water

MDPI

Article

Numerical Differences and Computational Performance Across Floating-Point Precision Configurations in a GPU-Accelerated 1D–2D Urban Flood Model

Sang-Bo Sim and Hyung-Jun Kim *

Department of Hydro Science and Engineering Research, Korea Institute of Civil Engineering and Building Technology, Goyang-si 10223, Gyeonggi-do, Republic of Korea; sangbo@kict.re.kr
* Correspondence: john0705@kict.re.kr; Tel.: +82-32-510-0276

Abstract

Coupled sewer–surface flood simulations are computationally demanding. We evaluated twelve surface-flow and exchange-discharge precision combinations in a graphics processing unit (GPU)-accelerated urban flood model for Jinju, Korea, using rain-on-grid (RoG) and rain-on-subcatchment (RoS) inputs. In 6 h hydraulic tests with 64-bit double-precision (FP64) exchange calculations, reduced-precision surface settings gave final water-depth root mean square errors (RMSEs) of 0.172–1.187 mm for RoG and 0.746–2.867 mm for RoS relative to the full-FP64 reference. Using 32-bit single precision (FP32) in selected exchange operations gave a maximum within-setting RMSE of 11.347 mm relative to FP64 exchange in the 48 h test. In a separate performance experiment, ten repeated 1 h runs per condition kept coupling data in GPU memory. Using FP32 for state storage and all targeted surface-flow operations, with FP64 exchange, achieved median speed-ups relative to the full-FP64 configuration of 1.883 (95% interval from repeated sampling: 1.876–1.887) for RoG and 1.205 (1.199–1.212) for RoS. The results identified this combination as an efficient candidate for the tested catchment. Local water-depth and exchange-volume differences should be assessed when applying it.

Keywords: urban flooding; coupled 1D–2D model; mixed precision; GPU acceleration; flow exchange; water balance

1. Introduction

Urban flooding arises from the time-dependent interactions between surface runoff and inundation, inflow into the storm-sewer network, exceedance of the sewer conveyance capacity, manhole surcharge, and return flow to the surface. Coupled models that dynamically link a one-dimensional (1D) sewer network with two-dimensional (2D) surface flow have been developed to represent these bidirectional exchanges and the subsequent water propagation across the surface. Leandro et al. [1] and Seyoum et al. [2] demonstrated that different sewer–surface coupling approaches produce different inundation results. Fan et al. [3] developed a model coupling 1D sewer, 2D surface-flow, and rainfall–runoff components and validated it against laboratory tests, numerical test cases, and a real drainage catchment. Guo et al. [4] noted that high-resolution topography and detailed drainage infrastructure improved the spatial representation of hydraulic processes but also increased the computational cost. Therefore, reducing runtime while retaining detailed 1D–2D interactions is an important requirement for urban flood forecasting and multi-scenario analysis.

Check for updates

Academic Editor: Athanasios Loukas

Received: 9 September 2026
Revised: 7 October 2026
Accepted: 13 October 2026
Published: 15 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Research aimed at reducing this computational cost has followed two main approaches: increasing the throughput via parallel computing and reducing arithmetic costs by varying numerical precision across computational components. Parallel implementations distribute workload across multi-core central processing units (CPUs), graphics processing units (GPUs), and heterogeneous CPU–GPU systems. Brodtkorb et al. [5] implemented and verified a GPU shallow-water solver with wetting-and-drying treatment. Smith and Liang [6] and Vacondio et al. [7] evaluated the computational performance of CPU- and GPU-based finite-volume flood solvers and improved their wetting-and-drying algorithms. Noh et al. [8] combined shared-memory and message-passing parallelism for hyper-resolution 1D–2D urban flood modeling. Morales-Hernández et al. [9] developed TRITON for execution on CPUs and on single or multiple GPUs, and examined how communication and I/O affect multi-GPU scaling. Buttinger-Kreuzhuber et al. [10] bidirectionally coupled a GPU-accelerated surface-flow solver with a 1D sewer-network model, whereas Caviedes-Voullième et al. [11] assessed the performance and scalability of SERGHEI-SWE across several CPU and GPU systems. Kim et al. [12] analyzed multi-GPU scaling in a 2D urban flood model as a function of mesh partitioning and workload per GPU. More recently, Zhu et al. [13] developed a GPU-accelerated sewer–surface model and showed that sewer computation influences coupled-model performance. These studies establish the importance of evaluating acceleration across the coupled workflow. However, in a coupled 1D–2D model, the extent to which faster 2D surface-flow computation reduces the total wall-clock time also depends on sewer routing, flow exchange, temporal synchronization, and I/O.

Mixed precision reduces the arithmetic and memory costs by assigning 32-bit single precision (FP32) and 64-bit double precision (FP64) according to the numerical sensitivity of each computational component. In a hydrodynamic model, different precisions can be used to store conserved variables, such as water depth and unit-width discharge, evaluate their tendencies, and accumulate quantities over long simulations. Klöwer et al. [14] characterized the error propagation in reduced-precision shallow-water calculations and proposed mitigation strategies. Li et al. [15] performed the main numerical operations of a large-scale hydrodynamic simulation in FP32, while retaining FP64 for accumulation. Ackmann et al. [16] used selective precision in the linear solver of a semi-implicit global shallow-water model, retaining higher precision in sensitive operations. Lai et al. [17] compared Gill, Kahan, and quasi-double-precision compensation in idealized one- and two-dimensional shallow-water models, demonstrating that the accuracy and computational cost depend on the compensation method. These approaches provide a basis for examining solver precision and accumulation; bidirectional sewer–surface exchange adds head-dependent flow direction and capacity limits. Hydrostatic balance, wet–dry treatment, flow-direction decisions and accumulation can be sensitive to rounding, producing differences in depth and discharge among precision settings. The proportion of whole-run time spent in each operation determines the available runtime benefit.

The unresolved question for a coupled urban flood model is how precision choices in surface-flow updates and sewer–surface exchange interact at the model level. Existing studies address parallel coupling, sensitive solver operations, or compensated accumulation, whereas exchange adds head comparisons, flow-direction selection, and capacity limits. Li et al. [18] examined synchronization between CPU-based sewer and GPU-based surface models operating at different timescales. Fappiano et al. [19] investigated coupled-model reproducibility when sewer and surface input data are inconsistent. Building on this work, the present study varies surface-flow and exchange-discharge precision separately and assesses depth, inundation extent, manhole exchange, surface-domain mass closure, and total runtime under common inputs. Depth, inundation, exchange and closure pro-

vide complementary numerical checks, while whole-run timing captures the saving after coupling overhead.

We investigated these issues by using the KICT Urban Flood model (KICT-UF). KICT-UF was developed as a dual-drainage model that dynamically couples a 1D sewer network with an unstructured 2D surface flow [20]; the present implementation follows the 1D–2D coupling and synchronization procedure described in Sim and Kim [21]. In a previous study, KICT-UF was evaluated against laboratory tests and a real catchment [22]. Two rainfall-input methods were used: rain-on-grid (RoG), in which rainfall was applied directly to each 2D element, and rain-on-subcatchment (RoS), in which subcatchment runoff was routed first to the 1D sewer network. As these methods produce different results and spatial distributions for wet 2D elements, as well as different temporal patterns for the sewer–surface exchange, they provide contrasting hydraulic and computational workloads for evaluating precision configurations.

The present study quantified how precision in the surface-flow solver and 1D–2D exchange calculations affects hydraulic outputs and computational performance. Hydraulic experiments evaluated twelve surface/exchange combinations through depth, inundation extent, manhole exchange and surface-budget closure under RoG and RoS. A separate one-hour performance experiment measured component and whole-run timing across the precision configurations under common execution settings. The analyses describe the numerical behavior and computational cost of precision choices for the tested catchment and rainfall inputs.

2. Materials and Methods

2.1. KICT-UF

KICT-UF simulates urban flooding by bidirectionally coupling the 2D surface flow generated by rainfall–runoff processes with a 1D sewer-network flow based on the U.S. Environmental Protection Agency Storm Water Management Model (EPA-SWMM) version 5.2 [23]. Its computational workflow comprises 2D surface-flow simulation, 1D sewer-network routing, 1D–2D flow exchange through manholes, and temporal synchronization between the two domains.

2.1.1. 2D Surface Flow on an Unstructured Mesh

The 2D surface-flow component of KICT-UF solves the conservative depth-averaged shallow-water equations (SWEs) on unstructured triangular elements:

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} + \frac{\partial G(U)}{\partial y} = S, \tag{1}$$

$$U = [h,\ hu,\ hv]^{\mathrm{T}};\ F = \left[hu,\ hu^2 + \frac{gh^2}{2},\ huv\right]^{\mathrm{T}};\ G = \left[hv,\ huv,\ hv^2 + \frac{gh^2}{2}\right]^{\mathrm{T}}, \tag{2}$$

where h is the water depth, u and v are the depth-averaged velocities in the x- and y-directions, respectively, and g is the gravitational acceleration. Conserved variables are defined as cell averages over each triangular element. The source term S represents rainfall, enabled surface losses and 1D–2D exchange in the continuity equation, together with bed slope and friction in the momentum equations. The boundary conditions are imposed through the exterior states and numerical fluxes at the domain boundaries and are not included in S. The numerical fluxes across the element interfaces are calculated using the Harten–Lax–van Leer–Contact (HLLC) approximate Riemann solver. On non-flat topography, numerical oscillations that may arise from a discrete imbalance between the numerical flux and bed-slope source term are reduced using the improved hydrostatic

reconstruction method of Hou et al. [24], and bed friction is calculated using Manning's roughness coefficient.

The surface solver uses a depth cutoff configured by TOLLIM = 10^{-6} m, a momentum dry threshold of max(TOLLIM, 10^{-3} m), and a separate Courant–Friedrichs–Lewy (CFL) depth switch at 10^{-4} m. These solver thresholds are distinct from the 1 mm depth mask used to evaluate numerical differences. The recorded CFL coefficient was 0.25. DTMIN is initialized to 1.0 s for minimum reductions and the dry-cell fallback. The outflow is limited by available water. Surface steps are clipped to the remaining coupling window. The output scheduler uses a half-step tolerance; Section 2.4.5 specifies the retained timestamp alignment. The precision assignments in Table 1 apply within these common numerical rules. All examined runs recorded direct initialization to zero surface depth. Exterior boundary opening was followed by solid-wall treatment of building obstructions and adjacent faces.

Table 1. Floating-point precision configurations for the 2D surface-flow solver.

Operation or Storage	P1	P2	P3	P4	P5	P6
Conserved-state storage	FP32	FP32	FP32	FP32	FP64	FP64
Water-depth calculation and hydrostatic reconstruction	FP32	FP64	FP64	FP64	FP64	FP64
Mesh geometry, CFL, and source terms	FP32	FP32	FP64 [a]	FP64 [a]	FP64	FP64
HLLC wave-speed calculation	FP32	FP32	FP32 [b]	FP32	FP32	FP64
Remaining numerical-flux calculations	FP32	FP32	FP32 [b]	FP64	FP64	FP64
Friction term	FP32	FP32	FP32	FP64	FP64	FP64

Note: [a] Momentum/depth quotients are evaluated in FP32. [b] HLLC fluxes use FP64 on protected faces. Implementation details are given in Section 2.2.

