669	–	–	0.0
P1-ET-VT-001	Open-water limit	–	0.0	27.0	-4.44×10^{-14}
P1-ET-VT-002	Soil demand limited	–	36.0	0.0	9.24×10^{-14}
P1-ET-VT-003	Soil storage limited	–	27.0	0.0	-4.44×10^{-14}
P1-ET-VT-004	Partial root access	–	45.0	0.0	0.0
P1-ET-VT-005	Internal ETP mode	–	18.0	12.0	3.55×10^{-15}
P1-ET-VT-006	Masks and dry cells	–	12.0	0.0	4.44×10^{-15}

These tests validate the implemented ET formula pathway and the actual ET extraction bookkeeping for controlled Phase 1 regimes. They do not validate field-scale ET, crop coefficients, aerodynamic parameter realism, remote-sensing ET products, or calibrated seasonal water balance.

10 Groundwater and Recharge

The groundwater/recharge validation uses the v-tilted catchment as the shared Phase 1 spatial context, but isolates groundwater dynamics from surface routing. The implemented groundwater coupling now uses two clocks. Vadose drainage, recharge production, and capillary rise are evaluated every surface-model step. Saturated groundwater storage and lateral groundwater heads are updated only when the groundwater scheduler is due. The default scheduler target is one day, shortened by the groundwater stability estimate, a maximum water-table-change cap, or the final model step. This avoids forcing a slow groundwater process to run at the same time step as overland routing while preserving the water-volume ledger.

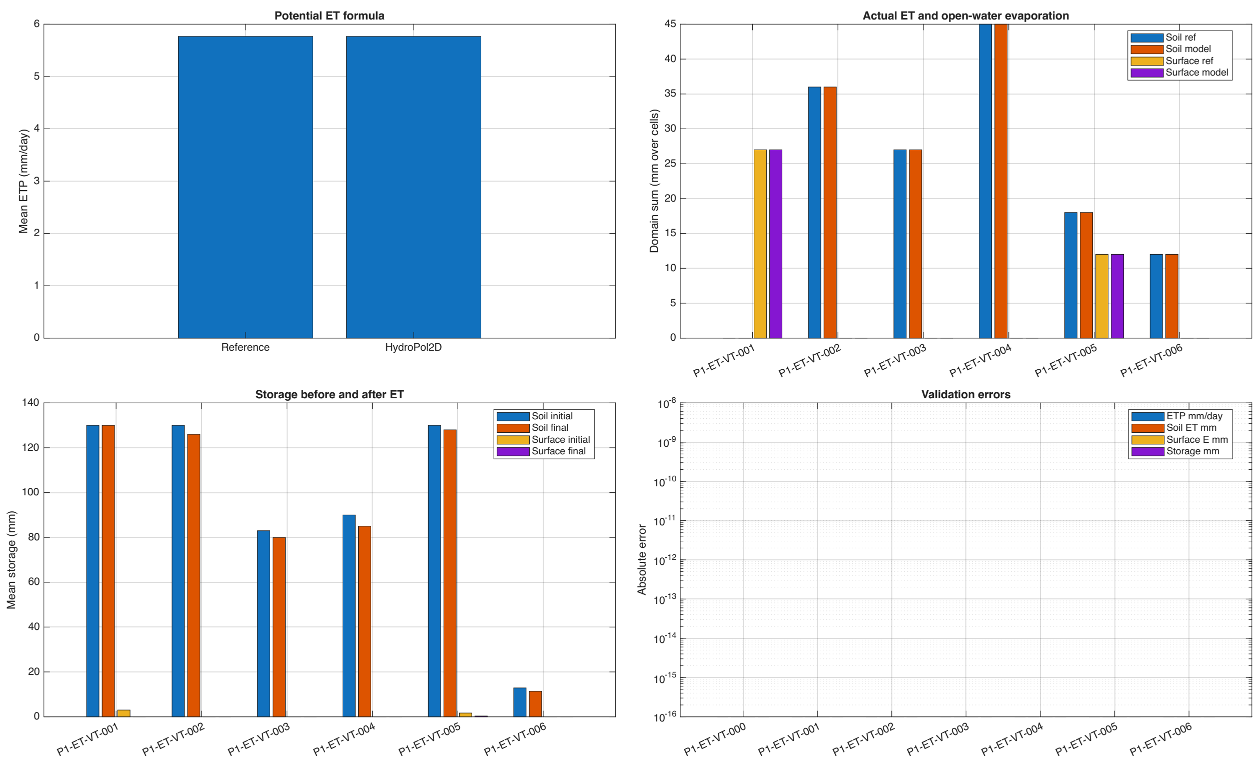

Figure 3: V-tilted evapotranspiration validation overview. The panels compare potential ET, actual soil ET, open-water evaporation, storage before and after ET, and absolute validation errors. Errors at zero are plotted at the display floor.

10.1 Groundwater Method

Four v-tilted groundwater cases were run with `run_hydropol2d_vtilted_groundwater_cases.m`. The first three cases test local groundwater storage and recharge/capillary bookkeeping. The fourth case activates lateral groundwater flow through the Boussinesq solver and compares the asynchronous result against a fine-timestep internal numerical reference.

For the local recharge case, a known recharge depth $R\Delta t$ is accumulated and applied to groundwater storage. With no lateral groundwater flow and no exfiltration, the analytical water-table response is

$$\Delta h = \frac{R\Delta t}{S_y}, \tag{12}$$

where S_y is the groundwater specific yield. The same expression is used to check the final head after six hourly groundwater updates.

For the scheduler-equivalence case, the same recharge forcing is run twice: once with legacy per-surface-step groundwater updating and once with the asynchronous accumulation scheduler. Because the local groundwater response is linear in accumulated recharge, the final groundwater storage and water-table elevation should match to numerical precision.

For the capillary-rise case, shallow groundwater supplies dry root-zone storage. The implemented simple upward flux is capped by available groundwater, layer storage deficit, saturated conductivity, dryness, and a 2 m capillary extinction depth:

$$q_c = K_s \max\left(1 - \frac{z_{wt}}{z_c}, 0\right) \max\left(1 - \frac{S_l}{S_{l,\max}}, 0\right), \tag{13}$$

with $z_c = 2\,\mathrm{m}$. The expected water-table drawdown is the transferred capillary water depth divided by S_y.

For the Boussinesq recharge-mound case, recharge is applied to the channel strip of a small v-tilted grid. The asynchronous solution is compared against the same model run with per-step groundwater updating. This is a numerical-reference benchmark, not a field validation case.

10.2 Groundwater Results

All four groundwater/recharge cases passed. Table 7 summarizes the diagnostics. The analytical recharge-rise case had a final mean-head error of -1.42×10^{-14} m and a mass residual of 1.96×10^{-11} m^3. The scheduler-equivalence case matched the legacy per-step update with head RMSE 1.71×10^{-13} m. The capillary-rise case transferred 32.4 m^3 from groundwater to vadose storage and closed mass to -1.45×10^{-12} m^3. The Boussinesq recharge-mound case matched the fine-timestep reference with head RMSE 1.83×10^{-8} m and maximum head error 9.90×10^{-8} m.

Table 7: Groundwater/recharge validation diagnostics for the v-tilted Phase 1 cases.

Case	Regime	Head RMSE (m)	Max head error (m)	Storage error (%)	Residual (m^3)
P1-GW-ASYNC-001	Local recharge	0.0	0.0	0.0	1.96×10^{-11}
P1-GW-ASYNC-002	Scheduler equivalence	1.71×10^{-13}	1.71×10^{-13}	1.67×10^{-11}	1.84×10^{-10}
P1-GW-CAP-001	Capillary rise	0.0	0.0	0.0	-1.45×10^{-12}
P1-GW-MOUND-001	Boussinesq mound	1.83×10^{-8}	9.90×10^{-8}	5.23×10^{-13}	3.09×10^{-11}

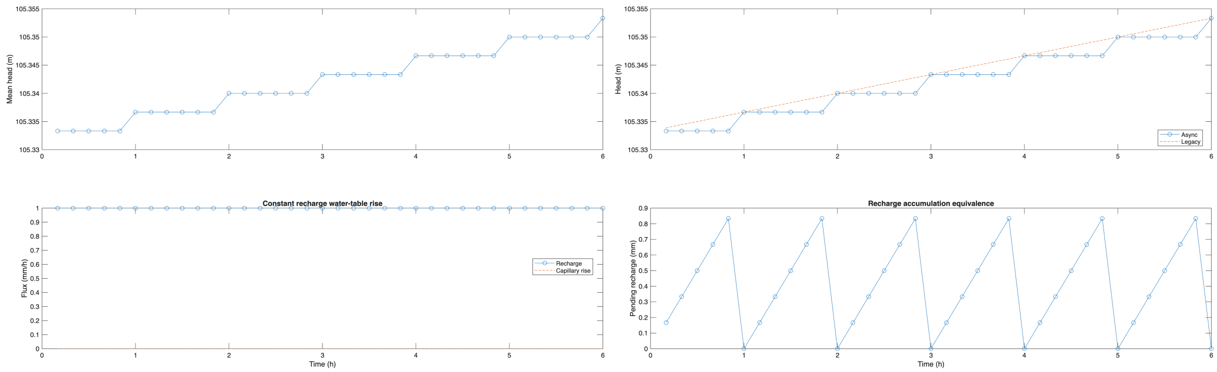

Figure 4: Groundwater recharge scheduler diagnostics. Left: analytical constant-recharge water-table rise. Right: asynchronous recharge accumulation compared with legacy per-step groundwater updating.

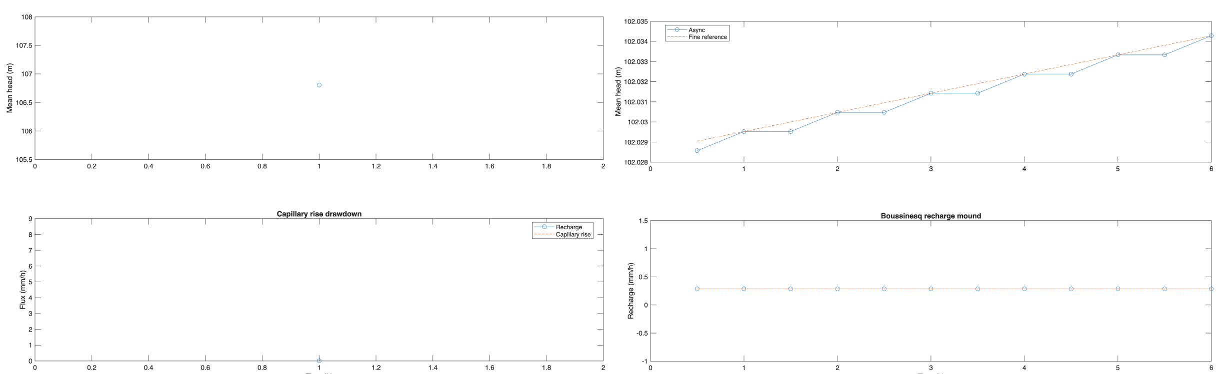

Figure 5: Groundwater capillary and lateral-flow diagnostics. Left: capillary rise from shallow groundwater to root-zone storage and corresponding water-table drawdown. Right: asynchronous Boussinesq recharge mound compared with the fine-timestep numerical reference.

10.3 Analytical Hillslope Groundwater Benchmarks

The preceding groundwater cases validate recharge bookkeeping, capillary exchange, scheduler behavior, and one lateral-flow numerical-reference comparison. To add stronger evidence for lateral groundwater dynamics, two idealized hillslope aquifer cases were added under `Groundwater/Hillslope_Analytical`. These cases call `Boussinesq_2D_explicit.m` directly with controlled boundary conditions: no-flow at the hillslope divide, a fixed-head drain at the toe, no lateral gradient across rows, uniform aquifer properties, and no surface-routing interaction.

The steady case uses the Dupuit unconfined aquifer solution for a horizontal hillslope with uniform recharge R, hydraulic conductivity K, length L, and fixed drain saturated thickness H_d:

$$H(x) = \left[H_d^2 + \frac{R}{K}\left(L^2 - x^2\right)\right]^{1/2}. \tag{14}$$

The numerical test initializes the finite-volume steady profile implied by HydroPol2D's interface fluxes and compares it to the continuous Dupuit profile at three grid resolutions. This checks both equilibrium preservation and convergence toward the analytical solution. Drain discharge is checked against total recharge to the hillslope.

The transient case uses the first eigenmode of the linearized Boussinesq equation, consistent with the small-perturbation hillslope solutions discussed by Verhoest and Troch, Troch et al., and Rupp and Woods [11, 9, 8]. A small perturbation of saturated thickness,

$$u(x,0) = A\cos\left(\frac{\pi x}{2L}\right), \tag{15}$$

decays as

$$u(x,t) = A\cos\left(\frac{\pi x}{2L}\right)\exp\left[-D\left(\frac{\pi}{2L}\right)^2 t\right], \qquad D = \frac{KH_0}{S_y}, \tag{16}$$

where H_0 is the background saturated thickness. The perturbation amplitude is kept small so the nonlinear HydroPol2D Boussinesq solver is tested in the regime where the linearized analytical solution is valid.

Table 8: Analytical hillslope groundwater benchmark diagnostics.

Case	Benchmark	Head RMSE (m)	Max head error (m)	Flow metric
P1-GW-HILL-STEADY-001	Steady Dupuit	6.04×10^{-3}	9.75×10^{-3}	drain error = 3.74×10^{-4}%
P1-GW-HILL-TRANSIENT-001	Linearized Boussinesq	1.94×10^{-5}	4.46×10^{-5}	NSE = 0.99985

Together, these tests validate recharge accumulation, the asynchronous saturated-groundwater scheduler, local water-table rise/drawdown using S_y, the simple layered capillary-rise closure, a small Boussinesq recharge-mound response, and idealized lateral hillslope groundwater dynamics against analytical solutions. They do not validate regional aquifer parameters, field baseflow, groundwater pumping, confined aquifer behavior, or calibration against observed wells.