The preserved launch configurations specified a uniform Manning's coefficient of 0.0125 s $m^{-1/3}$, free-outfall exterior boundaries and disabled 2D surface losses. Under these settings, the uniform coefficient overwrites the land-cover roughness array after input reading. The free-outfall path used a dry exterior state at the interior bed elevation and clipped inward boundary fluxes. SWMM subcatchment infiltration followed its Green-Ampt input setting. The launch configurations selected integrated inlet-head evaluation; the performance runs integrated the inlet relation directly over each surface substep.

2.1.2. 1D Sewer-Network Flow Based on EPA-SWMM 5.2

The 1D sewer-network component uses the source code of EPA-SWMM version 5.2 [23]. KICT-UF uses the EPA-SWMM 5.2 dynamic-wave routing procedure to calculate the pipe discharge and hydraulic head at the sewer nodes. The governing relations for dynamic-wave routing comprise the continuity and momentum equations.

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0, \tag{3}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial \left(\frac{Q^2}{A}\right)}{\partial x} + \frac{gA\partial H}{\partial x} + gAS_f = 0, \tag{4}$$

where A is the flow cross-sectional area, Q is the pipe discharge, H is the hydraulic head, and S_f is the friction slope, which is calculated using Manning's equation and accounts for the pipe geometry, length, slope, roughness, and local losses. Continuity at a sewer node includes inflow and outflow through connecting pipes, subcatchment runoff, and

1D–2D exchange discharge [25]. Equations (3) and (4) express the governing relations for dynamic-wave routing; the discretization and solution procedures follow EPA-SWMM 5.2. 1D sewer-network calculations were performed using FP64 for each precision configuration.

2.1.3. 1D–2D Flow Exchange and Coupling Interval

Exchange at a storm-sewer manhole is determined from the sewer head and the water state of its linked surface elements. Surface-to-sewer inflow uses the inlet capacity in Equation (5), subject to surface availability and conveyance and storage limits. In the opposite direction, surcharge calculated during sewer routing is distributed over the linked surface elements. The surface-to-sewer direction is positive. Equations (5)–(7) describe inlet capacity, allocated volume and its transfer to the sewer ledger in this direction.

Equation (5) gives the Q64 weir-form inlet relation. Each node was assigned a numerical area proxy of 1 m^2; the effective width was its equivalent circular perimeter. The explicit inlet-area parameter was zero, selecting the weir branch. Surface depth was evaluated over the linked region, and sewer head was referenced to the node rim. The driving head was limited by both the surface depth and the surface-to-sewer head difference. The discharge coefficient belongs to the inlet relation; a separate coefficient is used for the surcharge-rate fallback. The proxy represents an idealized coupling geometry shared by all nodes. The preserved parameter files specify 0.5 for both the inlet and surcharge-fallback coefficients.

$$\begin{gathered} b_{\mathrm{j}} = \sqrt{4\pi A_{\mathrm{proxy,j}}},\ A_{\mathrm{proxy,j}} = 1\ \mathrm{m}^2 \\ h_{\mathrm{w,j}}(\mathrm{t}) = \max\left[0,\ \min\left(h_{\mathrm{s,j}},\ h_{\mathrm{s,j}} - h_{\mathrm{m,j}}(\mathrm{t})\right)\right] \\ Q_{\mathrm{inlet,j}}(\mathrm{t}) = C_{\mathrm{d}}\ b_{\mathrm{j}} \sqrt{2g} h_{\mathrm{w,j}}(\mathrm{t})^{3/2} \end{gathered} \tag{5}$$

$$\begin{gathered} V_{\mathrm{inlet,j,k}} = \int_{t_{\mathrm{k}}}^{t_{\mathrm{k}}+\Delta t_{\mathrm{s,k}}} Q_{\mathrm{inlet,j}}(\mathrm{t})\mathrm{d}t \\ 0 \le V_{\mathrm{alloc,j,k}} \le \min\left(V_{\mathrm{inlet,j,k}},\ V_{\mathrm{avail,j,k}},\ V_{\mathrm{pipe,j,k}},\ V_{\mathrm{store,j,k}}\right), \end{gathered} \tag{6}$$

$$\Delta V_{\mathrm{1D,allocated,j,k}} = V_{\mathrm{alloc,j,k}};\ \textstyle\sum_{\mathrm{i}\in M_{\mathrm{j}}} \Delta V_{\mathrm{2D,exchange,i,k}} = -V_{\mathrm{alloc,j,k}}, \tag{7}$$

In Equations (5)–(7), j identifies the manhole, k the surface substep, and M_{j} the set of surface elements linked to manhole j. The duration of substep k is $\Delta t_{\mathrm{s,k}}$. For the performance runs, inlet capacity is integrated over that substep using sewer head interpolated between coupling endpoints. The allocation bounds summarize the inlet, surface-water, effective pipe and remaining node-storage capacities before numerical tolerances and rounding; distribution can reduce the allocated volume further. Pipe capacity accounts for competing runoff demand, and node-storage capacity uses the numerical area proxy and evaluated head. Equation (7) specifies equal-and-opposite exchange-volume allocations to the sewer ledger and the linked surface-element updates. The volume remains pending until the next sewer-routing interval; the transfer rate is rescaled when that interval changes. Unadmitted volume is returned in a subsequent surface call, and pending volume is included in terminal accounting. Q64 uses FP64 for selected exchange calculations; Q32 uses FP32 for selected operations. Sewer-network calculations and the application and accumulation of the capacity-limited exchange volume remain in FP64. Q32–Q64 comparisons quantify exchange-precision effects within each surface configuration.

Q32 applies to selected exchange operations. In the coupling path used for performance measurements, FP32 head/integration operations and rounding to FP32 representation coexist with FP64 available-volume reductions, capacity-limit functions with FP64 arguments, distribution and cumulative ledgers.

Coupling intervals of 10 and 60 s were selected with reference to the accuracy–runtime relationships reported in previous fixed-time synchronization studies [18,21]. The same conditional coupling interval rule is applied to each precision configuration. An interval of

10 s was used at the beginning of each simulation. Thereafter, the interval remained at 10 s, while manhole surcharge, surface inflow, or sewer inflow persisted over the preceding coupling interval. Otherwise, it was increased to 60 s.

KICT-UF was previously evaluated with laboratory runoff data and observed flood traces in the Sangpyeong district [22]. Laboratory hydrographs used constant rainfall intensities of 30, 50 and 80 mm h^{-1}. The present experiments retain this model framework and apply common inputs within each precision comparison.

2.2. Experimental Design and Floating-Point Precision Configurations

The simulations used FP64 and FP32. The six surface-flow configurations (P1–P6) assign arithmetic precision within a common set of governing equations, mesh and sewer network. They assigned different precisions to state storage, hydrostatic reconstruction, mesh geometry, CFL-controlled time stepping, source terms, HLLC fluxes and Manning's friction (Table 1). P1 combines FP32 state storage with FP32 in the selected surface-arithmetic paths. P2–P4 retain the conserved variables in FP32, but evaluate different groups of operations in FP64, allowing configuration-level numerical differences associated with alternative precision assignments to be compared. P5 is a sensitivity configuration that retains the surface state and all other targeted operations in FP64 while evaluating only the HLLC wave-speed estimates in FP32. P6 performs all targeted surface-flow operations in FP64 and serves as a surface-flow reference. Each surface configuration was combined with Q64, in which the exchange discharge is calculated in FP64, and Q32, in which selected exchange operations use FP32, yielding 12 combinations. All combinations in the 1 and 6 h tests were compared with those of P6-Q64. The effect of exchange precision was evaluated by directly comparing Q32 and Q64 within the same surface configuration.

P1–P4 store the surface state in FP32; the common depth update applies flux and source increments in FP64 temporaries before storage rounding. P3 uses FP64 HLLC on boundary faces and faces adjacent to a static-risk cell; other faces use the FP32 raw-flux path. The mask marks boundaries, coupling centers and cells satisfying a bed-elevation-jump/face-length criterion. Its source-default threshold is 0.10 and permits an environment override; the effective historical P3 threshold was not recorded.

P3/P4 generic CFL expressions and velocity limiting of FP32 states include momentum/depth quotients evaluated in FP32 before widening. P4 reconstructs water-surface elevation from bed elevation plus depth in FP64; P5 uses the stored elevation. Their comparison includes these reconstruction and intermediate-operation differences. The separate P4 storage intervention preserves its specified arithmetic choices (Section 2.4.4).

The precision configurations were informed by studies that separated the precision of state variables from that of accumulated quantities [14], examined the sensitivity of water depth and free-surface elevation to round-off [6,26], and distinguished between the precision of primary numerical operations and accumulation [15]. P2–P4 retain different groups of coupled surface-flow operations in FP64, as detailed in Table 1. P5 was designed to isolate the numerical differences produced by evaluating only the HLLC wave-speed estimates in FP32 relative to P6.

2.3. Study Area and Rainfall–Runoff Simulation Conditions

2.3.1. Study Area and Computational Mesh

The study area was the Sangpyeong district of Jinju, Gyeongsangnam-do, Republic of Korea, a low-lying urban catchment of approximately 2.42 km^2 adjacent to the Nam River. Industrial, residential and transportation land uses produce a high proportion of impervious cover, making sewer conveyance and internal drainage important during intense rainfall. The 2D domain contains 2,790,139 nodes and 5,572,301 unstructured

triangular elements resolving buildings and major topographic boundaries. Native mesh coordinates use EPSG:5179 (KGD2002/Unified CS), with meter units. The 1D network contains 542 manholes and 536 pipes; 528 storm-sewer manholes exchange water with linked surface elements. Figure 1 shows the study area, domain, manholes and mesh.

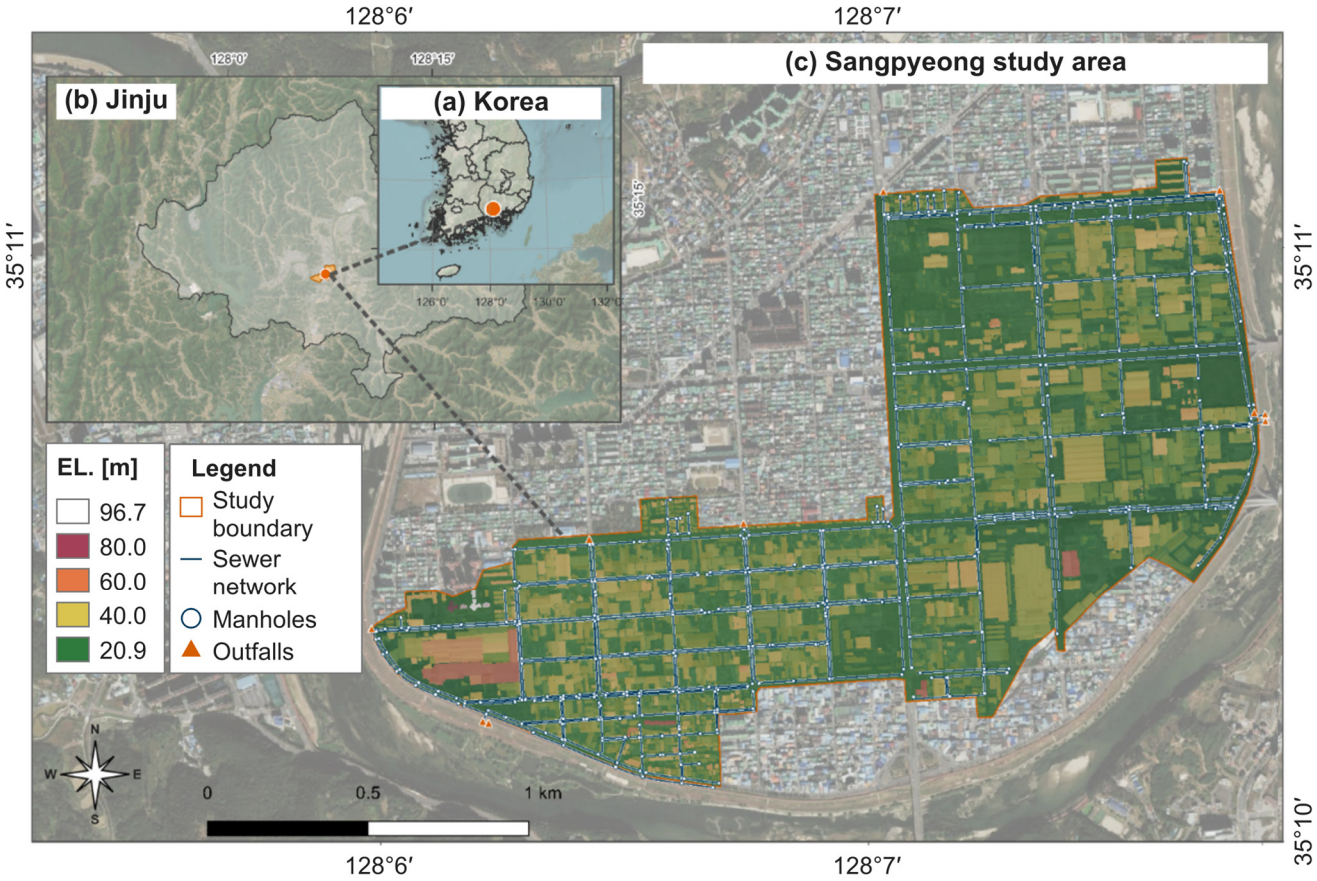


Figure 1. Study area and computational domain.