11 Hydrodynamics

Full-momentum hydrodynamic validation was run first, before local inertial, diffusive, kinematic, or CA/D8 routing alternatives. The main model loop was updated so that `flag_full_momentum` is respected as a mutually exclusive routing selector rather than being overwritten by the local-inertial

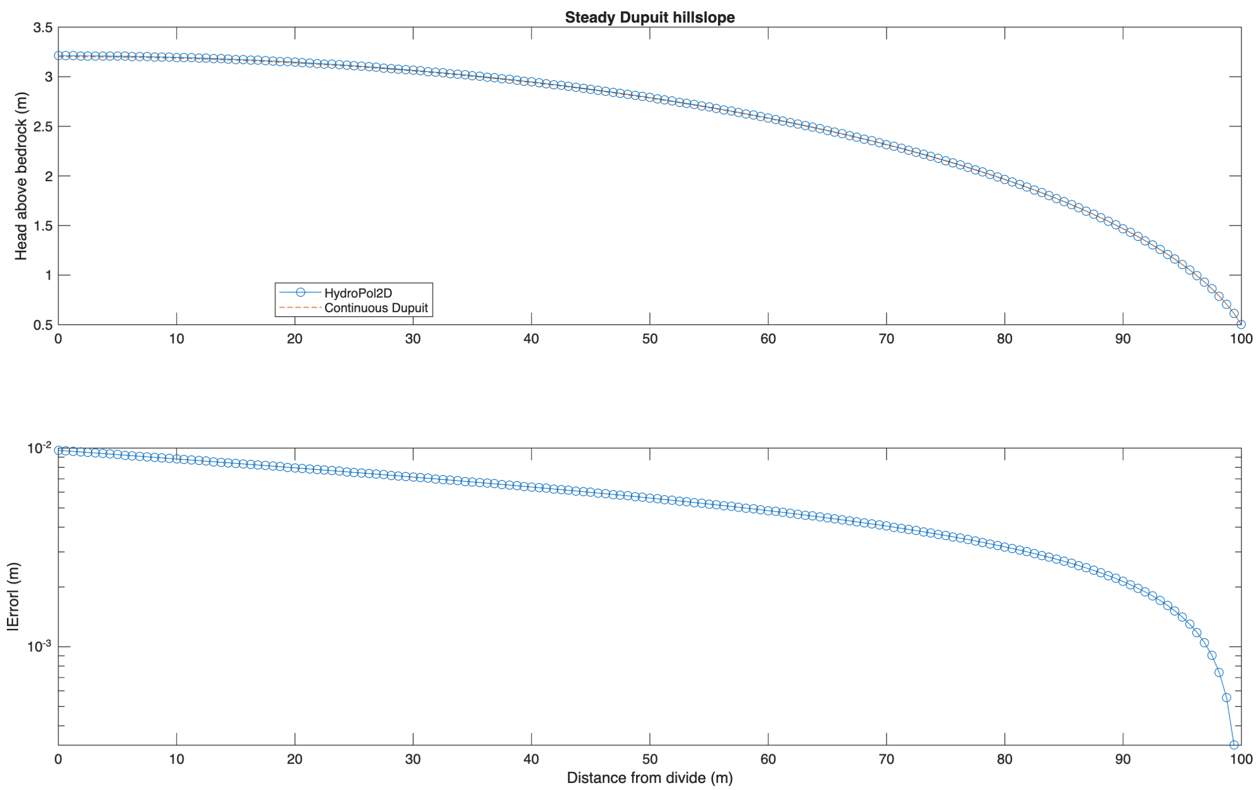

Figure 6: Steady Dupuit hillslope benchmark. HydroPol2D is compared with the continuous analytical profile on the finest grid, and the lower panel shows the absolute profile error.

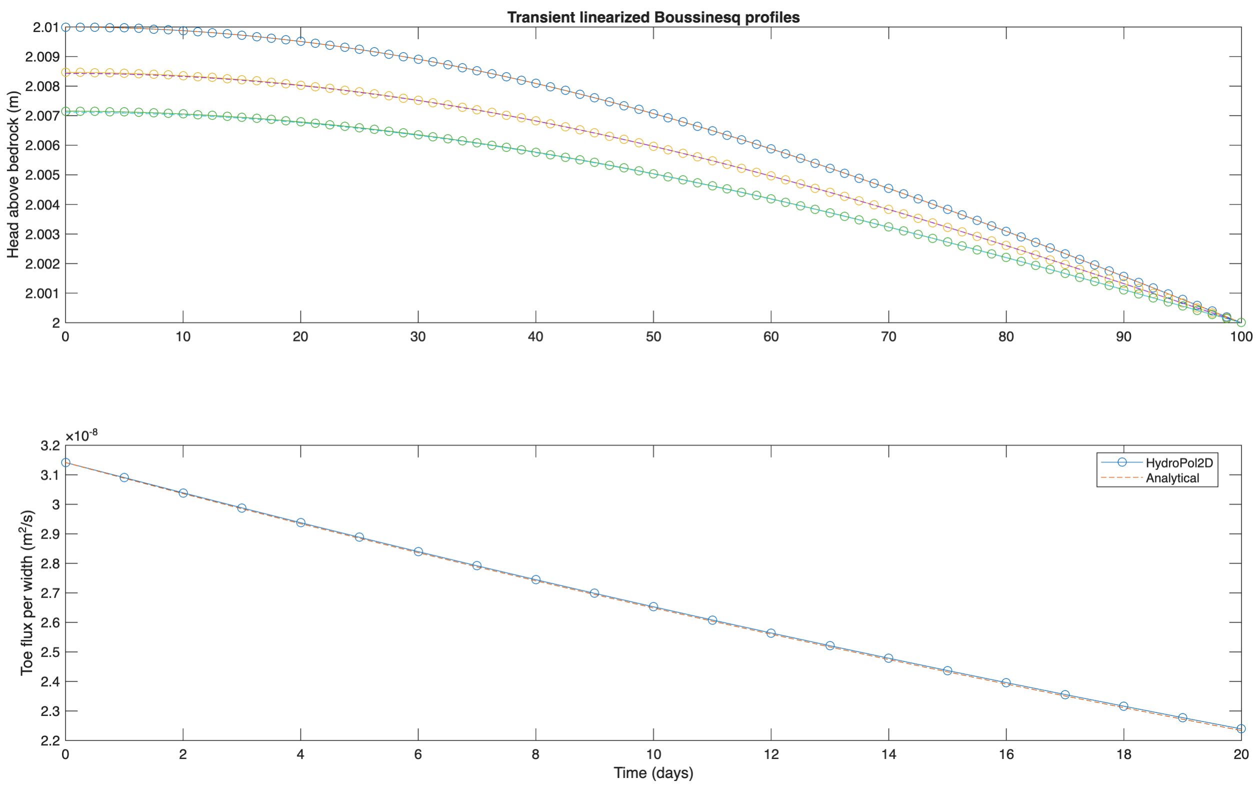

Figure 7: Transient linearized Boussinesq hillslope benchmark. The upper panel compares HydroPol2D and analytical head profiles during perturbation decay; the lower panel compares modeled and analytical toe flux per unit width.

default. The primary validation driver is `HydroPol2D_Model/Validation/Hydrodynamics/FullMomentum_Phase1/` the non-breaking-wave analytical benchmark is run by `HydroPol2D_Model/Validation/Hydrodynamics/NonBreak`

Five current-code cases were run. The tilted-plane case compares the full-momentum hydrograph with the analytical kinematic-wave limit for shallow uniform sheet flow. This is not an exact full Saint-Venant solution, but it is a useful limiting benchmark when inertia and pressure-gradient effects are small. The v-tilted case uses the Phase 1 catchment geometry as a conservation and bilateral-symmetry dynamics check. The Ritter case is the primary discontinuous-flow analytical hydrodynamic benchmark, comparing frictionless dry-bed dam-break profiles with Ritter's solution [5]. The non-breaking-wave case provides a smoother traveling-wave analytical benchmark with Manning roughness. The prescribed-stage case validates boundary enforcement and volume bookkeeping under full-momentum routing.

Table 9: Full-momentum hydrodynamics and prescribed-stage boundary metrics.

Case	Primary RMSE	Diagnostic value	Rel. L_2	NSE	Mass error (%)
Tilted plane hydrograph	0.0264	0.1067	0.0311	0.9961	7.80×10^{-12}
V-tilted symmetry/conservation	3.50×10^{-18}	4.2340	–	–	2.74×10^{-12}
Ritter dam-break profile	0.0305	0.0717	0.0567	0.9992	1.26×10^{-13}
Non-breaking wave profile	0.0116	0.1680	0.0260	0.9997	4.5638
Stage-hydrograph boundary	0.0000	0.0718	–	–	1.35×10^{-13}

For the mixed-evidence rows, "diagnostic value" is the peak outlet discharge for the V-tilted case and the downstream stage-response amplitude for the stage-hydrograph case; for analytical rows it is the maximum absolute error.

The tilted-plane benchmark passed with NSE = 0.996 and relative L_2 error = 0.031. The peak magnitude error was only 0.0072 %; the large peak-time difference reflects the analytical curve reaching its steady plateau early and the numerical hydrograph having an extended near-plateau, so it is reported but not used as the primary acceptance criterion. The V-tilted case conserved mass to machine precision and preserved left/right symmetry. The Ritter dam-break benchmark passed with maximum profile RMSE = 0.0305 m, relative L_2 error = 0.0567, and NSE = 0.999. The non-breaking-wave benchmark passed with profile RMSE = 0.0116 m, relative L_2 error = 0.0260, NSE = 0.9997, and maximum wetting-front offset of 35 m, which is seven grid cells on the 5 m benchmark mesh. Its storage residual is reported against the analytical profile volume over the active finite domain, rather than a closed water-balance ledger, and remained below the 5 % acceptance threshold.

11.1 Local-Inertial and Cellular-Automata Routing

The same Phase 1 sequence was repeated for local-inertial and cellular-automata routing using `HydroPol2D_Model/Validation/Hydrodynamics/LocalInertial_CA_Phase1/run_local_inertial_ca_hydrodyna` Local inertial is treated as an approximation benchmark against the tilted-plane and Ritter analytical cases. Cellular automata is treated primarily as a routing and storage-redistribution method; analytical hydrograph/profile comparisons are reported as diagnostics because CA does not solve momentum.

The local-inertial plane, V-tilted, and stage-boundary cases passed. The local-inertial Ritter case did not pass the analytical threshold: the RMSE was 0.1823 m and the relative L_2 error was 0.3838. This case is therefore retained as diagnostic evidence and a limitation, not as a report-ready local-inertial dam-break validation. This result is physically reasonable because the Ritter solution

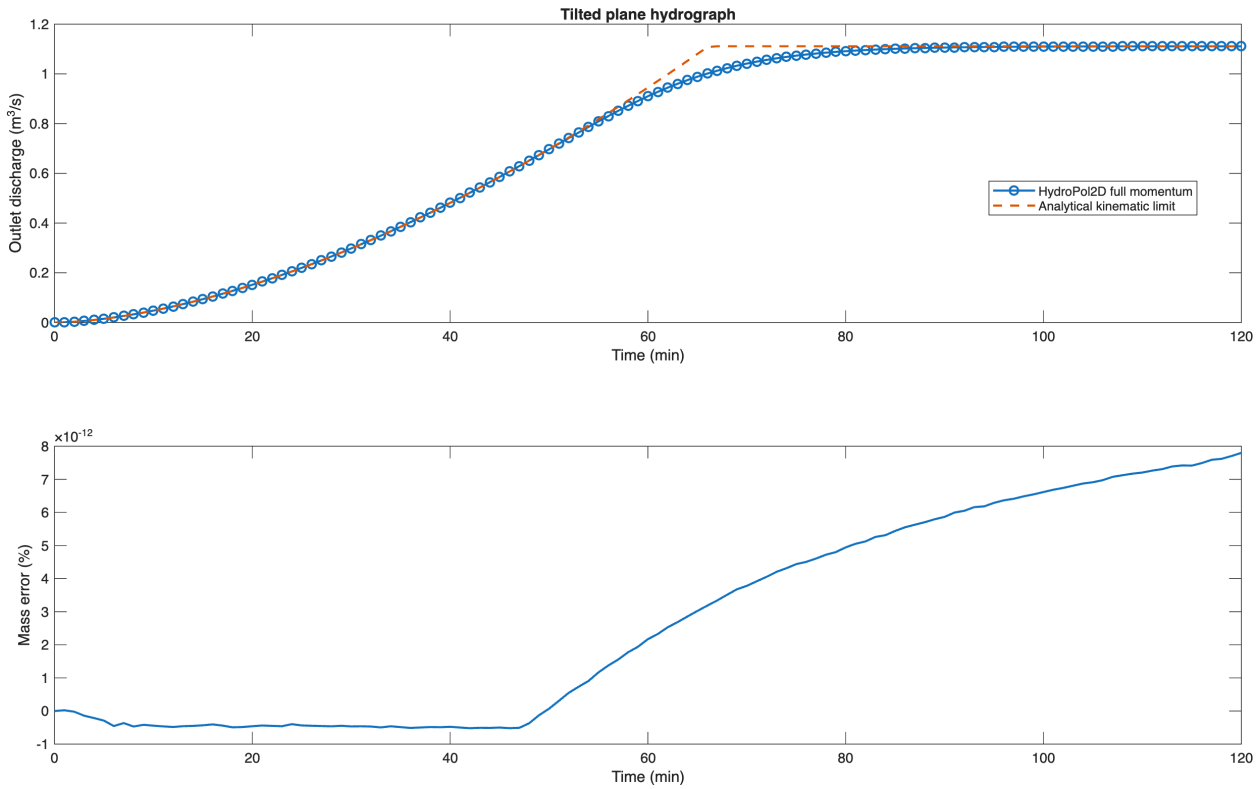

Figure 8: Full-momentum tilted-plane hydrograph compared with the analytical kinematic-wave limiting solution, with mass-balance residual.

Table 10: Local-inertial and cellular-automata hydrodynamic metrics.

Method	Case	RMSE	Rel. L_2	NSE	Mass error (%)	Status
Local inertial	Plane	0.0169	0.0200	0.9984	7.47×10^{-12}	pass
Local inertial	V-tilted	0.0000	–	–	3.66×10^{-12}	pass
Local inertial	Ritter	0.1823	0.3838	0.9552	5.05×10^{-14}	diagnostic fail
Local inertial	Stage BC	0.0000	–	–	1.43×10^{-13}	pass
CA	Plane	0.0172	0.0204	0.9983	7.69×10^{-12}	pass
CA	V-tilted	0.0000	–	–	4.04×10^{-12}	pass
CA	Ritter	0.1592	0.2957	0.9875	7.58×10^{-14}	diagnostic pass
CA	Stage BC	0.0000	–	–	2.10×10^{-13}	pass

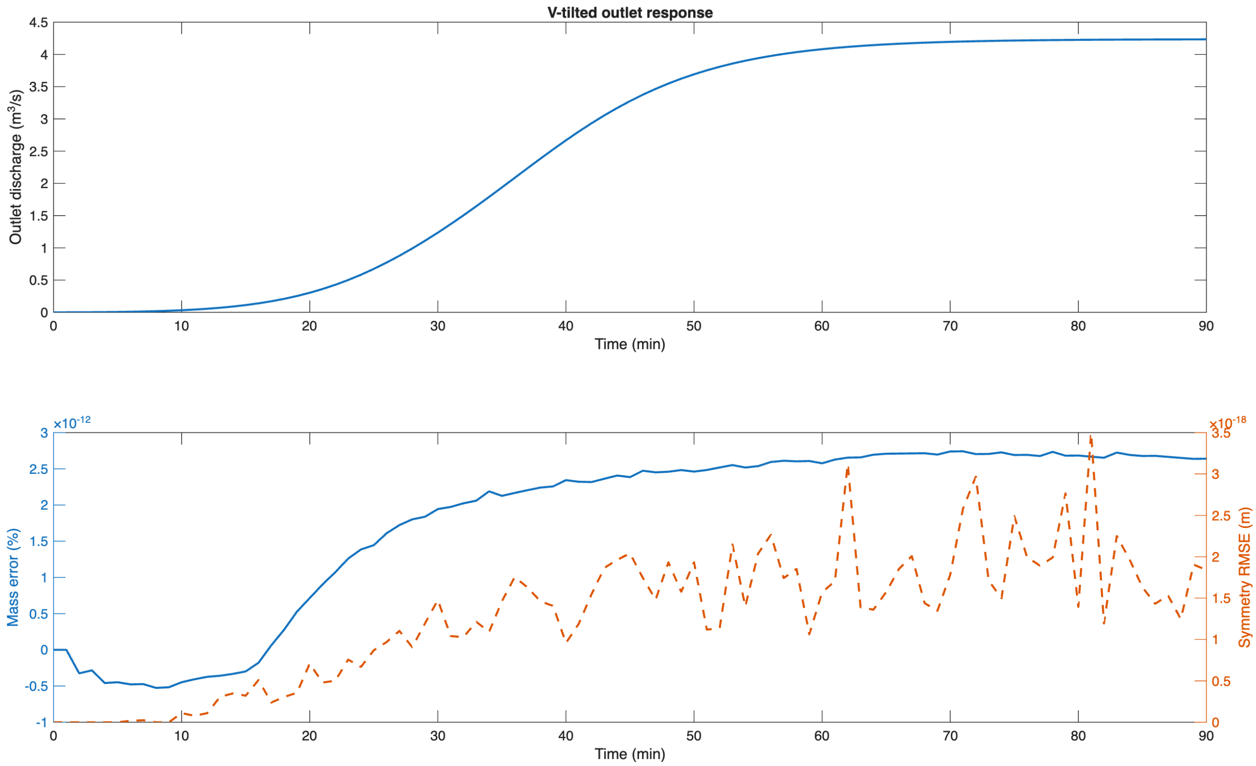

Figure 9: Full-momentum V-tilted catchment response. The case checks outlet response, mass conservation, and left/right symmetry on the shared Phase 1 catchment geometry.

is a full-momentum dry-bed dam-break benchmark, while the local-inertial approximation omits advective acceleration. CA passed the conservation, symmetry, and boundary-bookkeeping checks; its Ritter and plane analytical comparisons are useful diagnostics but are not interpreted as strict momentum validation.

11.2 Canonical V-tilted Benchmark Hydrograph

The original V-tilted benchmark hydrograph was added as a full hydrodynamic routing benchmark. The benchmark values come from `compare_flows.m`: rainfall intensity is $10.8\,\mathrm{mm\,h^{-1}}$, rainfall duration is 90 min, and the outlet hydrograph is digitized at 10–20 minute intervals through 240 min. On the canonical 20 m Phase 1 V-tilted raster domain, the rainfall volume is $26\,244\,\mathrm{m^3}$; the digitized benchmark outlet volume through 240 min is $26\,331\,\mathrm{m^3}$, effectively a no-loss event within digitization uncertainty.

The benchmark was rerun through the full HydroPol2D bypass/preprocessing pathway rather than the earlier lightweight routing-module shortcut. The supplied configuration uses the rainfall file, the bypass roughness pattern with side-slope roughness $n = 0.015$ and center-strip roughness $n = 0.15$, two outlet cells, and the 180 min bypass simulation window. The full preprocessing placed the outlets at row 1, columns 41 and 42 on the canonical raster. This is important because the earlier diagnostic shortcut flipped the raster and used the downstream row as the outlet, which caused the event volume to remain stored on the domain and produced the previously reported failed hydrograph.

The same full-config V-tilted case was then run with the full-momentum, local-inertial, cellular-automata, kinematic, and diffusive routing options. All five runs used the same 20 m raster, rainfall,

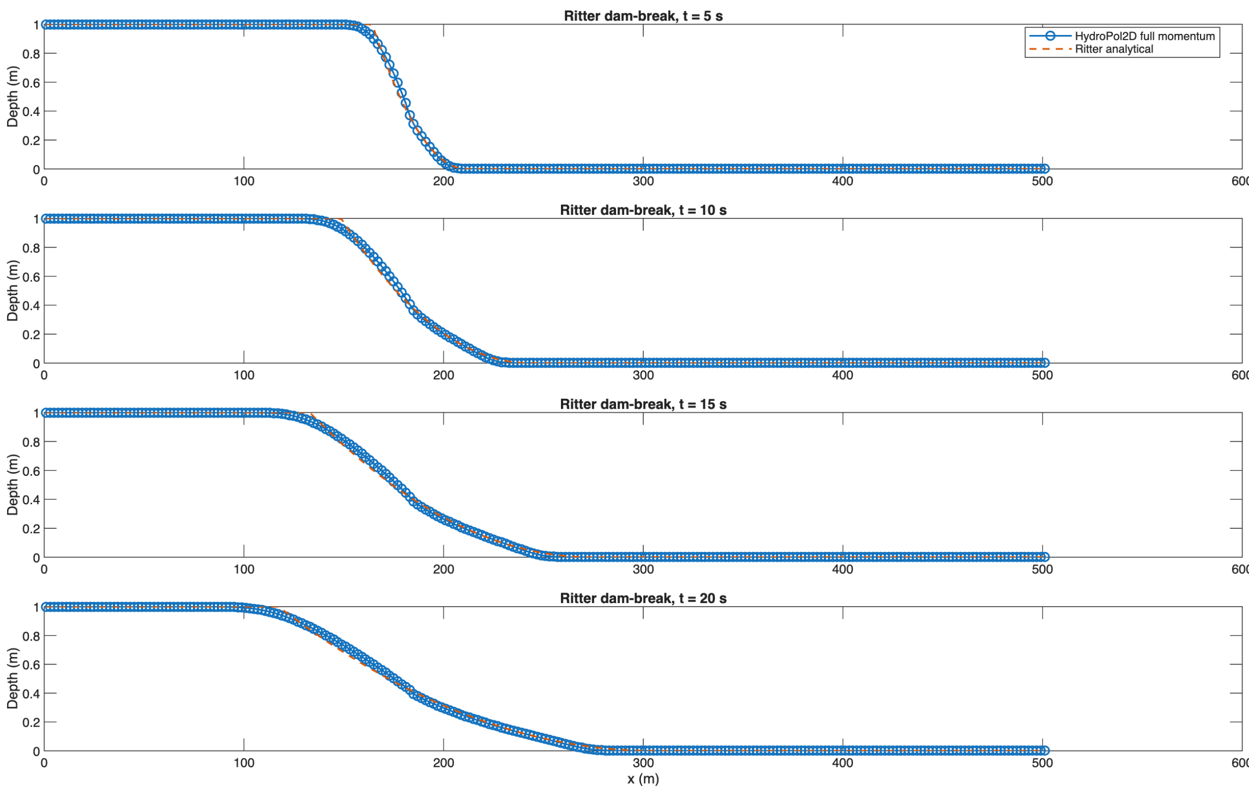

Figure 10: Ritter dry-bed dam-break benchmark. HydroPol2D full momentum is compared with analytical depth profiles at multiple times.

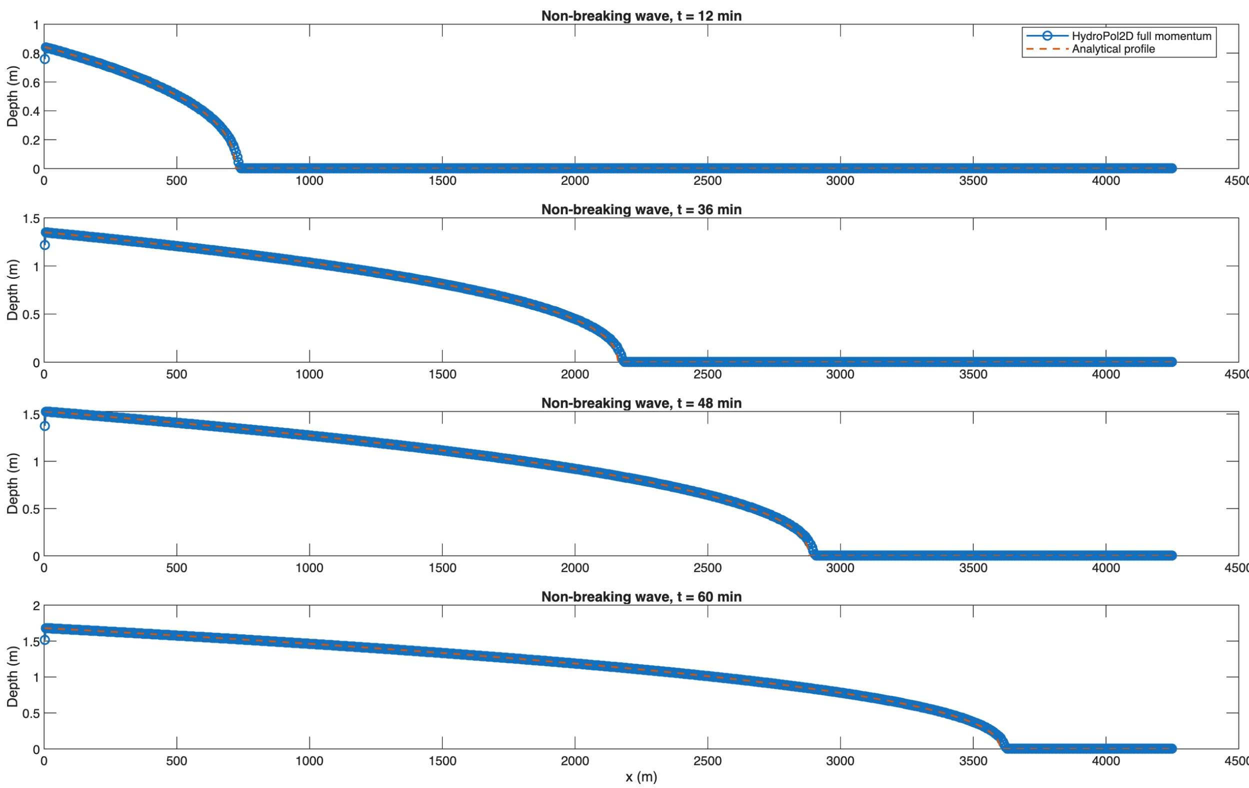

Figure 11: Non-breaking-wave benchmark. HydroPol2D full momentum is compared with the analytical Manning traveling-wave profile at 12, 36, 48, and 60 minutes.

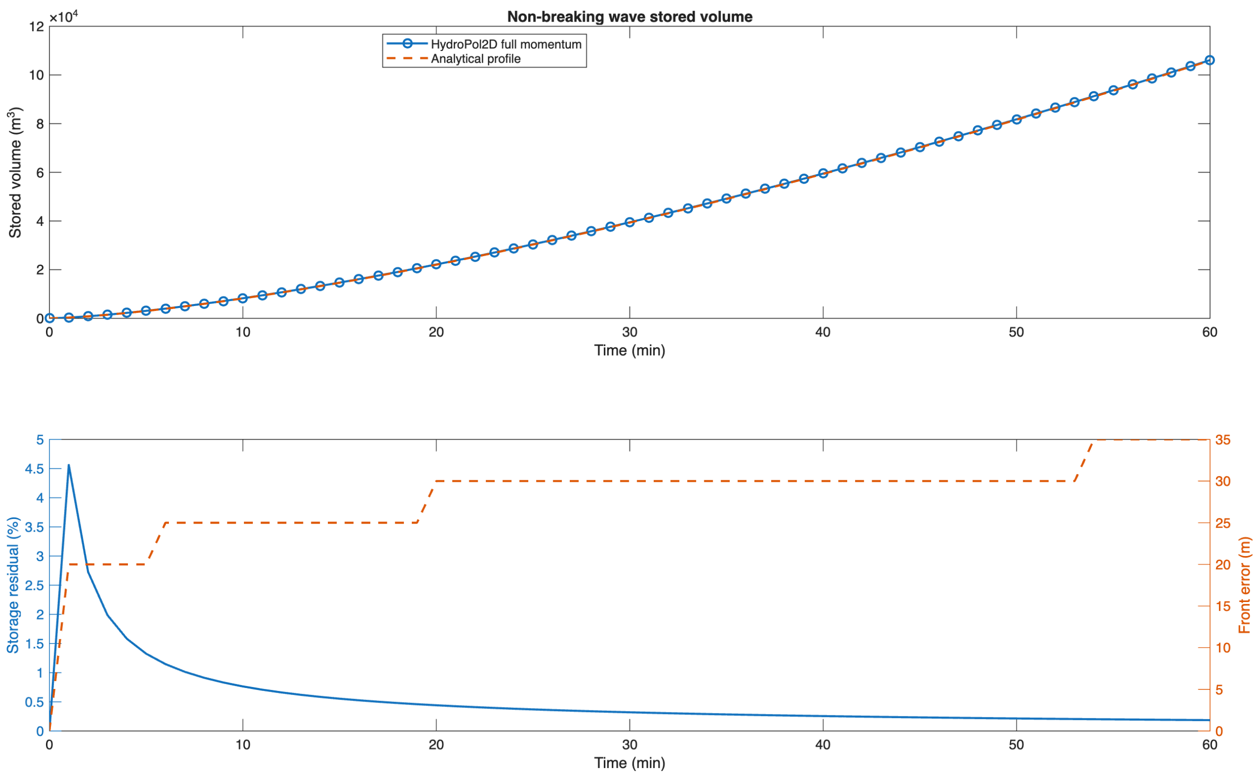

Figure 12: Non-breaking-wave stored-volume and wetting-front diagnostics. Storage residual is computed relative to analytical domain storage.

roughness, outlet cells, time-step controls, and disabled-loss hydrologic settings. Scores are computed over the common 0 min–180 min window because that is the end time in the supplied bypass configuration. The benchmark hydrograph still has a recession tail after 180 min; therefore, outlet-volume error over this truncated window is reported as an event-volume diagnostic, while the model water balance is evaluated from rainfall, outlet volume, and final stored water.

All routing options reproduced the benchmark hydrograph shape well, with NSE values from 0.9694 to 0.9817 and peak discharges within 2.1 % of the digitized benchmark peak. The full-momentum and local-inertial runs left more water stored on the domain at 180 min, giving larger truncated-window outlet-volume errors than the reduced routing options. However, every run closed the rainfall–outlet–storage ledger to −0.0181 %, which confirms that the differences are routing/recession differences rather than a mass-balance failure.

Table 11: Canonical V-tilted benchmark hydrograph diagnostics.