Figure 1 was prepared in QGIS using Esri World Imagery (https://services.arcgisonline.com/ArcGIS/rest/services/World_Imagery/MapServer; accessed on 29 September 2026). The imagery is credited to Esri, Vantor, Earthstar Geographics and the GIS User Community, © 2026 Esri and its licensors, and is used under Esri's terms for static maps.

2.3.2. Rainfall-Input Methods and Simulation Duration

Under RoG, rainfall is applied directly to the 2D numerical elements; therefore, a larger portion of the surface domain becomes wet, and the surface-flow workload increases. Under RoS, the subcatchment runoff enters the 1D sewer network. This generally produces fewer wet 2D elements, but can increase the relative contribution of sewer routing and 1D–2D exchange, including manhole surcharge and return flow to the surface. Therefore, RoG and RoS represent two contrasting rainfall-input methods and computational workloads. The numerical differences and computational performance were compared under both conditions.

The 1 h test used a design storm with a total depth of 80 mm, corresponding to the 1 h local flood-protection target for Jinju. A 1 min hyetograph was constructed using the Huff third-quartile temporal distribution, which is commonly adopted in Korean flood-design practices [27–29]. This input is hereafter referred to as the 1 h, 80 mm design storm. For the 6 h test, an external 277.5 mm hyetograph derived from an intense rainfall event that affected the Seoul metropolitan region on 8 August 2022, was applied to the

Sangpyeong catchment. For the 48 h test, a 60 min rain-free initial period was followed by a rainfall record from 01:00 on 20 September to 23:50 on 21 September 2024, with a total depth of 307.4 mm. The 1 h test examined short-duration numerical differences and repeated computational performance during the design storm. The 6 h test evaluated differences in water depth, inundation extent, exchange volume, and water balance under a high-intensity rainfall scenario, whereas the 48 h test examined the temporal evolution of water-depth differences during rainfall and subsequent drainage. Figure 2 shows the three rainfall hyetographs.

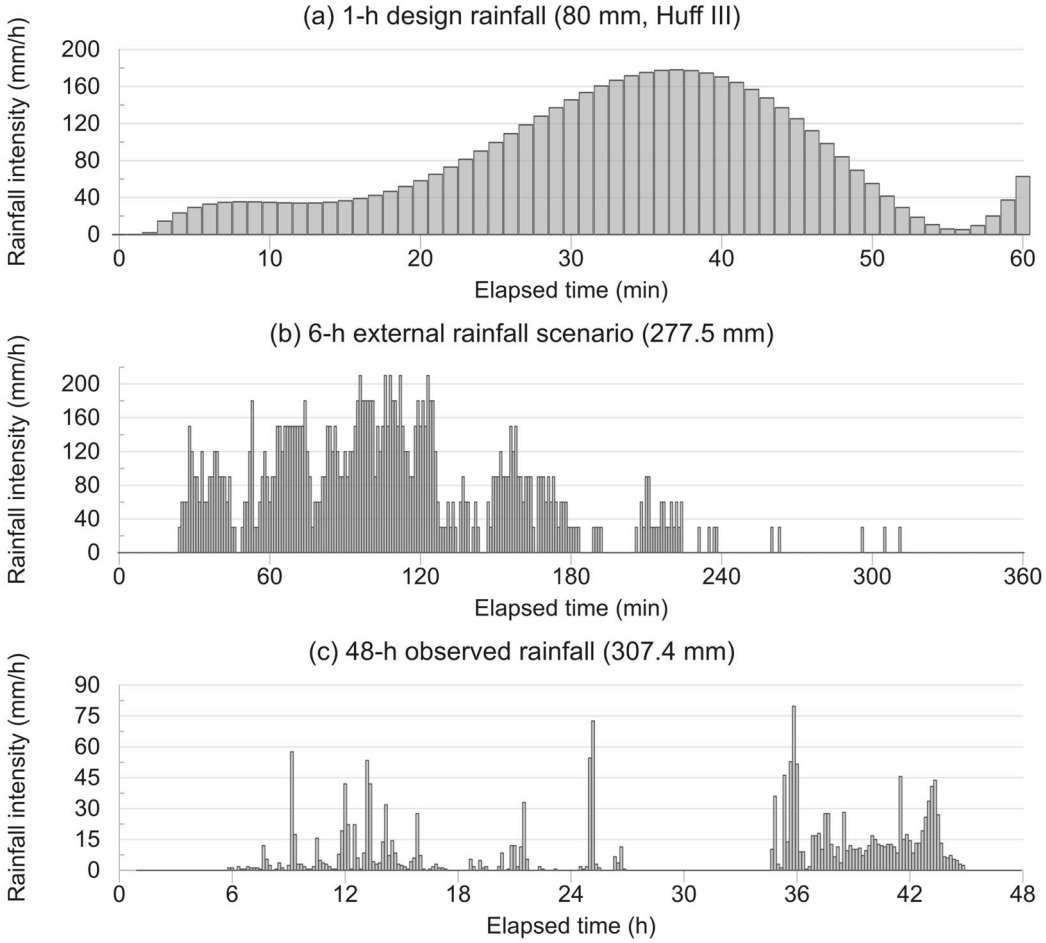


Figure 2. (**a**) 1 h, 80 mm design storm; (**b**) 6 h external rainfall scenario; (**c**) 48 h observed rainfall record.

The 48 h simulation included P6-Q64, P1-Q64, and P1-Q32. P1-Q64 was included to examine a reduced-precision surface configuration retaining FP64 exchange calculations, alongside the full-FP64 reference P6-Q64 and the FP32-exchange setting P1-Q32. P1-Q64 and P6-Q64 were compared to assess the combined influence of the surface-flow precision assignments; P1-Q32 and P1-Q64 were compared to assess the additional difference associated with exchange-discharge precision. Under RoS, subcatchment runoff enters the sewer network before surcharge supplies water to the surface, followed by return flow and drainage recession. This case therefore examines the evolution of numerical differences over an extended sewer–surface interaction period. Table 2 summarizes the simulation matrix.

Table 2. Simulation matrix and purpose of each rainfall case.

Test	Simulation Conditions	Precision Combinations	Primary Purpose
Numerical comparison under the 1 h design storm	80 mm, Huff third-quartile distribution; RoG and RoS	P1–P6 × Q64/Q32	Reference-relative numerical differences under the 1 h, 80 mm design storm
Repeated performance test under the 1 h design storm	80 mm, Huff third-quartile distribution; RoG and RoS; minimal-output configuration	P1–P6 × Q64/Q32	Ten repeated runs per condition; means, SDs, medians, IQRs and median runtime ratios with intervals estimated by repeated sampling.
Numerical comparison for the 6 h case	External rainfall input based on the 8 August 2022, Seoul metropolitan rainfall event (277.5 mm; 361 values at 1 min intervals); RoG and RoS	P1–P6 × Q64/Q32	Differences in water depth, cumulative net exchange volume, and 2D surface-domain water balance over the full rainfall scenario
Water-depth difference evolution in the 48 h RoS case	RoS rainfall record from 01:00 on 20 September to 23:50 on 21 September 2024 (307.4 mm)	P6-Q64, P1-Q64, and P1-Q32	Temporal water-depth difference evolution associated with the surface-flow and exchange-precision configurations
Additional Q64 spatial diagnostics	80 mm/60 min design storm; RoG and RoS; six depth snapshots	P1-Q64 and P6-Q64	Four snapshot runs and four paired controls; depth, inundation, exchange and budget comparisons between P1-Q64 and P6-Q64.

The 6 h analysis included all 12 P1–P6 × Q64/Q32 combinations for both RoG and RoS, yielding 24 simulations. Each simulation used an initially dry surface with zero surface-water storage, 528 storm-sewer manholes, identical input data, and common output times. The water depth, inundation extent, cumulative net exchange volume at each manhole, and 2D surface-domain water balance were compared at 36 output times from 10 to 360 min at 10 min intervals.

2.4. Evaluation Metrics and Reproducibility

2.4.1. Water Depth, Inundation Extent, Exchange Volume, and 2D Surface-Domain Water Balance

P6-Q64, in which all targeted operations were performed in FP64, was used as the reference simulation for comparison among precision configurations. The root mean square error (RMSE) of water depth relative to the reference was calculated as an element-area-weighted metric over the union of elements with a water depth of at least 1 mm in either simulation of the comparison pair. Inundation agreement was measured by dividing the area flooded in both simulations by the area flooded in either simulation, using a 10 mm depth threshold (intersection over union, IoU). IoU is reported to four decimal places. The difference in surface-water storage volume was also calculated. The 6 and 48 h water-depth fields were compared at 36 and 288 output times, respectively, both at 10 min intervals. For each storm-sewer manhole, the exchange volume was reported as cumulative directional volumes (sewer-to-surface and surface-to-sewer) and as the cumulative net exchange

volume with the sign retained by the flow direction. The weighted absolute percentage error (WAPE) in cumulative net exchange volume at each manhole was calculated using Equation (8). For comparison with P6-Q64, the denominator was taken as the sum of the absolute cumulative net exchange volumes for all manholes in P6-Q64. For Q32–Q64 comparison within the same surface configuration, the corresponding Q64 values were used in the denominator. RMSE, IoU and WAPE quantify reference-relative numerical differences among the tested precision configurations. The 95th and 99th percentiles in Section 3.2 used equal weights for elements with finite depths in the pairwise ≥1 mm wet union and linear interpolation between ordered absolute differences.

$$\mathrm{WAPE}\ (\%) = 100 \times \frac{\sum_j \left| V_{\mathrm{net,j,cand}} - V_{\mathrm{net,j,ref}} \right|}{\sum_j \left| V_{\mathrm{net,j,ref}} \right|}. \tag{8}$$

For the field-scale domain, the 2D surface-domain water-balance residual R_{2D} was calculated from the change in surface-water storage volume ΔV_s, rainfall P_s, sewer-to-surface and surface-to-sewer exchange volumes ($E_{1\to2}$ and $E_{2\to1}$, respectively), boundary inflow and outflow (B_{in} and B_{out}, respectively), surface losses L_s, dry-cell clipping volume C_{dry}, negative-depth correction volume C_{neg}, and the internal-face numerical-flux cancelation residual F_{int}. Accordingly, R_{2D} is defined as follows:

$$R_{2D} = \Delta V_s - [P_s + E_{1\to2} - E_{2\to1} + B_{in} - B_{out} - L_s - C_{dry} + C_{neg} + F_{int}]. \tag{9}$$

The total absolute bidirectional exchange volume is defined as $E_{abs} = |E_{1\to2}| + |E_{2\to1}|$, and the exchange-volume-normalized residual is expressed as a percentage: $R_{norm} = 100\,|R_{2D}|/E_{abs}$.