Routing mode	RMSE $m^3\,s^{-1}$	NSE –	Peak error %	Volume error %	Mass error %	Status
Full momentum	0.303	0.9730	0.052	7.50	-0.0181	pass
Local inertial	0.287	0.9759	2.05	5.04	-0.0181	pass
Cellular automata	0.323	0.9694	0.008	3.95	-0.0181	pass
Kinematic	0.249	0.9817	0.020	3.23	-0.0181	pass
Diffusive	0.255	0.9809	0.063	3.20	-0.0181	pass

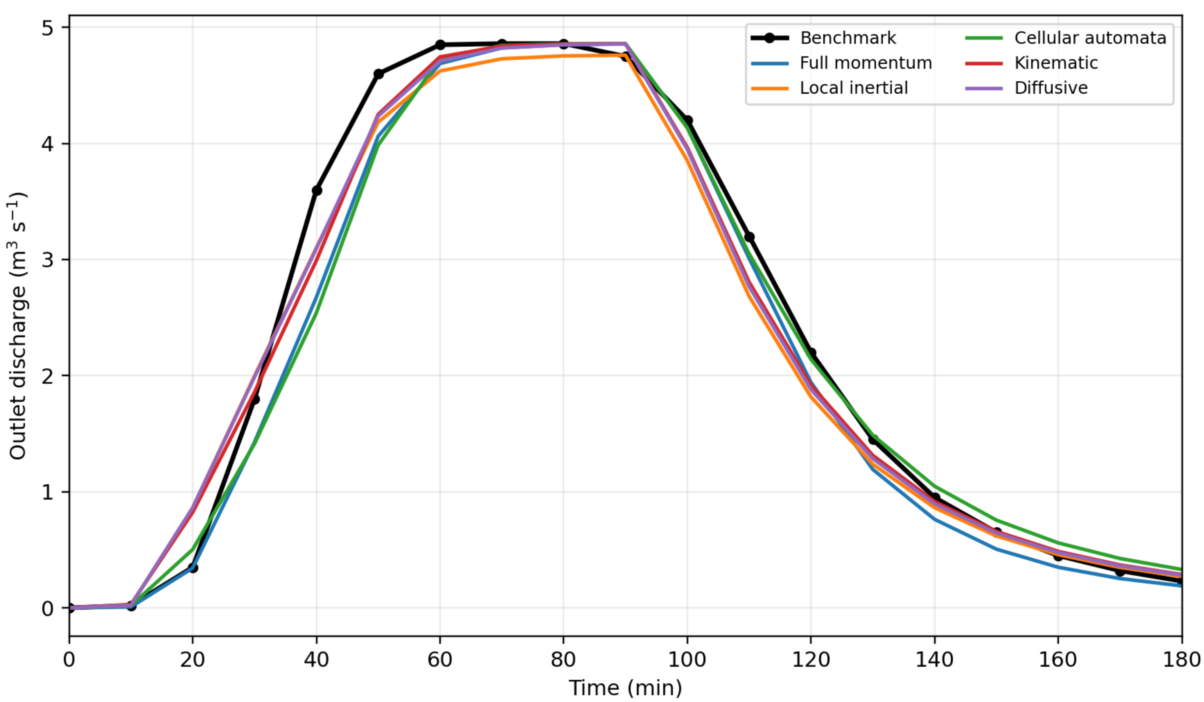

Figure 13: Canonical V-tilted benchmark hydrograph compared with the full-config 20 m HydroPol2D full-momentum, local-inertial, cellular-automata, kinematic, and diffusive routing runs.

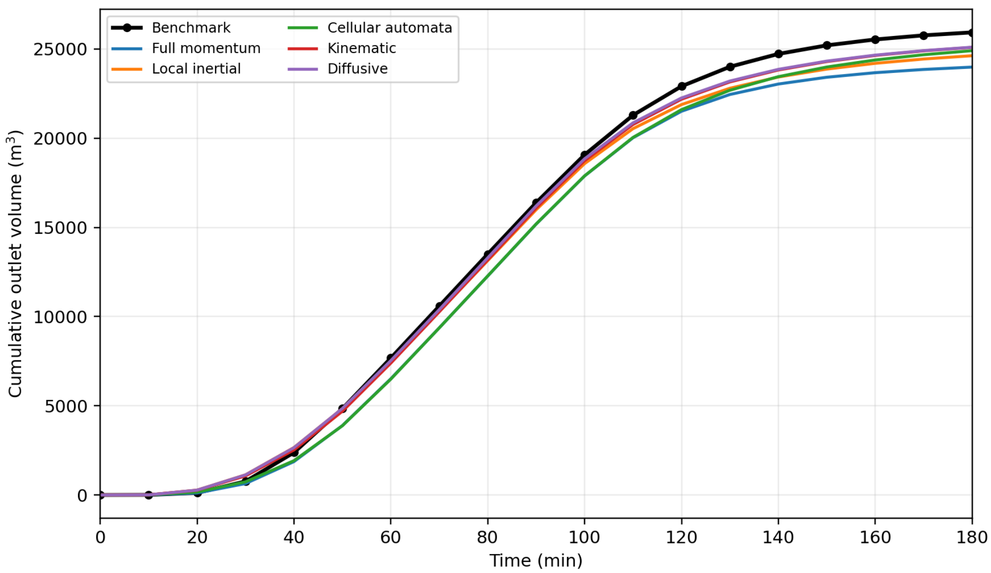

Figure 14: Canonical V-tilted cumulative outlet-volume diagnostics for all full-config routing runs. The model ledger including final storage closes to $-0.0181\,\%$ for every routing option.

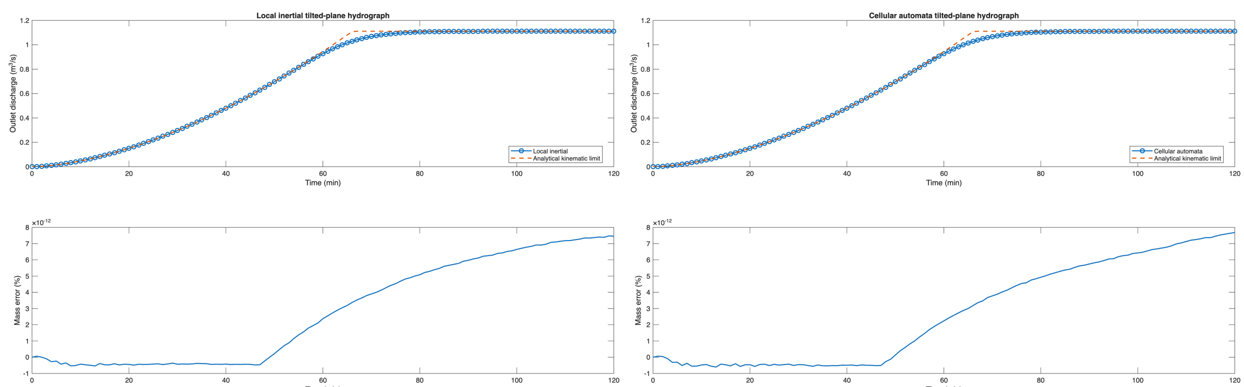

Figure 15: Tilted-plane hydrograph diagnostics for local-inertial routing (left) and cellular-automata routing (right).

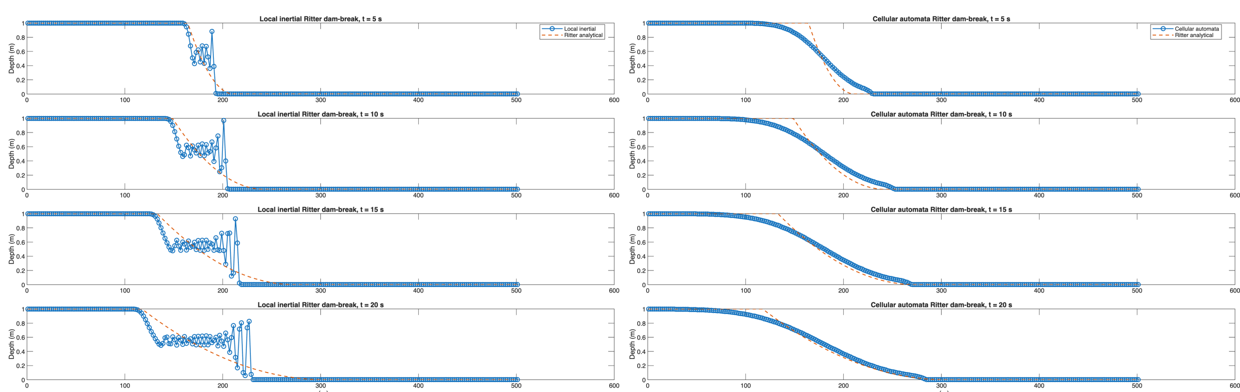

Figure 16: Ritter dam-break profile diagnostics for local-inertial routing (left) and cellular-automata routing (right). The local-inertial result is retained as a diagnostic failed benchmark.

12 Routing Alternatives

The kinematic and diffusive routing audit found two implementation-level problems. First, `Kinematic_Wave_Model.m` declared the wrong primary function name, so MATLAB could not reliably dispatch it as the kinematic solver. Second, the old kinematic/diffusive bodies used copied slope logic with inconsistent row/column directions and a Manning flux expression that did not cleanly represent either bed-slope kinematic routing or water-surface-slope diffusive routing. These functions were replaced by a shared conservative D4 face-flux core. Kinematic routing now uses bed slope and upstream cell depth; diffusive routing uses water-surface slope with hydraulic reconstruction at cell faces. Both are updated by conservative finite-volume storage accounting, a draining limiter, outlet-volume bookkeeping, and internal substepping for steep fronts.

The corrected solvers were applied to the same Phase 1 sequence using `HydroPol2D_Model/Validation/Hydrody` The tilted-plane case is the primary analytical validation for the kinematic wave. The diffusive wave is compared with the same plane solution as an approximation benchmark and is also checked by conservation/symmetry on the v-tilted catchment. Ritter dam-break profiles are retained only as diagnostics because neither reduced equation solves the full dynamic-wave problem.

Table 12: Corrected kinematic and diffusive routing metrics.

Method	Case	RMSE	Rel. L_2	NSE	Mass error (%)	Status
Kinematic	Plane	0.0156	0.0184	0.9986	7.66×10^{-12}	pass
Kinematic	V-tilted	0.0000	–	–	4.18×10^{-12}	pass
Kinematic	Ritter	0.2408	0.5069	0.9245	0.0000	diagnostic pass
Kinematic	Stage BC	0.0000	–	–	1.32×10^{-13}	pass
Diffusive	Plane	0.0172	0.0203	0.9984	7.75×10^{-12}	pass
Diffusive	V-tilted	0.0000	–	–	3.73×10^{-12}	pass
Diffusive	Ritter	0.5670	1.1151	-0.0803	1.53×10^{-12}	diagnostic fail
Diffusive	Stage BC	0.0000	–	–	1.58×10^{-13}	pass

The kinematic tilted-plane hydrograph reproduced the analytical kinematic-wave limit with NSE 0.9986 and relative L_2 error 0.0184. The diffusive plane run was similarly close for this mild rainfall-runoff case, with NSE 0.9984. Both solvers conserved mass to machine precision and preserved exact bilateral symmetry on the v-tilted catchment. The diffusive Ritter diagnostic failed, as expected for a reduced diffusive-wave approximation under a sharp dry-bed dam-break discontinuity; it is not used as validation evidence for diffusive routing.

13 Neal (2012) Channel-Subgrid Hydraulics

HydroPol2D represents channels narrower than a raster cell with the rectangular channel formulation of [4]. The channel is embedded below the coarse floodplain elevation. Channel and floodplain discharges are computed separately on each orthogonal face, then combined in the local-inertial continuity update. The supported configuration uses `flag_inertial=1`, `flag_subgrid=1`, and `flag_overbanks=1`.

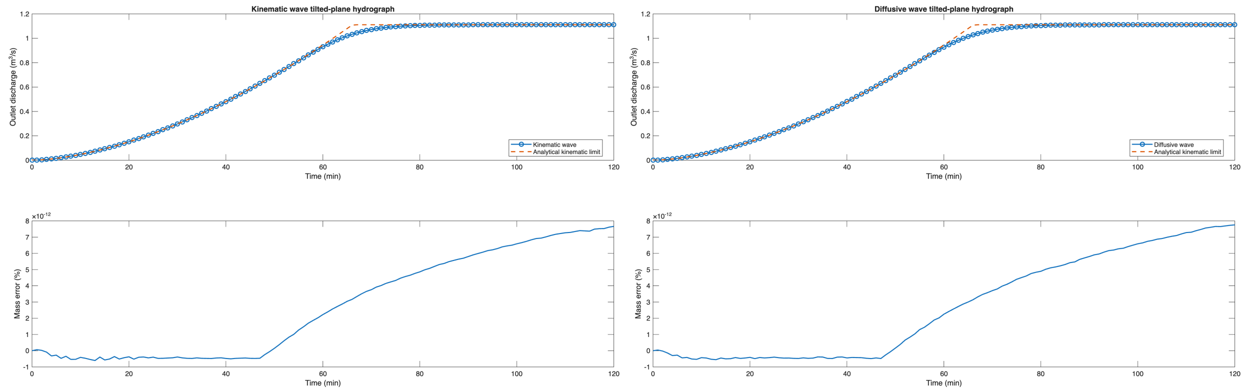

Figure 17: Tilted-plane hydrograph validation for corrected kinematic routing (left) and corrected diffusive routing (right).

13.1 Channel and Floodplain Formulation

Let Δx be the coarse-cell width, z_f the floodplain elevation, H the channel bank height, and w the channel width, where $0 < w < \Delta x$. The channel-bed elevation is

$$z_c = z_f - H. \tag{17}$$

The representative depth h is measured above z_c. For adjacent cells i and j, the channel water surface $y_c = z_c + h$, face slope, and face depth are

$$S_c = \frac{y_{c,j} - y_{c,i}}{\Delta x}, \qquad h_{c,f} = \max\left[\max(y_{c,i}, y_{c,j}) - \max(z_{c,i}, z_{c,j}), 0\right]. \tag{18}$$

The shared channel width, flow area, and hydraulic radius are

$$w_{c,f} = \min(w_i, w_j), \qquad A_c = w_{c,f} h_{c,f}, \qquad R_c = \frac{A_c}{w_{c,f} + 2h_{c,f}}. \tag{19}$$

The channel discharge is updated as

$$Q_c^{t+\Delta t} = \frac{Q_c^t - g A_c \Delta t S_c}{1 + g \Delta t n_c^2 |Q_c^t| / (R_c^{4/3} A_c)}, \tag{20}$$

where n_c is the channel Manning coefficient.

Floodplain depth is $h_f = \max(h + z_c - z_f, 0)$. The corresponding face slope S_f and face depth $h_{f,f}$ are computed from the floodplain water surface $y_f = z_f + h_f$. Floodplain unit discharge is

$$q_f^{t+\Delta t} = \frac{q_f^t - g h_{f,f} \Delta t S_f}{1 + g \Delta t n_f^2 |q_f^t| / h_{f,f}^{7/3}}, \tag{21}$$

where n_f is the floodplain Manning coefficient. Total floodplain and face discharges are

$$Q_f = q_f(\Delta x - w_{c,f}), \qquad Q = Q_c + Q_f. \tag{22}$$

Cell continuity is evaluated in volume space. The storage relation is

$$V(h) = \begin{cases} \Delta x\, w\, h, & h \le H, \\ \Delta x\, w\, H + \Delta x^2 (h - H), & h > H. \end{cases} \tag{23}$$

Routing fluxes and hydrologic sources or sinks update V first. The same piecewise relation is then inverted to recover h. Thus, the active surface area changes conservatively from $\Delta x w$ in-bank to Δx^2 overbank.