2.4.2. Computational Performance and Reproducibility

In the separate performance experiment, each precision combination was evaluated using ten measured runtimes for each rainfall-input method (n = 10). The 24 conditions provided 240 observations for the 60 min, 80 mm design storm. All performance conditions used the same approach to retain coupling data in GPU memory between updates; the initial surface-water depth was zero, and spatial output was disabled under the minimum-output settings. Within each condition, the records confirmed the same executable, source code, supporting files and input data. All runs used the same NVIDIA GeForce RTX 5090 (NVIDIA Corporation, Santa Clara, CA, USA), verified by GPU UUID, with 12 host threads. The implementation enabled accepted feedback and conflict-level parallel coupling with the adaptive 10/60 s schedule. No timing observation was excluded.

Linked records identify NVHPC Fortran 26.3.0, NVIDIA CUDA compiler 13.1.115, GNU C 15.2.0 and the project-modified EPA-SWMM 5.2.3 source lineage. Recorded Fortran options included -O2, -Kieee, -Mnofma, -Mnoflushz, -Mlre=noassoc, -Mvect=noassoc, -acc, -mp and -gpu=cc120,rdc,nofma,noflushz,nounroll. CUDA C++ used C++17 and compute_120/sm_120, with -O3 for the 1, 6 and 48 h hydraulic experiments and -O2 for the repeated performance and additional Q64 spatial-diagnostic runs; the performance builds defined HC_SURF_C1_FIELD_PARALLEL_LEDGER. Historical CPU, host OS, affinity, physical RAM and loaded-runtime identities were unrecorded. The physical settings in Section 2.1.1 are documented in preserved launch scripts. Complete process-environment records are available for the representative Q64 output and control runs; coverage of earlier runs is partial. The modified KICT-UF source code used in this study is not publicly available and will not be distributed.

The 1, 6 and 48 h hydraulic experiments used the original implementation, while all repeated performance conditions used the implementation described above. Separate P1-Q64/P6-Q64 output runs with the performance implementation supplied the spatial diagnostics in Section 2.4.5. Recorded GPU UUIDs matched the performance device for all 24 one-hour hydraulic runs and 10 of the 24 six-hour runs; UUIDs were not recorded for the remaining six-hour runs or the three 48 h RoS runs.

Wall time in COUPLING_TIME_ANAL.OUT was measured from entry to swmm_run, before initialization, through the timer record preceding swmm_close, after reporting. It includes initialization, simulation and reporting in that interval, but excludes final close and timer-file writing; it is neither the full-process duration nor the shell-wrapper duration. Separate warm-up or JIT overhead was not isolated. Tables A1 and A2 report mean, sample standard deviation (ddof = 1), median and interquartile range (IQR) for the ten observations per condition. Quartiles use linear interpolation. Speed-up is the reference median wall time divided by the candidate median wall time. The median runtime ratio was recalculated 20,000 times by drawing observations with replacement from the recorded timing data (seed 20260928). The 2.5th and 97.5th percentiles of these ratios formed the 95% interval. Each repeated sample retained the recorded execution-group sizes: randomized run blocks shared across configurations were drawn together, while unpaired runs were drawn independently within their execution group. The intervals are exploratory and unadjusted for multiple comparisons. An interval containing 1 is consistent with equal median runtimes for the two configurations.

The 240 timer records and executable/input metadata were checked. Within each condition, printed solver-step counts, final surface storage, directional exchange totals and full-diagnostic residuals were identical across repeats. Node source/sink totals agreed with corresponding mass-budget records within printed precision and FP64 summation bounds. These checks cover retained summaries. Component timers can overlap; Section 3.8.6 supplies a separate exclusive host partition and GPU/API measurements.

Separate component profiling used the same 60 min, 80 mm inputs for RoG/RoS P1-Q64 and P6-Q64. Each condition had three control runs without profiling instrumentation, one marker-only run and three Nsight Systems runs (28 runs in total), on the same GPU and with the same numerical and minimum-output settings. Host-clock and NVTX instrumentation retained the numerical statements and CUDA kernels and added no synchronization. A CLOCK_MONOTONIC nested interval stack subtracted child intervals to obtain an exclusive partition of the main host control thread. Model wall time extended from swmm_run entry to completion of swmm_close; writing the new timing CSV was outside this interval. Fractions were computed within each trace run and then averaged. These profiling observations were not pooled with the $n = 10$ performance samples.

Nsight activity records supplied GPU-kernel and explicit-copy durations. Attribution used CUDA correlation identifiers linked to NVTX ranges. Synchronization API intervals were combined by interval union on the host thread. GPU and host activities were presented separately because they can overlap. The phase counts matched the host-call counts despite a general NVTX-collection warning. Unified Memory migration was unavailable under the environment's trace restriction. All 28 runs had matching saved step counts, mass budgets, final storage and node summaries within each condition. Legacy-timer Nsight/control median ratios were 1.0153–1.0593; marker-only/control ratios were 0.9941–1.0197, with one marker-only observation per condition. All observations were retained.

2.4.3. Numerical Trajectories and Storage Rounding

Additional RoS diagnostics used P1-Q64, P4-Q64 and P5-Q64 over a 60 min prefix of the 6 h scenario. Eleven elements were selected from the 40 min P1/P4 difference field:

eight with large area-weighted squared differences and three comparison elements. A 31 min P1/P4 diagnostic monitored 14 elements associated with CFL divergence. These targeted records describe the selected numerical pathways. CFL values are candidates before coupling/termination clipping, and transition counts cover the selected cells.

The diagnostics recorded flux and state-update intermediates, step counts, coupling schedules and selected-cell transitions. P4 rounding volumes were calculated from actual stored depth minus the FP64 target, multiplied by element area and accumulated after flux and source updates. Diagnostic fields matched the original six-hour outputs at 10–50 min for the 60 min runs and at 10–30 min for the 31 min runs. Terminal budgets were retained as separate diagnostic records. Their runtimes were excluded from the performance sample.

2.4.4. Analysis of Hydraulic Differences and Runtime

Depth-conditioned metrics were computed on reference-depth masks $h \geq 0.10$ m, $h \geq 0.30$ m, and $0.10 \leq h < 0.30$ m. The reference was the corresponding Q64 run for within-configuration Q32–Q64 comparisons and P6-Q64 for comparisons among surface policies. RMSE, mean absolute error (MAE), and signed bias were area-weighted. Absolute-difference percentiles reported for these masks used the left inverse of the cumulative element-area distribution. Empty masks were recorded as NA. The thresholds describe depth bands for analysis, not universal safety limits. The postprocessing checked 1008 instantaneous depth fields from the 1 and 6 h experiments; instantaneous fields were not mixed with cellwise maximum-depth maps.

The original 24 six-hour accuracy runs were linked to their wall-clock logs, solver-step counts, and coupling schedules. These were single observations per configuration, including the accuracy-run output workload, not additional samples of the one-hour repeated performance experiment. Historical GPU UUIDs were recovered for 10 of the 24 runs and were not recorded for the remaining 14 in the supplied run metadata, so device identity could not be independently established from those records. Top-level timing shares were calculated as a percentage of wall-clock time: the surface module, sewer routing, and their unassigned remainder. Within the surface module, total coupling and the remaining surface work were distinguished; the nested coupling timer was not interpreted as pure exchange-discharge arithmetic.

For additional normalization, the absolute corrected 2D residual was divided by the accumulated surface-external inflow (direct surface rainfall plus sewer-to-surface transfer and boundary inflow) and expressed as a percentage. A direct-rainfall normalization was also calculated where surface rainfall was nonzero. Direct surface rainfall was zero under RoS, so that denominator was not used for RoS. Transfers from the sewer are external to the surface domain but internal to the full coupled system; these normalizations therefore assess only the 2D budget. Uncorrected boundary-inclusive and corrected residuals were kept separate. RMSE peak times were aligned with rainfall, inundated area at 10 mm, storage, and recorded exchange-process peaks. Depth outputs have a 10 min resolution; the earliest tied process maximum was used.

A separate six-hour RoS diagnostic promoted P4 state storage to FP64 while retaining Q64 exchange calculations and the specified P4 arithmetic policy, including explicit FP32 velocity division in the CFL and dynamic-limiter expressions. The widened state was linked to the existing FP64 cell-data layout. Auxiliary storage for maximum values used the same precision as the state variables; its visualization output was disabled. With the intervention disabled, six short-run depth fields matched the prior diagnostic bitwise; checks of the compilation settings and selected exchange-volume records also passed. The long run provided 36 finite, nonnegative depth fields and retained the 528 mapped nodes

used for exchange comparisons. Exchange-volume records were checked for selected nodes over the recorded time intervals and for aggregate totals.

For the intervention, RMSE was evaluated against P6-Q64 on the time-specific union of elements with depths of at least 1 mm in any of the original twelve configurations. This fixed comparison mask was shared by the original P4 and the storage-promoted run at each output time. It differs from the pairwise union mask in Section 2.4.1. The intervention estimates the combined response of state storage and subsequent time-step and branch decisions. Its runtime was excluded from performance comparisons.

2.4.5. Additional Q64 Spatial Diagnostics

Additional spatial diagnostics compared P1-Q64 with P6-Q64 under RoG and RoS for the 80 mm/60 min design storm. These runs used the performance implementation identified in Section 2.4.2. Four snapshot runs were paired with four minimum-output controls; within each precision/rainfall condition, the paired runs differed only in diagnostic output. Actual output time, time step, dry cutoff and solver-step count were recorded alongside each depth snapshot. P1 retains FP32 state storage; its snapshot depths are written in FP64 format for comparison. Final surface budgets, 542-node summaries, solver/CFL counts and saved coupling-controller records matched among runs with output disabled, runs with output enabled, and the retained performance runs. The eight diagnostic runs were analyzed separately from the $n = 10$ timing sample. Their process-environment records confirm uniform Manning's roughness of 0.0125 s $m^{-1/3}$, the free-outfall boundary path and disabled 2D surface losses.

The four output runs provided 24 depth fields at nominal 10/20/30/40/50/60 min. Depths follow the writer's active-cell and dry-cut convention. Element IDs and areas cover all 5,572,301 triangles, and the metrics use the complete fields. P1-Q64 was compared with P6-Q64 using the pairwise $\geq$1 mm wet mask, 1/10/100 mm inundation thresholds, depth-conditioned masks and exchange/budget definitions in Sections 2.4.1 and 2.4.4. P6-Q64 served as the internal FP64 reference for these diagnostics.

Depth fields were paired by scheduled output time, with actual output times retained in the accompanying records. Departures from nominal times ranged from -0.0194 to 2.82×10^{-11} s, and the P1/P6 final-time difference was below 3×10^{-11} s. Input hashes and numerical settings were checked against the retained performance records. The analyses used ten-minute depth/storage histories and final exchange/budget summaries.