13.2 Numerical-reference Protocol

Each retained case compares three local-inertial simulations: a fine grid that resolves channel and floodplain topography, an ordinary coarse grid without an embedded channel, and a coarse grid with the Neal (2012) formulation. The fine grid is a numerical reference, not an analytical solution. The comparison tests whether the embedded channel preserves hydraulic behavior that is lost when the grid is coarsened.

13.3 Simple In-bank Channel

Cases `P1-SUBGRID-NEAL-SIMPLE-001` and `P1-SUBGRID-NEAL-SIMPLE-002` use a 10 m-wide, 1 m-deep channel with bed slope 0.001 and Manning $n = 0.035$. The explicit grid has 10 m cells; the ordinary and Neal (2012) grids have 30 m cells. A constant $5\,\mathrm{m^3\,s^{-1}}$ inflow remains within bank.

For the 30-cell reach, the Neal (2012) hydrograph RMSE is $0.098\,\mathrm{m^3\,s^{-1}}$, compared with $0.677\,\mathrm{m^3\,s^{-1}}$ for the ordinary coarse grid. NSE increases from 0.561 to 0.991, and stage RMSE decreases from 0.689 m to 0.053 m. The embedded channel corrects the excessive flow width and storage represented by the ordinary coarse cells.

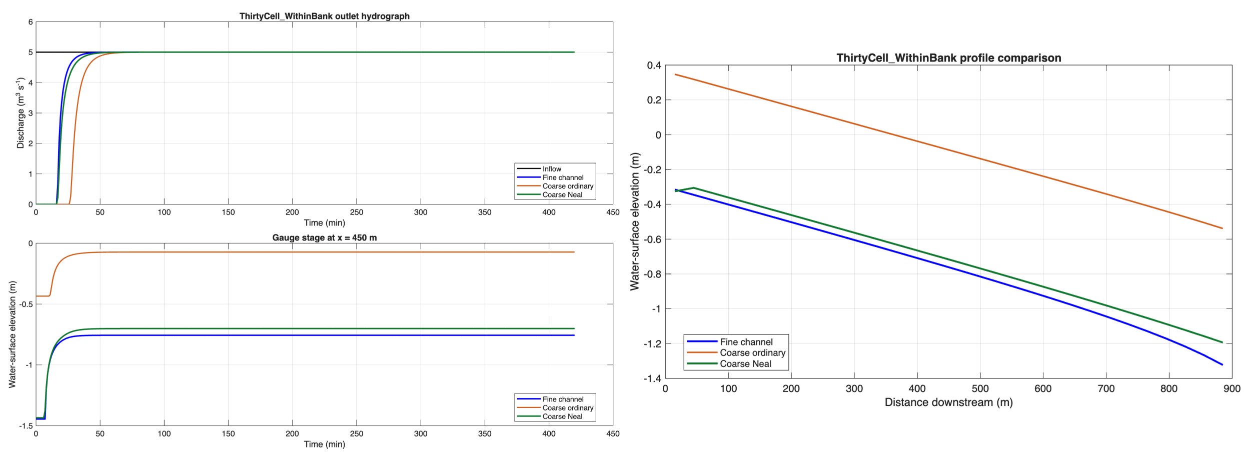

Figure 18: Simple in-bank channel benchmark. The Neal (2012) coarse grid reproduces the fine-grid hydrograph, internal stage, and longitudinal profile more closely than the ordinary coarse grid.

13.4 Composite Channel Under Constant and Transient Inflow

The composite case uses a 990 m reach with a 10 m-wide, 1 m-deep channel, 90 m lateral benches, slope 0.005, and Manning $n = 0.035$. The explicit reference uses 10 m cells, and both coarse simulations use 30 m cells. Constant $100\,\mathrm{m^3\,s^{-1}}$ inflow and a Nash hydrograph peaking at $100\,\mathrm{m^3\,s^{-1}}$ at 30 min were tested.

Table 13: Composite-channel metrics relative to the 10 m explicit numerical reference.

Forcing	Coarse model	Outlet RMSE ($\mathrm{m^3\,s^{-1}}$)	Outlet NSE	Internal RMSE ($\mathrm{m^3\,s^{-1}}$)	Storage RMSE ($\mathrm{m^3}$)
Constant	Ordinary	6.178	0.954	4.819	2439
Constant	Neal (2012)	4.189	0.979	3.346	734
Nash	Ordinary	7.263	0.949	5.770	3363
Nash	Neal (2012)	4.044	0.984	2.987	781

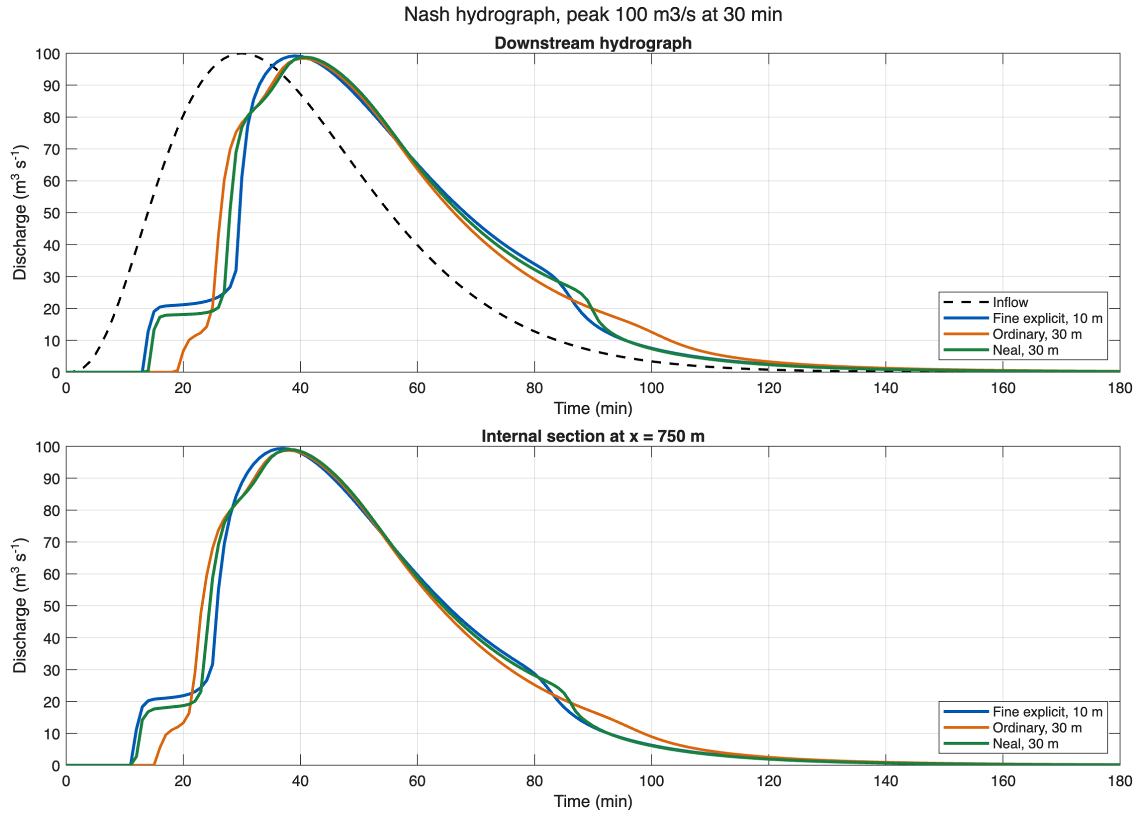

Figure 19: Transient composite-channel comparison. The 30 m Neal (2012) simulation follows the 10 m explicit outlet and internal hydrographs more closely than the ordinary 30 m simulation.

13.5 Staged Overbank Flow

Case `P1-SUBGRID-NEAL-STAGED-OVERBANK-001` evaluates channel-to-floodplain activation. The 1 km reach is 300 m wide with slope 0.005. Its cross section has a 50 m channel, two 25 m benches at 0.5 m above the channel bed, and an external floodplain at 1.0 m. Channel Manning n is 0.035 and floodplain Manning n is 0.10. The explicit reference uses 5 m cells; the ordinary and Neal (2012) grids use 100 m cells. The Nash inflow peaks at 200 $m^3\,s^{-1}$ at 45 min. Analytical uniform-flow calculations give activation discharges of 31.4 $m^3\,s^{-1}$ for the benches and 110.8 $m^3\,s^{-1}$ for the external floodplain.

Table 14: Staged-overbank metrics relative to the 5 m explicit numerical reference.

Metric	Ordinary 100 m	Neal (2012) 100 m
Outlet RMSE ($m^3\,s^{-1}$)	15.476	2.723
Outlet NSE	0.945	0.998
Internal RMSE ($m^3\,s^{-1}$)	12.053	3.043
Bench-depth RMSE (m)	0.180	0.018
Maximum-depth RMSE (m)	0.280	0.040
Wet-area CSI at 0.01 m	0.334	0.998

The Neal (2012) simulation activates the benches at 21 min, equal to the explicit reference, and

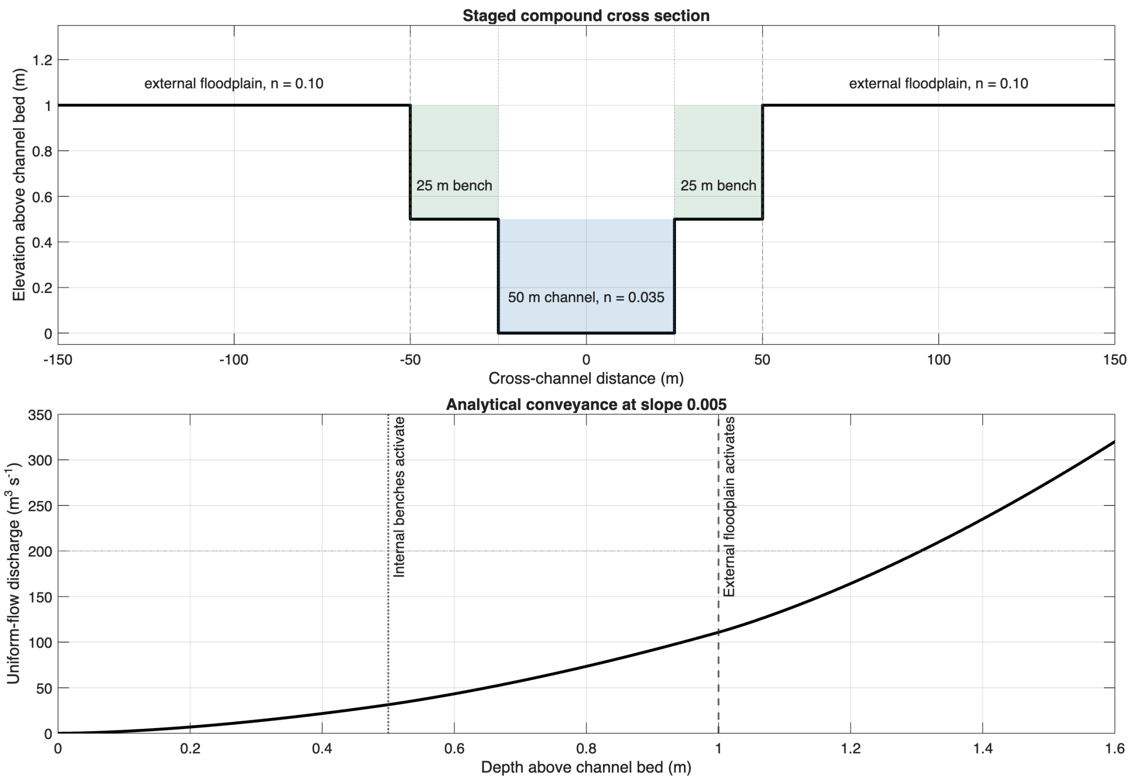

Figure 20: Staged compound-channel geometry and analytical uniform-flow rating curve. The inflow activates both the internal benches and the external floodplain.

the external floodplain at 35 min, compared with 34 min in the reference. Its outlet hydrograph, depth series, and maximum inundation extent are substantially closer to the reference than those from the ordinary 100 m grid. All retained Neal (2012) cases have mass residuals below $6 \times 10^{-13}\%$.

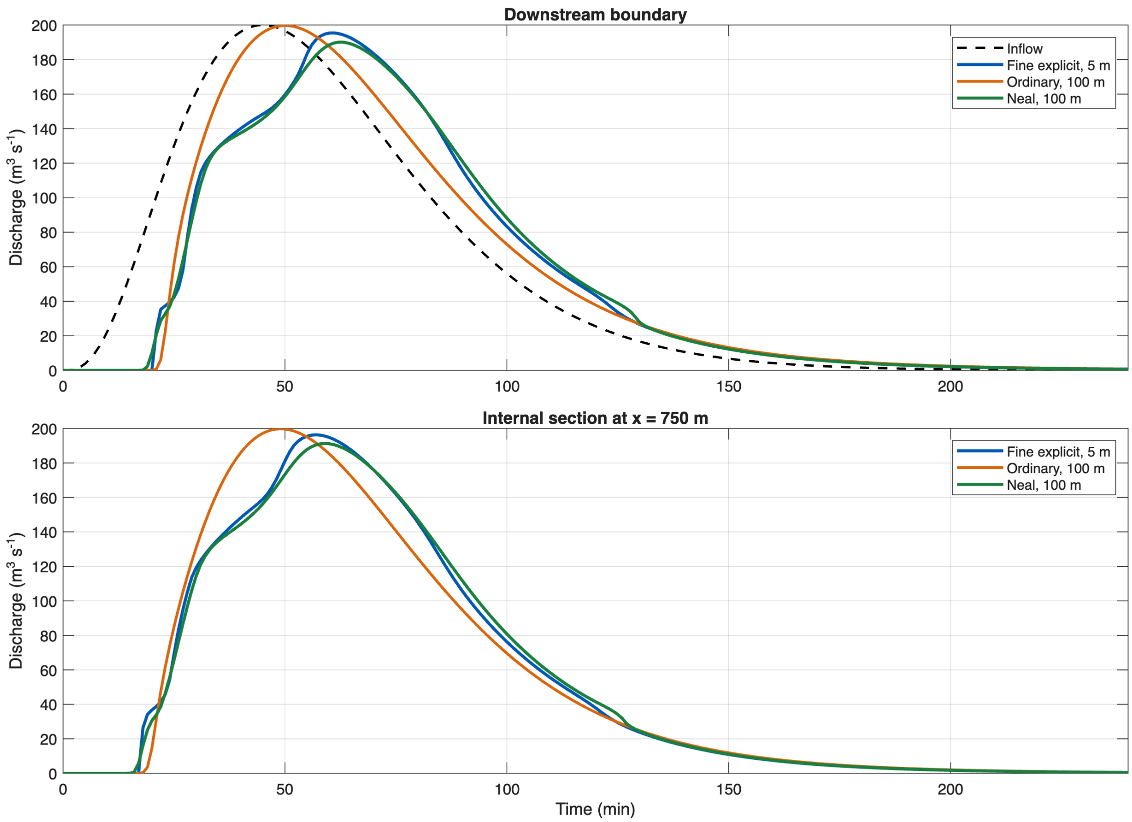

Figure 21: Staged-overbank outlet and internal hydrographs. The Neal (2012) result follows the fine-grid reference through activation and peak flow.