3. Results

3.1. Numerical Differences for the 6 h High-Intensity Rainfall Scenario

For the 6 h high-intensity rainfall scenario, we first retained FP64 for the exchange-discharge calculation and compared the surface-flow precision configurations with the P6-Q64 reference simulation. Then, we compared Q32 directly with Q64 within each surface-flow configuration to quantify the additional differences associated with the exchange precision changes.

3.1.1. Differences Among the Surface-Flow Precision Configurations

Under Q64 exchange, increasing the number of surface operations performed in FP64 did not consistently reduce water-depth differences from P6-Q64 (Figure 3). Final RMSEs for P1–P4 ranged from 0.520 to 1.187 mm under RoG and from 0.746 to 2.867 mm under RoS. P1 had final values of 0.520 and 0.746 mm, whereas P4 had 0.939 and 2.867 mm, respectively. P5, which retained FP64 state storage and used FP32 HLLC wave-speed estimates, had the smallest final RoG difference (0.172 mm) but a larger final RoS difference (2.513 mm) than P1. The maximum RMSEs over time for P1–P5 ranged from 2.637 to 3.269 mm under

RoG and from 4.482 to 5.919 mm under RoS. The different rankings demonstrate that the outcome depends on both the precision assignment and the rainfall-input pathway.

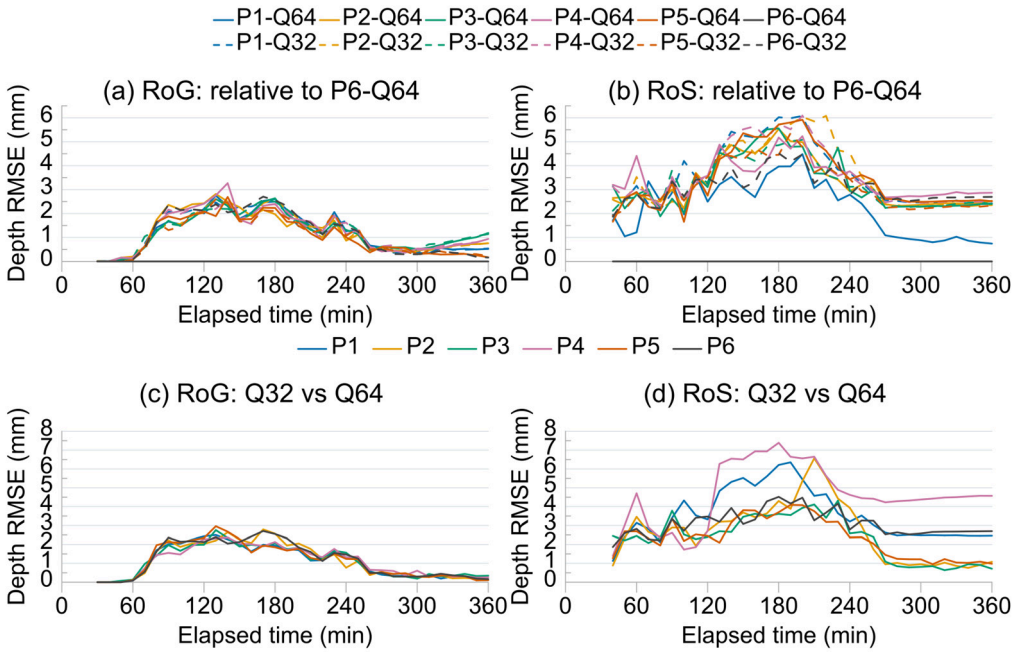


Figure 3. Temporal evolution of the water-depth RMSE for the 6 h simulation.

The P1-Q64/P6-Q64 RMSEs reached 2.745 mm at 130 min under RoG and 4.482 mm at 200 min under RoS. Figure 4 locates the corresponding element-wise differences, which are concentrated in different parts of the inundated domain for the two rainfall pathways.

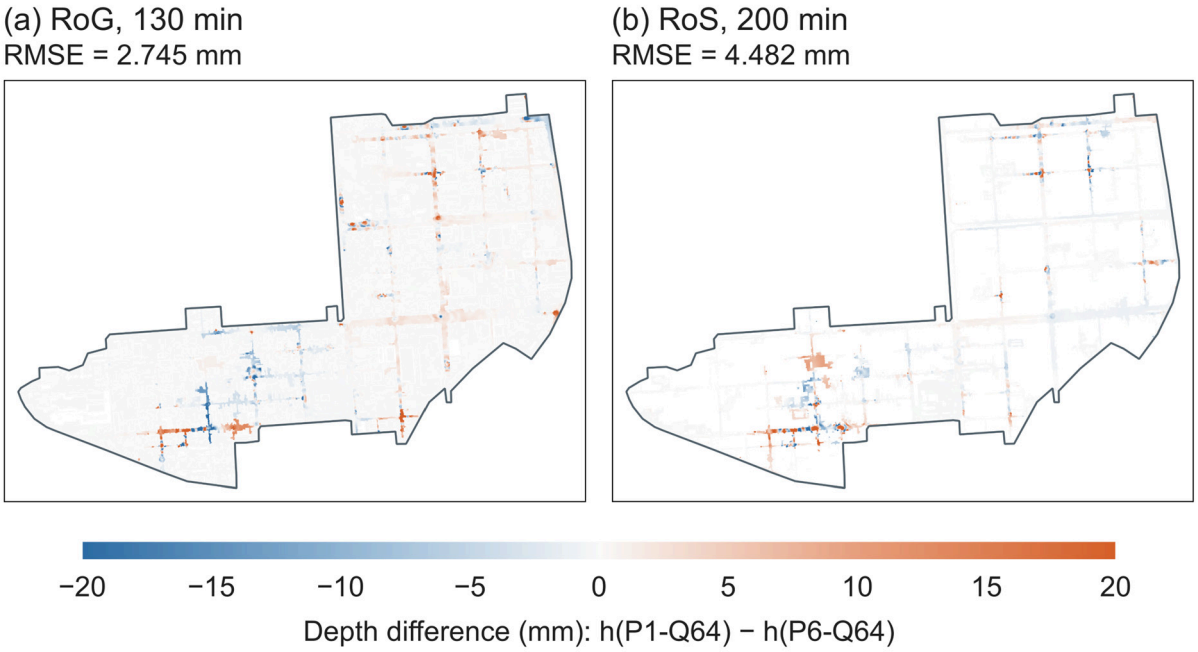


Figure 4. Spatial water-depth differences between P1-Q64 and P6-Q64 in the 6 h simulations.

3.1.2. Differences Among Exchange-Precision Settings

Changing exchange arithmetic from Q64 to Q32 produced larger water-depth differences under RoS than RoG (Figure 5). Final RMSEs ranged from 0.108 to 0.345 mm for RoG and from 0.716 to 4.571 mm for RoS; the corresponding maxima over time were 2.365–2.960 and 4.094–7.391 mm. Final 10 mm inundation IoU remained 0.9997–0.9999 and

0.9942–0.9985, respectively. The cumulative net-exchange WAPE ranged from 1.652% to 2.277% for RoG and from 0.674% to 1.623% for RoS. Under RoS, P4 had the largest final depth RMSE, whereas P2 had the largest manhole-level WAPE. Thus, spatial inundation agreement and sewer–surface exchange differences describe distinct aspects of the response.

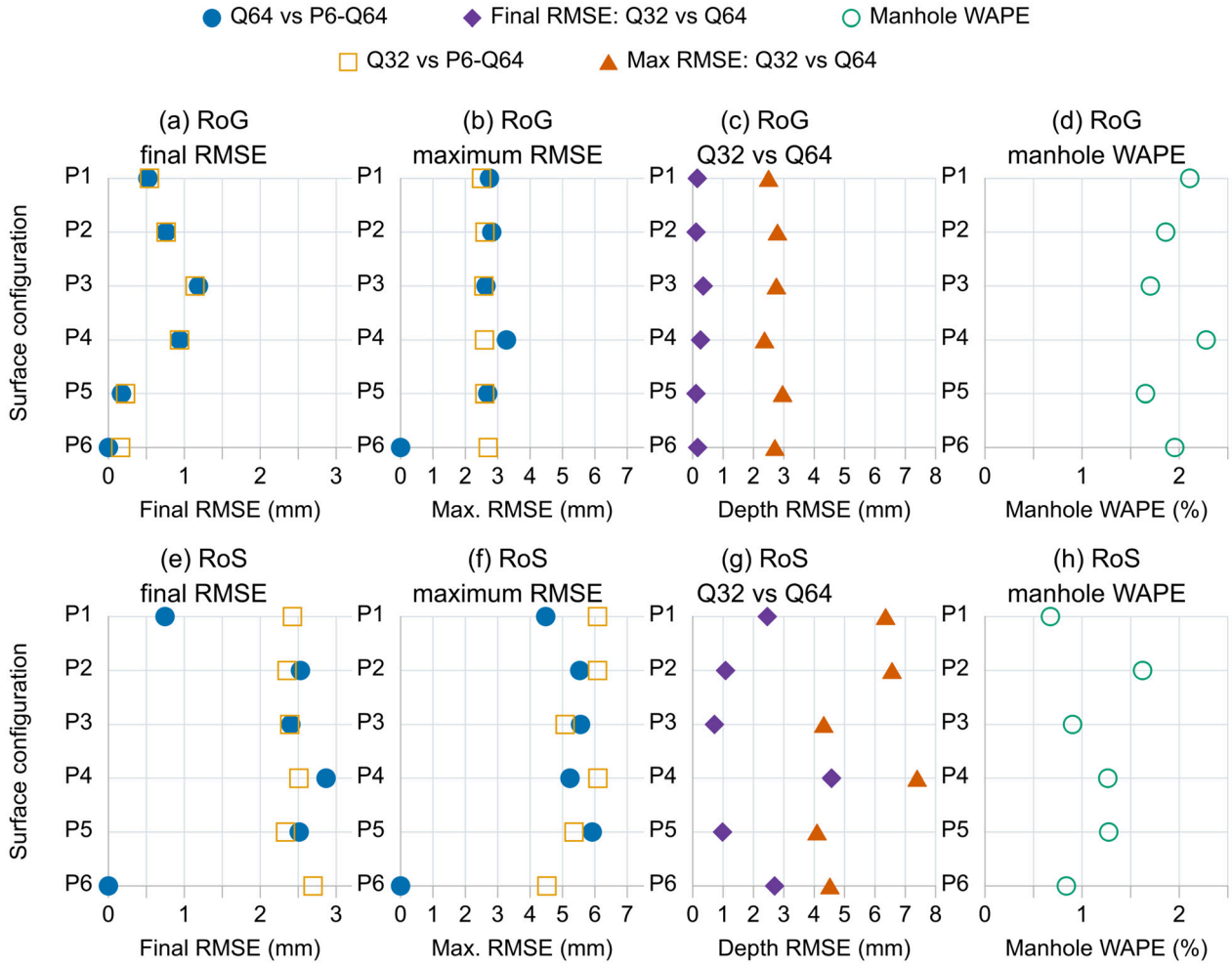


Figure 5. Water-depth and manhole exchange differences for the 6 h simulations.

3.2. Numerical Differences for the 1 h, 80 mm Design Storm

For the 1 h, 80 mm design storm, the maximum water-depth RMSEs for P1–P5 relative to P6-Q64 ranged from 1.056 to 1.418 and 2.983 to 7.110 mm for RoG and RoS, respectively. The end-of-simulation IoU for the 10 mm inundation extent ranged from 0.9988 to 0.9993 and 0.9930 to 0.9971, respectively.

For the 1 h, 80 mm design storm, the Q32–Q64 water-depth differences within the same surface-flow configuration were also larger for RoS than for RoG. The end-of-simulation RMSEs ranged from 0.850 to 1.492 and 3.580 to 8.410 mm for RoG and RoS, respectively, while the WAPE in cumulative net exchange volume at individual manholes ranged from 0.872 to 1.051% and 0.798 to 2.053%, respectively. Specifically, the P4-Q32 versus P4-Q64 comparison under RoS had an end-of-simulation area-weighted RMSE of 8.410 mm. For this same Q32–Q64 comparison, the 95th, 99th, and maximum element-wise absolute water-depth differences were 16.034, 42.718, and 289.127 mm, respectively. The maximum was substantially larger than the 99th percentile, indicating that the largest water-depth differences were confined to a small number of elements (Figure 6).

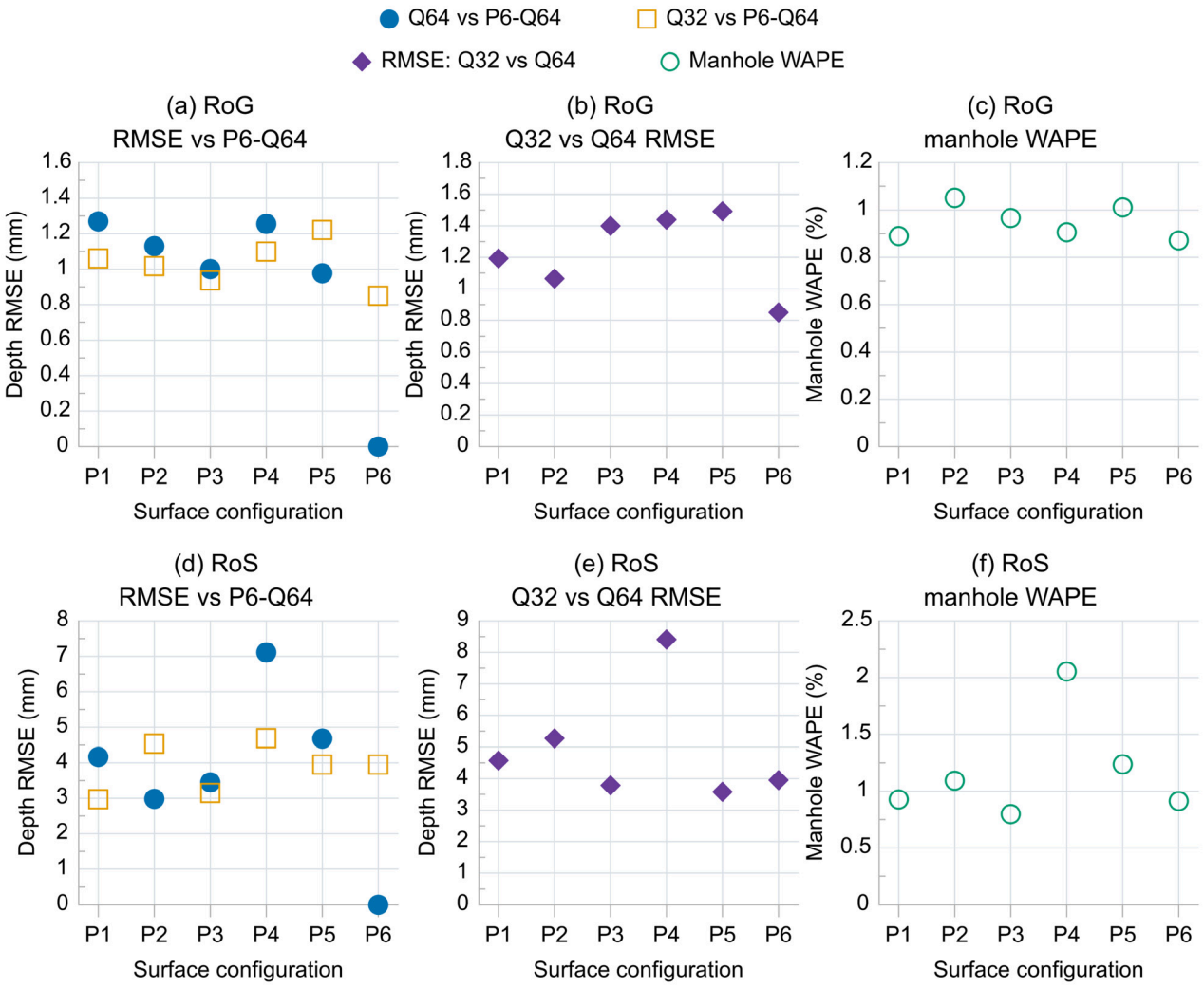


Figure 6. Water-depth and manhole exchange differences for the 1 h design storm.

3.3. Spatial Distribution of Water-Depth Differences Between Exchange-Precision Settings in the 6 h Simulations

Figure 7 maps Q32 minus Q64 depth for P1 and P4 with the pairwise wet mask in Section 2.4.1. P1 uses FP32 state storage and selected surface-arithmetic paths; P4 had the largest exchange-precision RMSE peak under RoS. RoG peaks occurred at 130/140 min for P1/P4 (2.498/2.365 mm); RoS peaks occurred at 190/180 min (6.353/7.391 mm). The later RoS peaks follow the sewer-first input pathway. Section 3.8.4 relates these times to the monitored exchange response.

3.4. Cumulative Exchange Volume at Manholes and 2D Surface-Domain Water Balance for the 6 h Simulations

The configurations with the largest depth RMSE, manhole-level exchange WAPE, and surface-budget residual differed. Relative to P6-Q64, net-exchange WAPE for P1–P5 ranged from 1.643% to 2.591% under RoG-Q64 and from 0.581% to 1.471% under RoS-Q64; the corresponding Q32 ranges were 1.632–1.972% and 0.862–1.419%. Because opposing exchanges can cancel each other out in the net volume, directional WAPE was also evaluated: sewer-to-surface differences were 5.393–7.061% for RoG and 0.861–1.911% for RoS, compared with 1.511–1.861% and 0.987–1.784% for surface-to-sewer exchange. Table 3 gives the corrected surface-budget residuals. The largest exchange-normalized value was 9.52×10^{-3}% for P3-Q64, while the absolute residuals for P5/P6 did not exceed 1.890×10^{-8} m^3. Close

corrected mass closure therefore coexisted with measurable differences in local depth and exchange volumes.

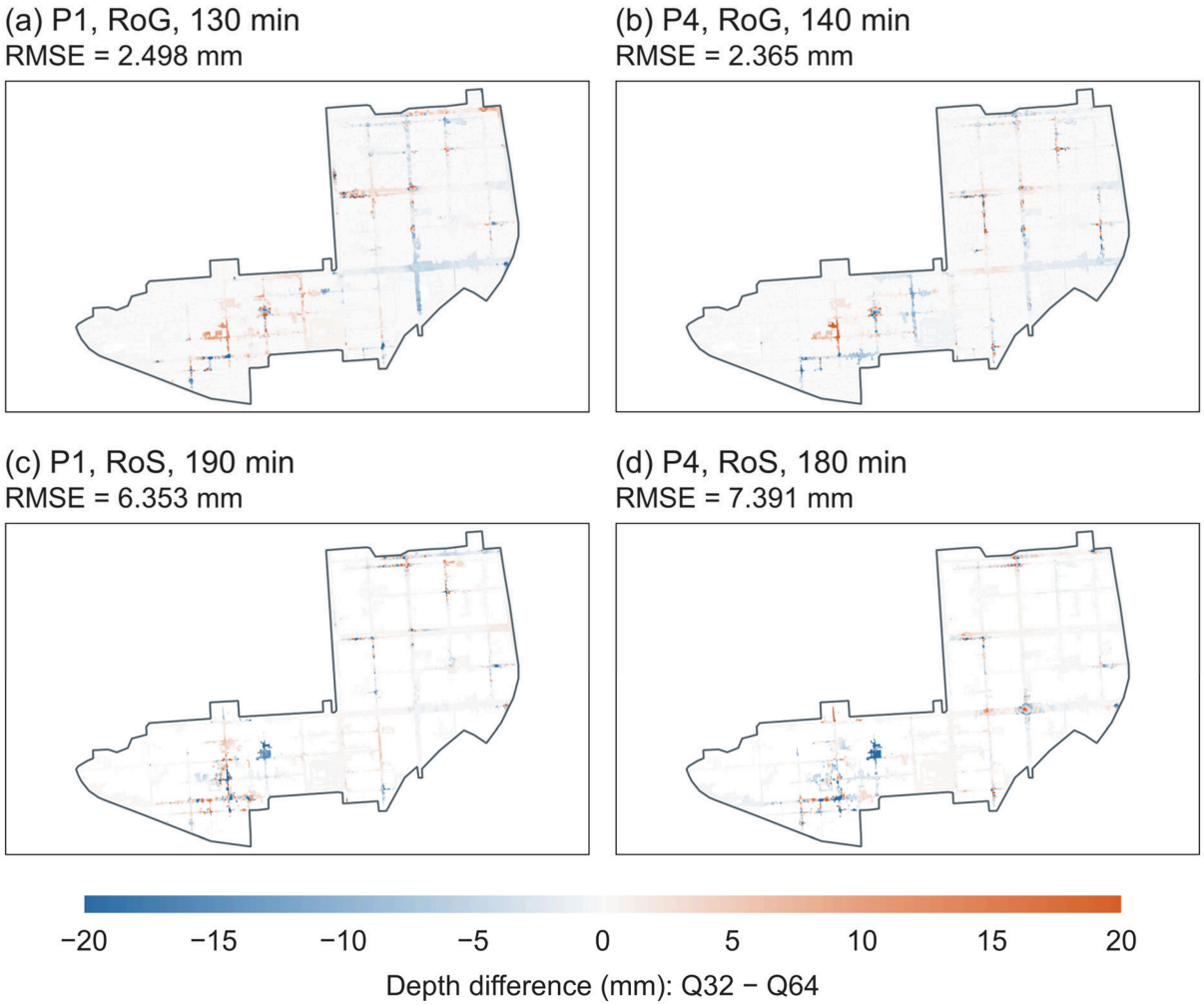


Figure 7. Spatial Q32–Q64 water-depth differences in the 6 h simulations.

Table 3. Water-balance residuals for the 2D surface domain in the 6 h simulations.