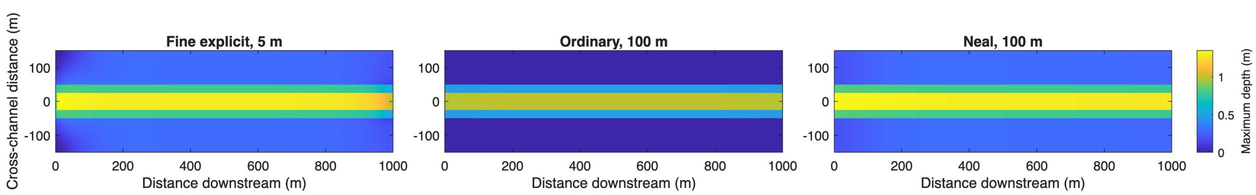

Figure 22: Maximum-depth maps for the staged-overbank case. The ordinary coarse grid underestimates floodplain wetting, whereas the Neal (2012) result approaches the fine-grid maximum extent.

The staged case also defines the present limitation. The Neal (2012) benches become dry at 126 min, compared with 196 min in the explicit reference. The external floodplain becomes dry at 152 min, compared with 237 min. One representative water level is stored in each coarse channel cell, so residual water on benches and floodplains cannot drain independently from the embedded channel. The controlled tests therefore support in-bank routing, floodplain activation, peak hydraulics, and maximum extent within the tested geometries. Applications controlled by late floodplain drainage require separate observational or fine-grid evaluation.

14 Reservoir and Control Structures

The reservoir validation isolates the control-structure rating-curve logic before it is coupled to a full catchment. The benchmark uses a constant-area reservoir with storage $S = Ah$ and a single outlet rating curve

$$Q = k \max(h - h_0, 0)^b, \tag{24}$$

where h is reservoir stage, h_0 is the sill stage, k is the rating coefficient, and b is the rating exponent. With no external inflow, the analytical reference is the exact depletion solution of

$$A\frac{dh}{dt} = -k(h - h_0)^b. \tag{25}$$

For $b \neq 1$, this gives

$$h(t) = h_0 + \left[(h(0) - h_0)^{1-b} + (b - 1)\frac{k}{A}t\right]^{1/(1-b)}. \tag{26}$$

The implemented test in `HydroPol2D_Model/Validation/Reservoir/RatingCurve_StorageDepletion/run_rese` uses $A = 20\,000\,\mathrm{m}^2$, $h(0) = 2.50\,\mathrm{m}$, $h_0 = 0.35\,\mathrm{m}$, $k = 2.2$, and $b = 1.5$. The explicit HydroPol2D rating-step update is compared against the exact stage, storage, and discharge series.

Table 15: Reservoir rating-curve validation metrics.

Case	Storage RMSE (m^3)	Outflow RMSE ($\mathrm{m}^3\,\mathrm{s}^{-1}$)	Relative L_2	Recovered k error (%)	Mass resi
P1-RES-001	0.2730	4.96×10^{-5}	7.80×10^{-6}	0.0000	4.3

The case passed all thresholds. The small storage difference is the expected explicit-integration truncation error relative to the continuous exact ODE, while the mass ledger itself closes to numerical precision because released volume plus remaining storage equals the initial storage.

15 Inflow Hydrograph Boundary Conditions

The inflow hydrograph boundary test imposes a known triangular hydrograph on the upstream boundary of a rectangular full-momentum domain. This is a boundary-condition validation rather than a dam-break validation: breach-style hydrographs, when needed, should be supplied through this same inflow-hydrograph mechanism. The truth source is the exact integral of the prescribed hydrograph. The routed downstream hydrograph is diagnostic because no analytical Saint-Venant solution is expected for the finite rough rectangular domain.

The implemented case in `HydroPol2D_Model/Validation/BoundaryConditions/InflowHydrograph_Rectangu` imposes a hydrograph that starts at 5 min, peaks at 25 min with $8\,\mathrm{m}^3\,\mathrm{s}^{-1}$, and returns to zero at 45 min. The exact inflow volume is $9600\,\mathrm{m}^3$. The full-momentum domain routes the wave to a downstream outlet while the diagnostic ledger checks

$$V_{\mathrm{in}} = V_{\mathrm{out}} + S_{\mathrm{domain}} + \epsilon. \tag{27}$$

The case passed the imposed-volume and water-balance criteria. At the end of the 3 h run, $9579.14\,\mathrm{m}^3$ had exited the outlet and $20.86\,\mathrm{m}^3$ remained stored in the domain, closing the ledger against the $9600\,\mathrm{m}^3$ imposed inflow volume.

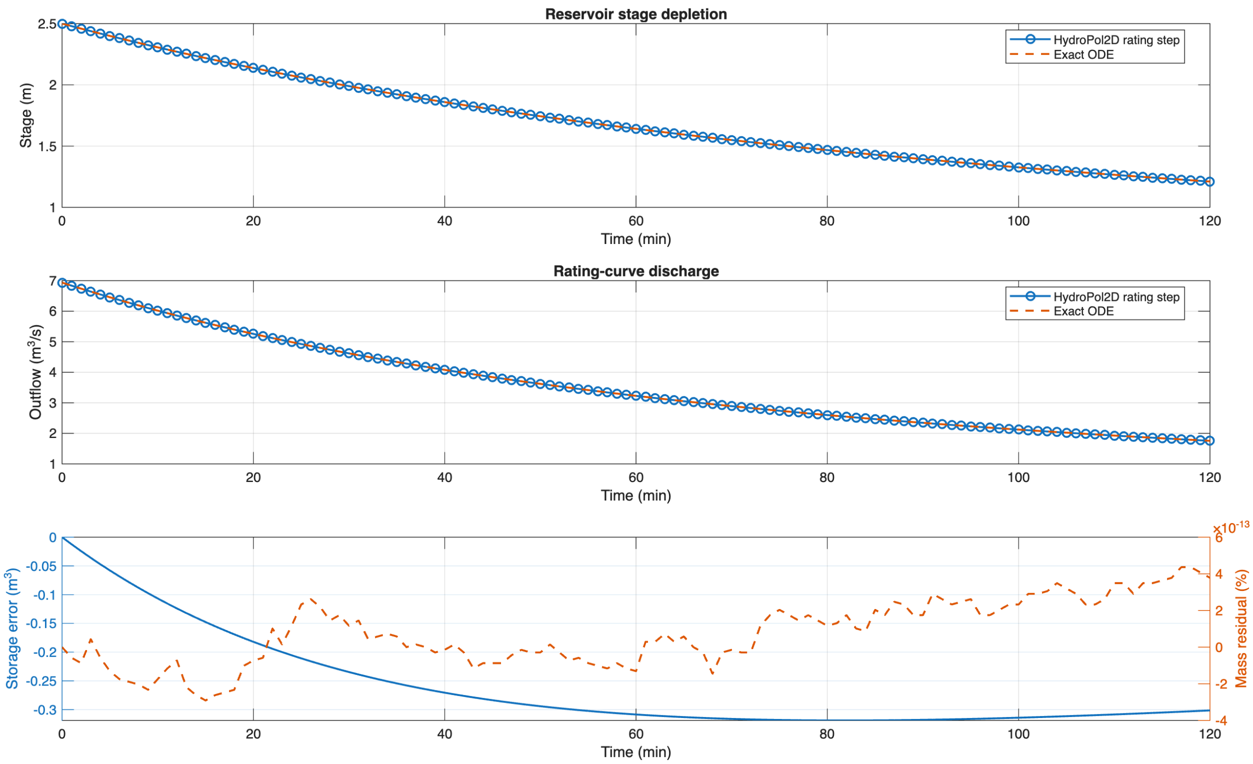

Figure 23: Reservoir rating-curve validation. The upper panel compares modeled and exact stage depletion, the middle panel compares rating-curve discharge, and the lower panel reports storage error and mass residual.

Table 16: Inflow-hydrograph boundary validation metrics.

Case	Imposed volume error (%)	Mass residual (%)	Outlet peak ($m^3\,s^{-1}$)	Peak lag (min)
P1-BC-INFLOW-001	3.41×10^{-13}	6.44×10^{-13}	7.3771	5.0

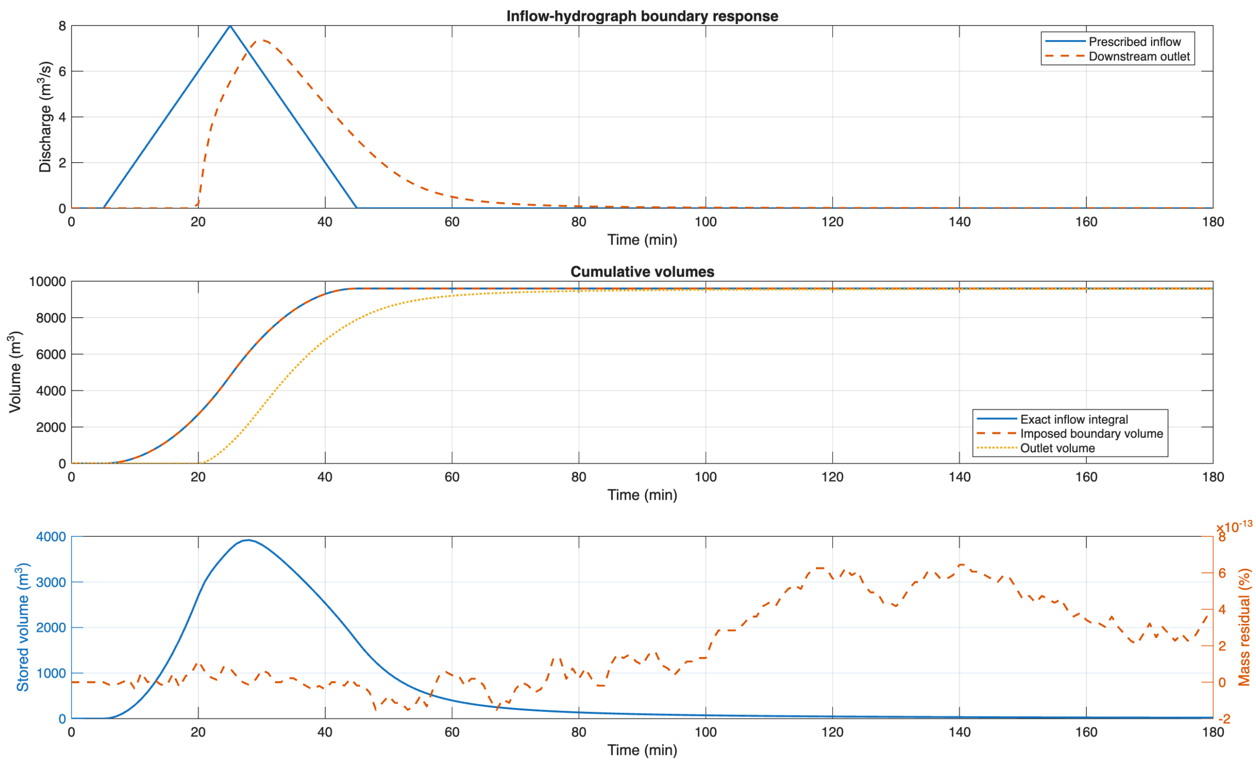

Figure 24: Inflow-hydrograph boundary validation under full-momentum routing. The upper panel shows the prescribed inflow and routed downstream hydrograph, the middle panel compares cumulative volumes, and the lower panel reports stored volume and mass residual.

16 Stage Hydrograph Boundary Conditions

The stage-hydrograph cases prescribe a sinusoidal upstream boundary depth in a rectangular domain. At every model step the upstream stage cells are reset to the prescribed value, the selected routing solver propagates the perturbation, and the diagnostic ledger tracks the volume inserted or removed by the stage boundary, outlet volume, and remaining domain storage. The prescribed boundary-stage enforcement error was zero to numerical precision for full momentum, local inertial, CA, kinematic, and diffusive routing. The downstream response amplitudes were 0.0718 m, 0.0723 m, 0.0170 m, 0.1514 m, and 0.0964 m, respectively, with all water-balance errors far below 0.1 %.

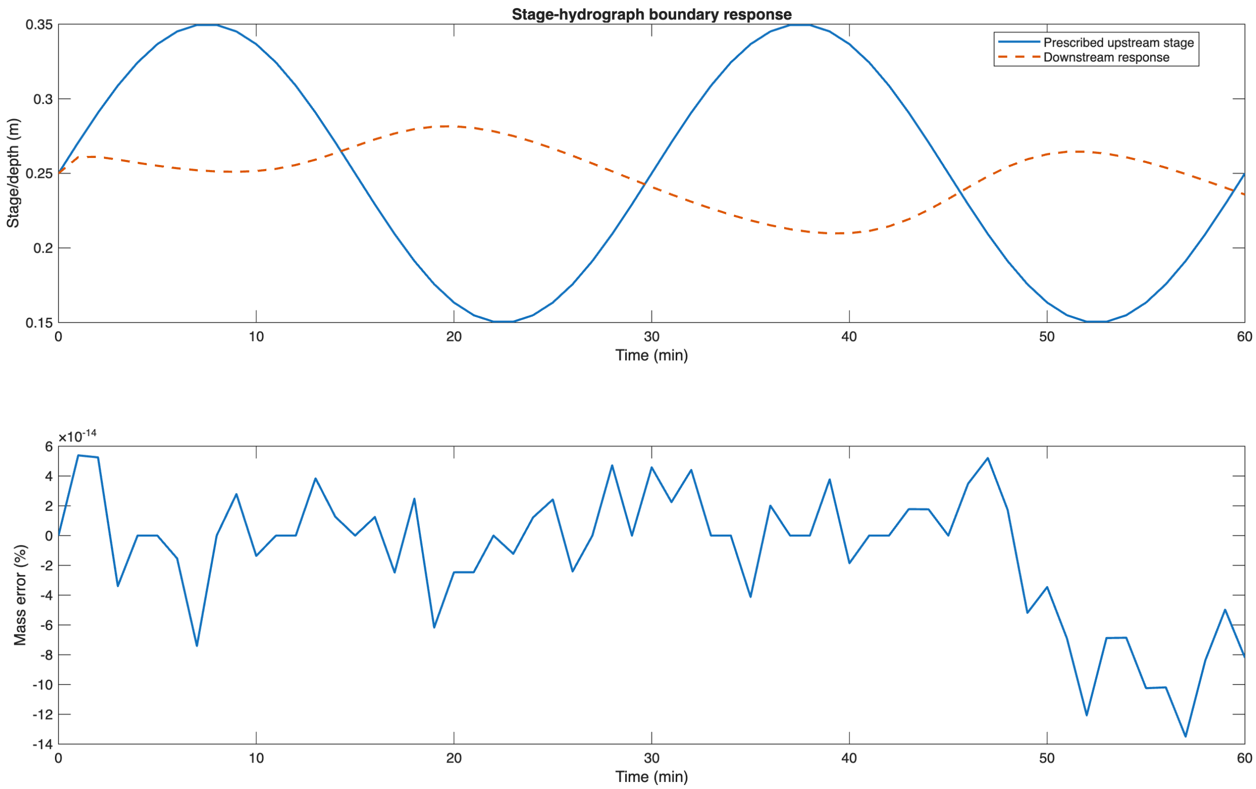

Figure 25: Prescribed stage-hydrograph boundary validation under full-momentum routing. The upper panel shows the imposed upstream stage and downstream response; the lower panel shows the mass-balance residual.