Condition	Surface-Flow Configuration	Q64 $\lvert R_2D \rvert$ (m^3)	Q32 $\lvert R_2D \rvert$ (m^3)	$\Delta \lvert R_2D \rvert = \lvert R_2D,Q32 \rvert - \lvert R_2D,Q64 \rvert$ (m^3)	Larger of the Q64 and Q32 Normalized Residuals (%)
RoG	P1	0.481	0.677	0.196	1.10×10^{-4}
RoG	P2	12.621	12.857	0.236	2.05×10^{-3}
RoG	P3	60.032	59.608	−0.424	9.52×10^{-3}
RoG	P4	26.519	27.624	1.104	0.00440
RoG	P5	1.006×10^{-8}	5.072×10^{-9}	-4.988×10^{-9}	1.610×10^{-12}
RoG	P6	1.362×10^{-8}	1.890×10^{-8}	5.277×10^{-9}	3.006×10^{-12}
RoS	P1	0.336	0.388	0.052	0.00004
RoS	P2	2.044	2.282	0.238	0.00024
RoS	P3	2.297	2.060	−0.237	0.00024

Table 3. *Cont.*

Condition	Surface-Flow Configuration	Q64 $\lvert R_2D \rvert$ (m^3)	Q32 $\lvert R_2D \rvert$ (m^3)	$\Delta \lvert R_2D \rvert = \lvert R_2D,Q32 \rvert - \lvert R_2D,Q64 \rvert$ (m^3)	Larger of the Q64 and Q32 Normalized Residuals (%)
RoS	P4	2.970	2.440	−0.530	0.00031
RoS	P5	1.554×10^{-9}	1.717×10^{-10}	-1.382×10^{-9}	1.619×10^{-13}
RoS	P6	1.598×10^{-9}	1.417×10^{-9}	-1.811×10^{-10}	1.664×10^{-13}

Note: R_{2D} is the 2D surface-domain water-balance residual, and $E_{abs} = \lvert E_{1\to2} \rvert + \lvert E_{2\to1} \rvert$ is the total absolute bidirectional exchange volume. The exchange-volume-normalized residual was calculated as $R_{norm} = 100 \lvert R_{2D} \rvert / E_{abs}$. The last column reports the larger Q64 and Q32 values for each surface-flow precision configuration.

Figure 8 compares Q32 with Q64 within each surface configuration. The upper panels show manhole-level WAPE in cumulative net exchange volume; the lower panels show the 95th and 99th percentiles and maximum absolute differences on logarithmic horizontal axes.

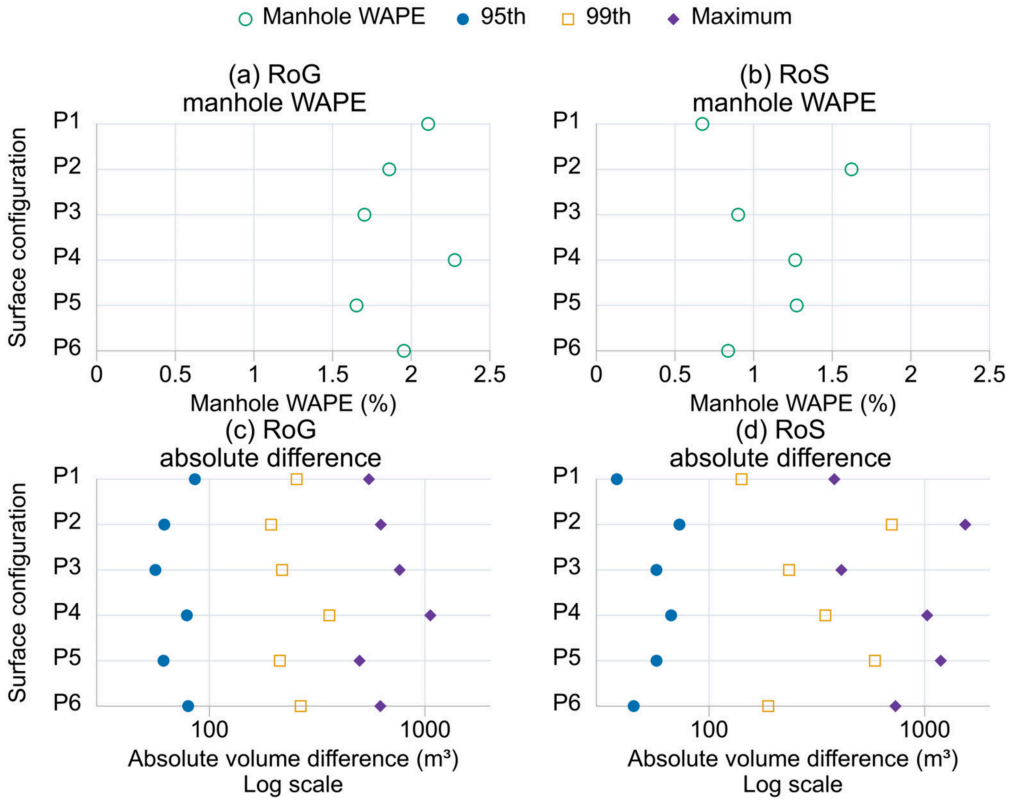


Figure 8. Q32–Q64 differences in cumulative net exchange volume for the 6 h simulations.

Figure 9 shows the absolute and exchange-normalized corrected 2D water-balance residuals on logarithmic vertical axes. These residuals quantify budget closure after the numerical correction terms have been included; dry-cell removal and other correction terms quantify the corresponding numerical adjustments.

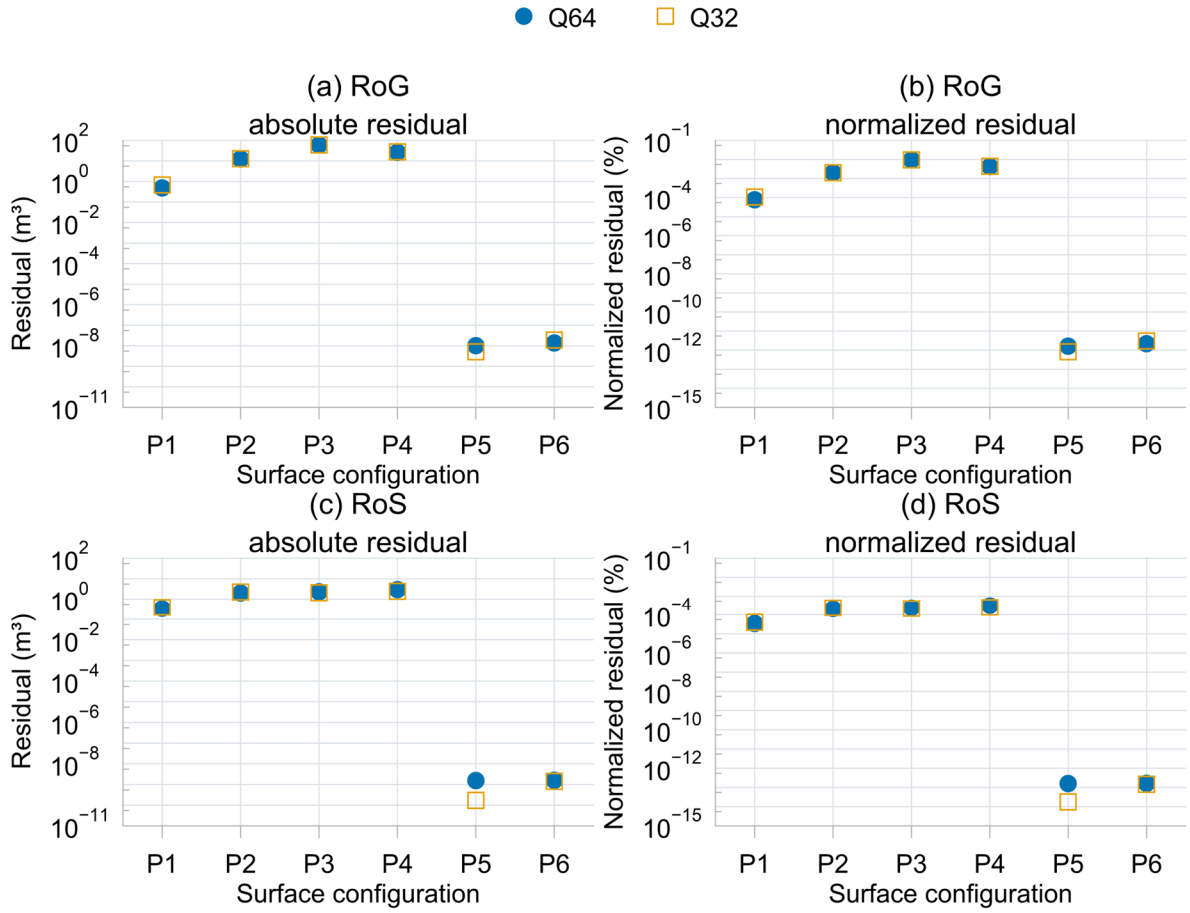


Figure 9. Absolute and exchange-volume-normalized water-balance residuals for the 2D surface domain in the 6 h simulations.

3.5. Computational Performance Across Precision Configurations

Tables A1 and A2 give the runtime statistics for all 24 conditions (n = 10), and Figure 10 shows median runtime ratios with intervals estimated by repeated sampling. For P1-Q64 relative to P6-Q64, the ratios were 1.883 (95% interval 1.876–1.887) for RoG and 1.205 (1.199–1.212) for RoS. The P1-Q64 and P6-Q64 median times were 490.522 and 923.628 s for RoG, and 156.916 and 189.124 s for RoS, respectively. These intervals support lower P1-Q64 runtime relative to P6-Q64 under the specified sampling procedure.

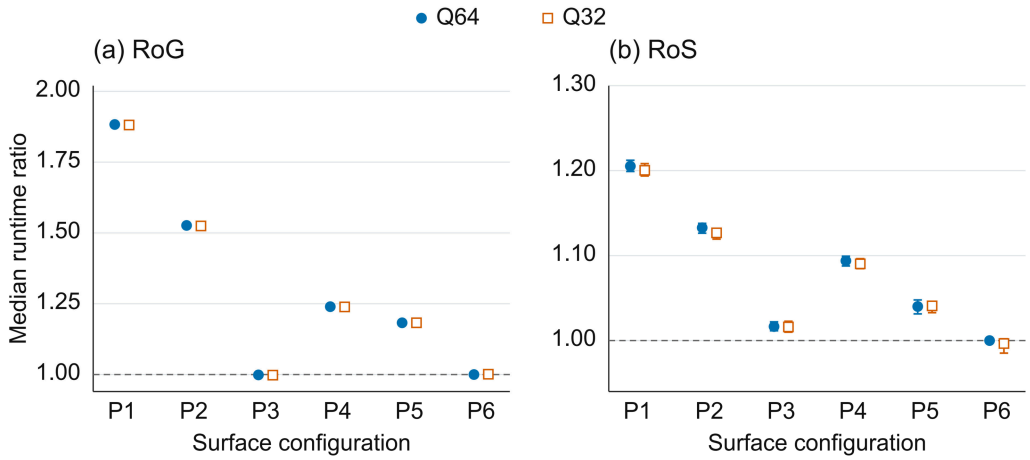


Figure 10. Median runtime ratios for the 1 h performance experiment (n = 10): (**a**) RoG; (**b**) RoS.

The ratios in Figure 10 are the P6-Q64 median runtime divided by the median runtime for each configuration, so the reference ratio is one. The bars show the exploratory 95% intervals defined in Section 2.4.2; RoG and RoS use different vertical scales.

Across the twelve settings, relative runtime variation (sample standard deviation divided by mean runtime, expressed as a percentage) ranged from 0.198 to 0.574% for RoG and 0.316 to 2.112% for RoS. Tables A1 and A2 give SDs and IQRs. For within-policy Q64/Q32 median-time ratios, all six RoG intervals spanned one. RoS P1/P2/P4 intervals lay below one, corresponding to slightly longer Q32 runtime; P3/P5/P6 intervals spanned one. Thus, FP32 exchange operations provided no consistent runtime saving in this implementation.

3.6. Temporal Evolution of Water-Depth Differences in the 48 h RoS Simulation

In the 48 h RoS test, depth differences increased during later rainfall and decreased during drainage (Table 4; Figure 11). The P1-Q64/P6-Q64 RMSE peaked at 5.666 mm at 2520 min and ended at 0.854 mm. P1-Q32/P6-Q64 and P1-Q32/P1-Q64 peaked at 9.823/11.347 mm at 2640 min and ended at 1.020/1.455 mm. The exchange-precision storage difference reached 166.007 m^3 at 2650 min. These event-specific trajectories show how rainfall and drainage modulate the differences.