17 Spatial Rainfall and Satellite Maps

The Phase 1 spatial rainfall case validates raster ingestion, temporal accumulation, unit conversion, and cellwise volume accounting before these rainfall maps are coupled to routing, infiltration, or ET. Stanford remains an applied invariant-rainfall example; it is not used as the spatial-rainfall truth case.

The implemented benchmark in `HydroPol2D_Model/Validation/SpatialRainfall/TinyRaster_TotalRainfa` writes three tiny GeoTIFF rainfall-intensity rasters on a 4 by 5 grid with 20 m cells. The rasters are prescribed in $\mathrm{mm\,h^{-1}}$ and assigned durations of 10 min, 20 min, and 15 min. The reference

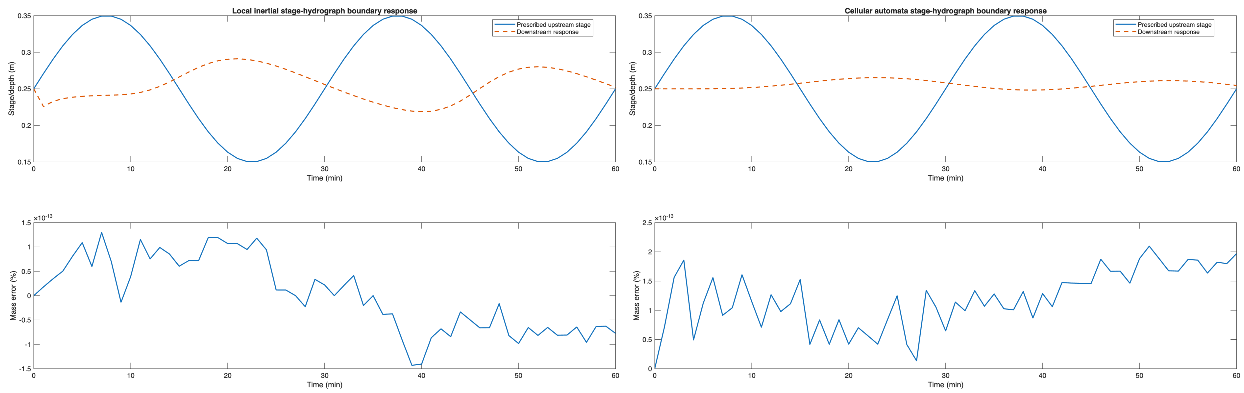

Figure 26: Prescribed stage-hydrograph boundary validation under local-inertial routing (left) and cellular-automata routing (right).

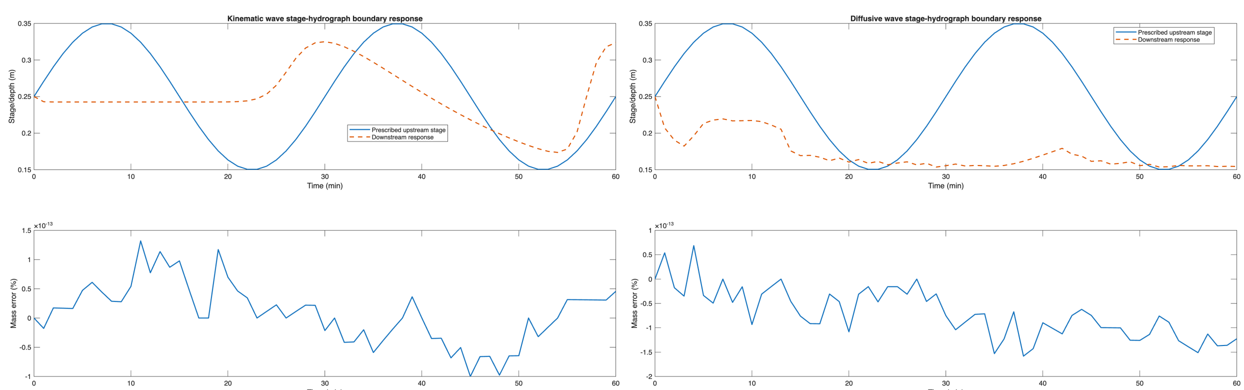

Figure 27: Prescribed stage-hydrograph boundary validation under corrected kinematic routing (left) and corrected diffusive routing (right).

accumulated depth in each cell is

$$P_{i,j} = \sum_k \frac{I_{i,j,k}}{1000} \frac{\Delta t_k}{60}, \tag{28}$$

and the domain volume is

$$V_P = \sum_{i,j} P_{i,j} \Delta x^2. \tag{29}$$

This case intentionally includes spatial gradients and zero-rainfall cells to expose row/column flips, domain averaging, missing no-data handling, and timestep-unit mistakes.

Table 17: Spatial rainfall raster validation metrics.

Case	Cellwise RMSE (m)	Max cell error (m)	Total volume (m^3)	Volume error (%)	Max
P1-RAIN-MAP-001	0	0	75.3333	0	

The case passed exactly: the raster-read accumulated depths matched the prescribed cellwise reference, each timestep volume matched the analytical integral, and the total rainfall volume residual was zero.

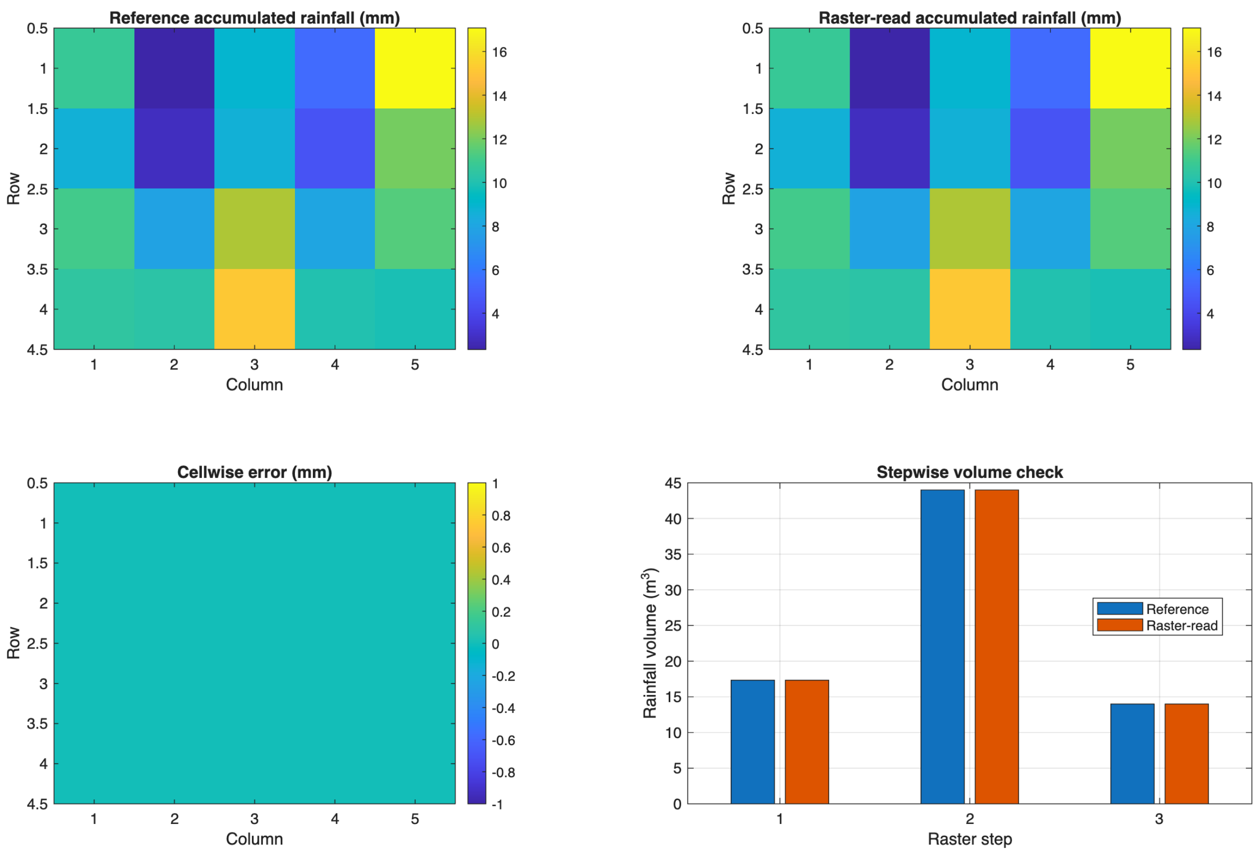

Figure 28: Spatial rainfall raster validation. The panels compare prescribed accumulated rainfall, raster-read accumulated rainfall, cellwise error, and stepwise rainfall volumes.

18 Water Quality

The Phase 1 water-quality case validates HydroPol2D's current mass-based buildup/washoff pathway before any field pollutograph calibration is attempted. The case uses physically plausible total suspended solids (TSS) parameter scales from the EPA SWMM buildup/washoff framework [6, 7], but the truth source is analytical mass balance rather than a SWMM numerical run.

The validation driver is `HydroPol2D_Model/Validation/WaterQuality/BuildUpWashOff_MassBalance/run_w` It calls HydroPol2D's `build_up_wash_off.m` directly in `flag_wq_model = 0` mass-based washoff mode. Three deterministic checks are included: dry-weather buildup initialization, single-cell exponential washoff, and conservative two-cell transfer/export. The TSS-like setup uses a $400\,\mathrm{m}^2$ cell, $30\,\mathrm{mm\,h}^{-1}$ runoff, a maximum buildup scale of $56.04\,\mathrm{kg\,ha}^{-1}$ (about 50 lb/acre), 5 day antecedent dry weather, and a washoff coefficient selected to give an initial concentration of about $120\,\mathrm{mg\,L}^{-1}$, within a plausible urban first-flush range.

Dry-weather buildup follows the SWMM-style exponential form

$$B_0 = C_1\,(1 - \exp(-C_2 ADD))\,\frac{A}{10000}, \tag{30}$$

where C_1 is the maximum buildup in $\mathrm{kg\,ha}^{-1}$, C_2 is the buildup rate in d^{-1}, ADD is antecedent dry days, and A is cell area in m^2. For the HydroPol2D mass-based washoff mode, the removed pollutant load is

$$W = C_3 Q^{C_4} B, \tag{31}$$

where B is remaining pollutant mass, Q is runoff discharge, and $C_4 = 1$ in this validation. Under constant runoff, the analytical single-cell solution is

$$B(t) = B_0 \exp(-\lambda t), \qquad \lambda = C_3 Q^{C_4}, \tag{32}$$

with t in hours for this driver. The two-cell transfer case uses identical first-order washoff in both cells. With equal initial masses B_0 in the upstream and downstream cells, the analytical cell masses are

$$B_1(t) = B_0 \exp(-\lambda t), \qquad B_2(t) = \exp(-\lambda t)\,(B_0 + \lambda B_0 t)\,, \tag{33}$$

and exported mass is obtained by conservation.

The validation exposed an implementation artifact: `build_up_wash_off.m` and `build_up_wash_off_8D.m` were rounding remaining pollutant mass to six decimal places at every substep. That quantization accumulated into an artificial mass error in high-resolution analytical tests. The routines now preserve full precision and only clean near-zero roundoff values. Table 18 reports the passing current-code metrics after that correction.

Table 18: Water-quality buildup/washoff validation metrics.

Case	Mass RMSE (kg)	Export error (%)	Mass residual (kg)	Negative mass count
Dry-weather buildup	0	0	0	0
Single-cell washoff	4.998×10^{-7}	2.483×10^{-4}	0	0
Two-cell transfer	6.272×10^{-8}	3.722×10^{-5}	0	0

All three cases passed the Phase 1 acceptance thresholds: mass RMSE below 1×10^{-6} kg, exported-load error below 0.01 %, zero mass residual, no negative mass, and no unexplained `NaN` values. Figures 29 and 30 show the analytical comparisons and conservation ledgers for the single-cell and two-cell cases.

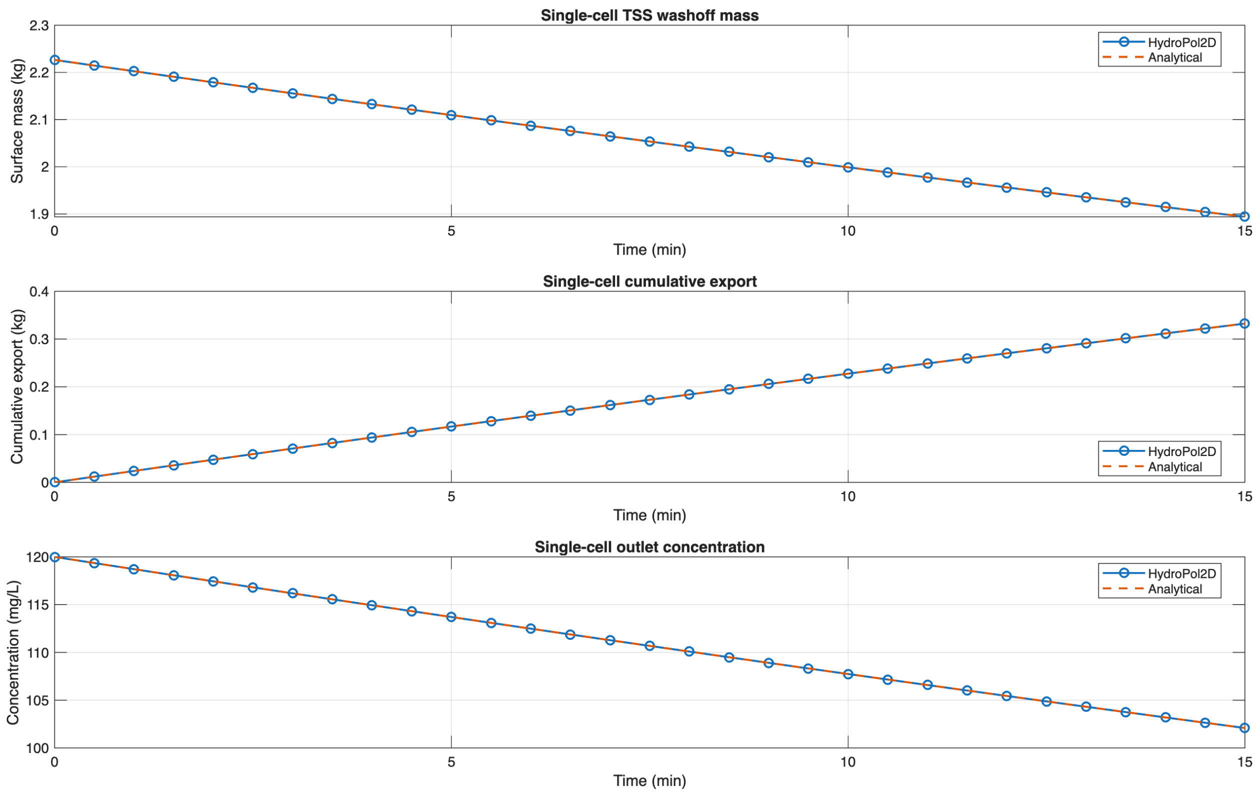

Figure 29: Single-cell water-quality validation. HydroPol2D's mass-based washoff is compared with the exact exponential remaining-mass, load, concentration, and cumulative-export solution under constant runoff.

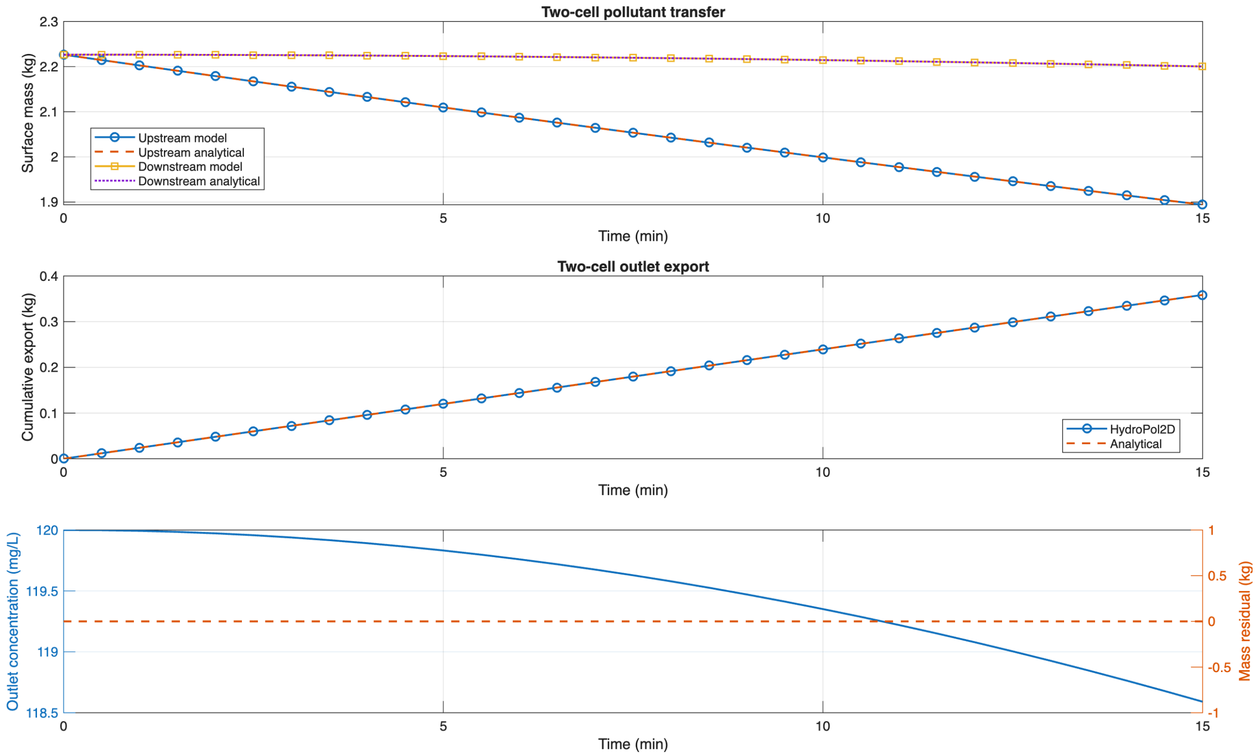

Figure 30: Two-cell water-quality validation. Upstream washoff enters the downstream cell, downstream washoff exits the domain, and remaining plus exported pollutant mass closes exactly.

19 Human Risk and Instability

The Phase 1 human-risk case validates the implemented algebra and classification logic, not field-scale public-safety prediction. The driver `HydroPol2D_Model/Validation/HumanRisk/DepthVelocity_Classific` tests two active HydroPol2D pathways. The first is the simple instability index used when `flag_human_instability = 1`. The second is the detailed force-balance classifier used when `flag_human_instability = 3`, which returns stable, sliding, toppling, or drowning classes.

For the simple instability index, the independent reference computes person weight, buoyancy, available friction, and hydrodynamic drag:

$$F_p = mg, \qquad F_b = w_1 w_2 h \rho_w g, \tag{34}$$

$$F_f = \mu \max(F_p - F_b, 0), \qquad F_d = \frac{1}{2} C_d \rho_w w_1 h v^2. \tag{35}$$

The reported risk index is

$$R = \min\left(\frac{F_d}{F_f}, 1\right), \tag{36}$$

with a deterministic edge convention: dry cells have $R = 0$, while wet cells with no available friction have $R = 1$. This convention avoids the physically ambiguous `0/0` case that previously produced `NaN` for still water when buoyancy eliminated friction.

For the detailed classifier, an independent implementation of the same slide, topple, and drowning force-balance equations was evaluated over a depth/velocity grid. The grid includes exact geometry-split and drowning-threshold cases. The geometry split is deterministic at equality by assigning the exact split depth to the lower-depth branch. The drowning threshold uses the implemented open condition $h > 13y/16$, so the exact threshold is not yet classified as drowning, while any value above it is.

Table 19: Human-risk and instability validation metrics.

Case	RMSE	Max error	Accuracy (%)	Boundary errors	NaN count
Simple instability index	0	0	100	0	0
Slide/topple/drowning classifier	0	0	100	0	0

The detailed classifier grid exercised all implemented output classes: 21 stable cells, 19 sliding cells, 2 toppling cells, and 12 drowning cells. All model classes matched the independent reference. Figures 31 and 32 show the model outputs and zero-error maps.

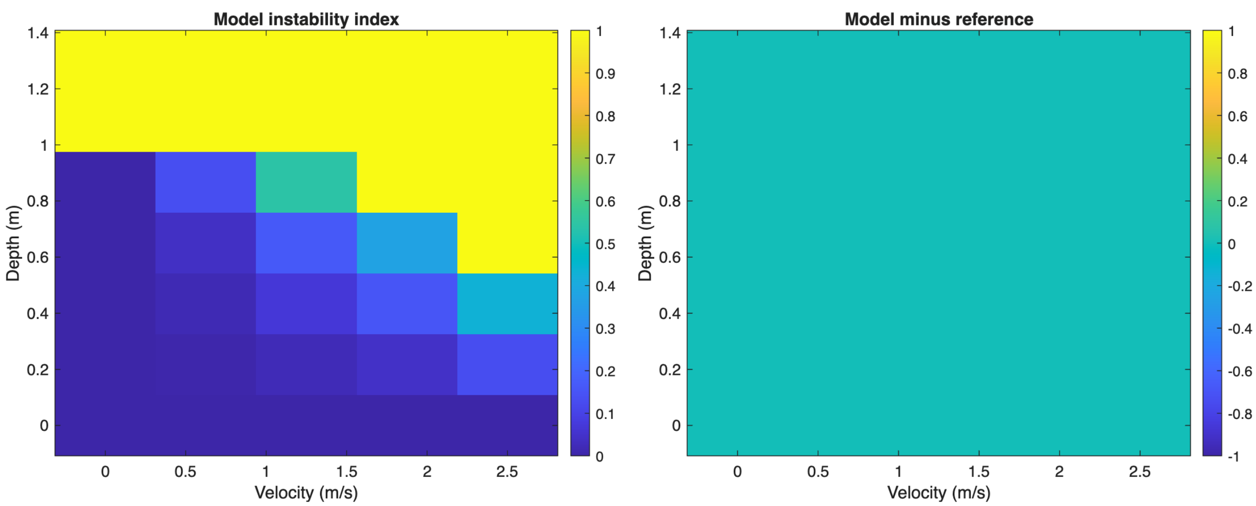

Figure 31: Simple human-instability validation. The left panel shows the HydroPol2D instability index over the deterministic depth/velocity grid; the right panel shows model-reference error.

20 Discussion

The Phase 1 structure deliberately separates isolated formula checks from coupled-domain dynamics. This prevents a correct final hydrograph from masking compensating errors among rainfall input, interception, infiltration, ET, groundwater exchange, routing, and boundary handling. At this stage, report-ready claims are limited to modules that have current-code outputs, documented analytical or independent-reference targets, exact bookkeeping checks, and standardized diagnostics.

21 Limitations

This Phase 1 report contains completed module results for canopy interception, snow, infiltration, evapotranspiration, groundwater and recharge, hydrodynamics, routing alternatives, Neal (2012) channel-subgrid hydraulics, reservoirs, boundary conditions, spatial rainfall, water quality, and human risk. Local-inertial routing did not pass the Ritter dam-break analytical threshold and should

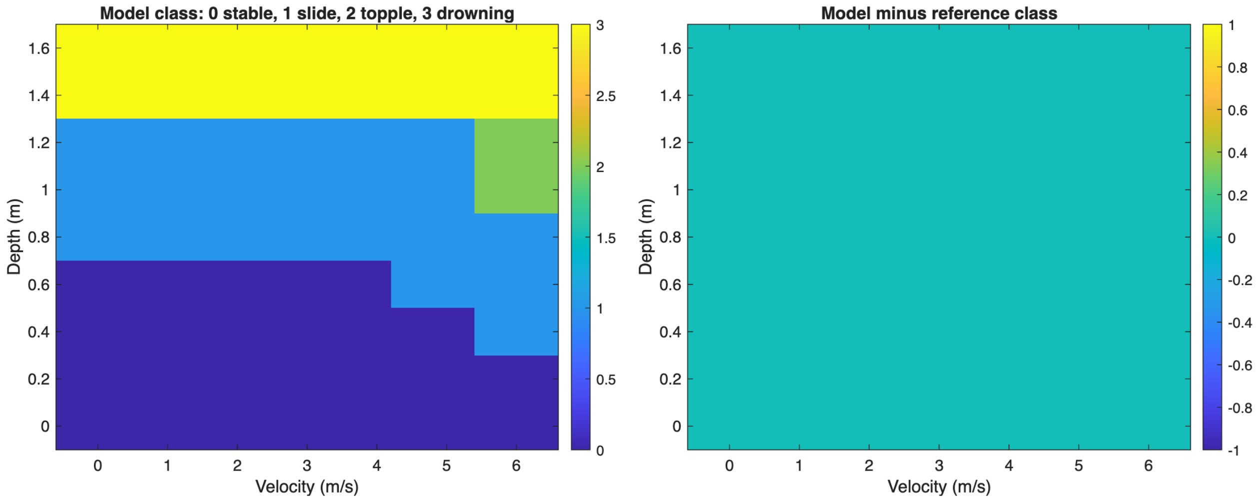

Figure 32: Detailed human-risk classifier validation. The left panel shows stable, sliding, toppling, and drowning classes; the right panel shows zero classification error against the independent force-balance reference.

not be cited as validated for dam-break dynamics. Kinematic and diffusive routing passed their reduced-equation plane, conservation, stage-bookkeeping, and V-tilted tests, but their Ritter cases remain diagnostic. The Neal (2012) channel-subgrid results use fine-grid local-inertial simulations as numerical references. They support the tested in-bank, channel-to-floodplain activation, peak-flow, and maximum-inundation behavior. They do not establish field-scale accuracy. The staged case also shows that one water level per coarse cell can make benches and floodplains drain too quickly during late recession.

22 Final Summary

Phase 1 validation proceeds from canopy interception through snow, infiltration, evapotranspiration, groundwater, hydrodynamics, reservoirs, boundary conditions, spatial rainfall, water quality, and risk classification. The completed cases pass their analytical, independent-reference, numerical-reference, or exact-bookkeeping criteria within the stated scope. The local-inertial and diffusive Ritter comparisons remain failed diagnostics, and the kinematic Ritter comparison remains diagnostic rather than a dynamic-wave validation claim. The Neal (2012) channel-subgrid formulation improves hydrograph, stage, storage, and inundation metrics relative to the ordinary coarse grid in all retained controlled cases. In the staged-overbank test, outlet NSE is 0.998, maximum-depth RMSE is 0.040 m, and wet-area CSI is 0.998. The main remaining limitation is early late-stage drainage from coarse floodplain cells. Field calibration and observational validation remain outside Phase 1.

References

[1] Richard G. Allen, Luis S. Pereira, Dirk Raes, and Martin Smith. *Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements*. FAO Irrigation and Drainage Paper 56. FAO, 1998.

[2] Robert F. Carsel and Russell S. Parrish. Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, 24(5):755–769, 1988.

[3] Y. Mualem. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, 12(3):513–522, 1976.

[4] Jeffrey Neal, Guy Schumann, and Paul Bates. A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resources Research*, 48(11):W11506, 2012.

[5] A. Ritter. Die fortpflanzung der wasserwellen. *Zeitschrift des Vereines Deutscher Ingenieure*, 36:947–954, 1892.

[6] Lewis A. Rossman. *Storm Water Management Model User's Manual Version 5.1.* U.S. Environmental Protection Agency, 2015.

[7] Lewis A. Rossman and Wayne C. Huber. *Storm Water Management Model Reference Manual Volume III: Water Quality.* U.S. Environmental Protection Agency, 2016.

[8] D. E. Rupp and R. A. Woods. On the use of the boussinesq equation for interpreting recession hydrographs from sloping aquifers. *Water Resources Research*, 42(12):W12421, 2006.

[9] P. A. Troch, C. Paniconi, and E. E. van Loon. Hillslope-storage boussinesq model for subsurface flow and variable source areas along complex hillslopes: 1. formulation and characteristic response. *Water Resources Research*, 39(11):1316, 2003.

[10] M. Th. van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44(5):892–898, 1980.

[11] N. E. C. Verhoest and P. A. Troch. Some analytical solutions of the linearized boussinesq equation with recharge for a sloping aquifer. *Water Resources Research*, 36(3):793–800, 2000.