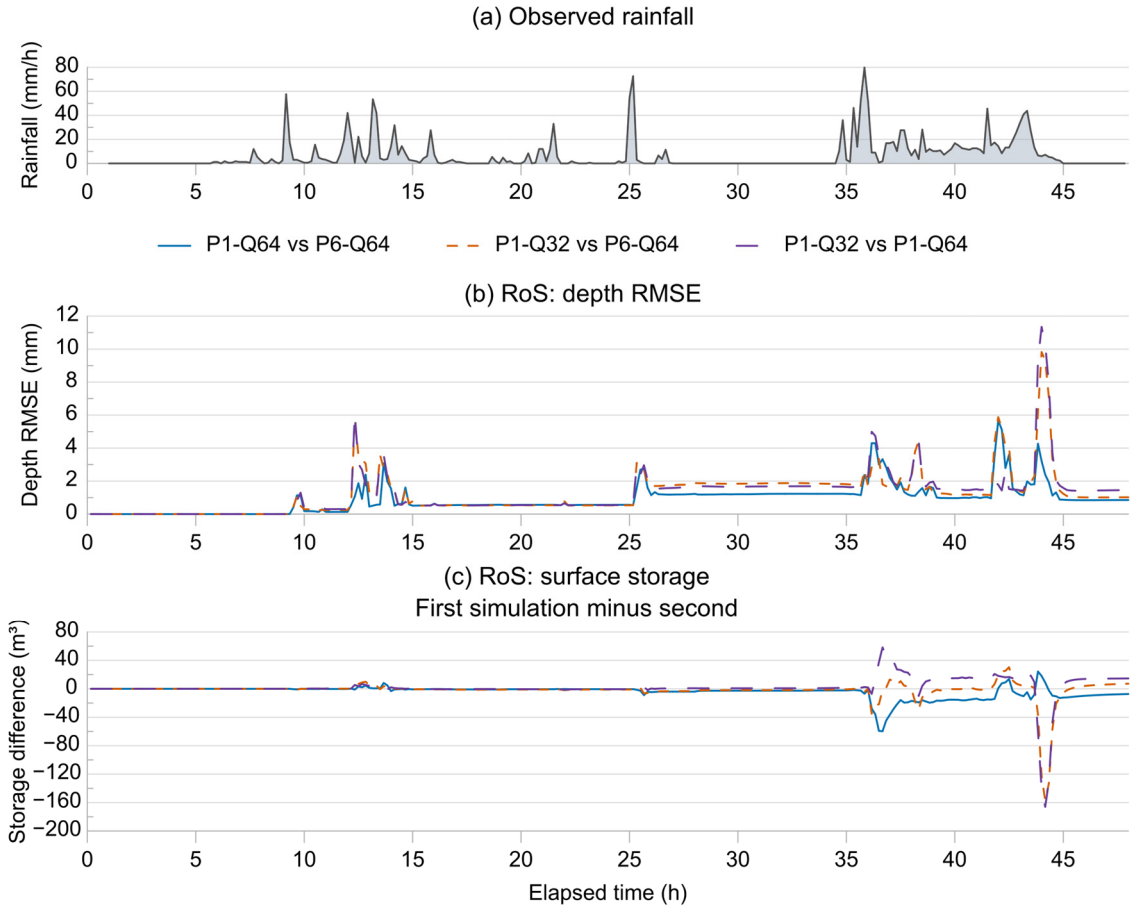


Figure 11. Water-depth and surface-storage differences for the 48 h RoS simulations.

Table 4. Water depth, inundation extent, and surface-water storage volume comparison between the 48 h RoS simulations.

Pairwise Comparison	Maximum RMSE (mm)	Time of Maximum RMSE (min)	End-of-Simulation RMSE (mm)	End-of-Simulation IoU for the 10 mm Inundation Extent	End-of-Simulation Difference in Surface-Water Storage Volume (First Configuration − Second Configuration) (m^3)
P1-Q64−P6-Q64	5.666	2520	0.854	0.9975	−7.358
P1-Q32−P6-Q64	9.823	2640	1.020	0.9975	+7.204
P1-Q32−P1-Q64	11.347	2640	1.455	0.9960	+14.561

The simulation outputs in Figure 11 are sampled at 10 min intervals. Surface-storage differences are calculated as the first configuration minus the second in each legend entry.

3.7. Precision-Dependent Numerical Trajectories

In the P1-Q64/P4-Q64 diagnostic, a monitored source-depth intermediate first differed at 1810 s while stored depth remained equal after rounding. Flux and stored-depth differences appeared at 1811 s under the same recorded step. At 1812 s, the global CFL candidates were 0.9656458497/0.9656515892 s. Fixed-state replay reproduced both with the policy-specific coordinate spans and CFL expressions, identifying sensitivity to arithmetic and geometric precision in the selected trajectory.

The 60 min RoS diagnostics required 32,805/32,725/32,806 steps for P1/P4/P5-Q64. Each had 220 coupling schedule entries (192 at 10 s and 28 at 60 s). This shared schedule accompanied different internal step counts. Schedule entries count prescribed coupling intervals; transition records count selected-cell changes.

In the P4 diagnostic, signed rounding volumes after flux and source updates were -0.0261135836 and -0.0004871151 m^3. Their sum, -0.0266006987 m^3, agreed with the corrected diagnostic residual within approximately 1.77×10^{-10} m^3, directly linking storage rounding to closure in this test. Dry-cut loss was 0.0424937542 m^3 and the boundary-inclusive uncorrected residual was -0.0691174777 m^3. Table 3 retains the separate full six-hour budgets.

3.8. Hydraulic Differences and Computational Performance

3.8.1. Depth-Conditioned Differences and the Local Maximum

For the one-hour RoS P4-Q32 versus P4-Q64 comparison at 60 min, reference depths of at least 0.10 and 0.30 m gave area-weighted RMSEs of 9.302 and 11.295 mm, respectively (Table 5). The corresponding area-weighted 99th percentiles of absolute differences were 49.185 and 54.460 mm. These statistics use reference-depth masks and area-weighted percentiles. The all-wet percentiles in Section 3.2 give equal weight to each element.

Table 5. Depth-conditioned differences for P4-Q32 relative to P4-Q64 at the end of the one-hour RoS design storm.

Reference Depth (m)	Elements	Area (m^2)	RMSE (mm)	Area-Weighted P99 (mm)
$\geq$0.10	519,778	225,041.886	9.302	49.185
$\geq$0.30	127,918	55,364.443	11.295	54.460
$0.10 \leq h < 0.30$	391,860	169,677.443	8.552	42.401

The reproduced 289.127 mm local maximum occurred in element 3,520,405 (0.452066 m^2): Q64 depth was 0.289127 m and Q32 depth was zero. Absolute differences $\geq$100/$\geq$200 mm occupied 305.505/6.890 m^2. The nearest mapped manhole was 57.667 m away. The maximum difference occurred where Q64 predicted inundation and Q32 predicted a dry cell. The selected trajectory diagnostics provide additional information on the numerical processes involved.

3.8.2. Six-Hour Runtime and Numerical Workload

Table 6 gives the original six-hour wall times and actual step counts. All RoG settings had 2035 coupling entries (2010 at 10 s; 25 at 60 s), and RoS had 2020 (1992 at 10 s; 28 at 60 s). Internal step counts varied by policy. The times correspond directly to the six-hour hydraulic outputs, with one observation per configuration and the historical device-record coverage stated in Section 2.4.4.

Table 6. Original six-hour accuracy-run wall-clock times and 2D solver-step counts (one observation per configuration).

Configuration	RoG Time (s)	RoS Time (s)	RoG Steps	RoS Steps
P1-Q64	7138.175	2346.835	1,133,265	545,962
P1-Q32	7370.428	2400.702	1,133,265	546,042
P2-Q64	8737.439	2482.524	1,133,255	546,115
P2-Q32	8923.095	2615.601	1,133,255	546,128
P3-Q64	11,666.841	2944.447	1,133,918	546,046
P3-Q32	12,335.863	3001.274	1,133,918	546,031
P4-Q64	9883.232	2796.869	1,133,919	546,150
P4-Q32	11,482.234	2785.266	1,133,919	545,946
P5-Q64	10,150.757	2969.182	1,134,117	546,046
P5-Q32	10,468.356	2887.014	1,134,117	546,054
P6-Q64	11,722.337	3683.693	1,134,117	546,176
P6-Q32	11,956.412	3032.415	1,134,117	546,049

Table 7 summarizes representative component shares from four original six-hour runs. Surface-module shares were 98.138–99.645% of wall time. The nested coupling share was larger for RoS than RoG in both P1-Q64 and P6-Q64, helping explain how reduced surface arithmetic transfers to whole-run savings. The separate one-hour performance experiment and component profiling are reported in Sections 3.5 and 3.8.6.

Table 7. Representative six-hour timing shares as percentages of total wall-clock time.

Case	Other Surface Work	Coupling Total	SWMM	Unassigned
ROG P1-Q64	72.547	26.878	0.130	0.445
ROG P6-Q64	82.171	17.473	0.079	0.276
ROS P1-Q64	65.068	33.070	0.390	1.472
ROS P6-Q64	62.231	36.518	0.276	0.974

In Table 7, other surface work is surface-module time minus nested coupling time and includes the remaining surface computations. The columns partition the measured top-level accounting; unassigned time is its wall-time remainder. Coupling includes synchronization/preparation, discharge calculation, distribution, application and accounting.

3.8.3. Complementary Normalizations of the Surface Budget

Table 8 adds normalizations of the corrected residual by direct surface rainfall and by surface-external inflow. The largest six-hour corrected residual relative to direct surface rainfall was 0.009298% for RoG P3-Q64. Relative to surface-external inflow, the maxima were 0.007613% for RoG P3-Q64 and 0.000530% for RoS P4-Q64. These small corrected values must be distinguished from the boundary-inclusive uncorrected budget: the corresponding absolute uncorrected residuals as a fraction of surface-external inflow ranged from 0.701374% to 0.723266% under RoG and from 0.000031% to 0.000494% under RoS. Explicitly accounted clipping and internal-flux terms explain why the two residual definitions differ.

Table 8. Additional normalizations of the absolute corrected six-hour surface-domain residual (%).

Configuration	RoG/Rainfall	RoG/Surface Inflow	RoS/Surface Inflow
P1-Q64	7.450×10^{-5}	6.116×10^{-5}	5.999×10^{-5}
P1-Q32	1.049×10^{-4}	8.609×10^{-5}	6.918×10^{-5}
P2-Q64	1.955×10^{-3}	1.603×10^{-3}	3.651×10^{-4}
P2-Q32	1.991×10^{-3}	1.634×10^{-3}	4.057×10^{-4}
P3-Q64	9.298×10^{-3}	7.613×10^{-3}	4.101×10^{-4}
P3-Q32	9.232×10^{-3}	7.563×10^{-3}	3.685×10^{-4}
P4-Q64	4.107×10^{-3}	3.364×10^{-3}	5.296×10^{-4}
P4-Q32	4.278×10^{-3}	3.509×10^{-3}	4.357×10^{-4}
P5-Q64	1.558×10^{-12}	1.280×10^{-12}	2.776×10^{-13}
P5-Q32	7.855×10^{-13}	6.447×10^{-13}	3.072×10^{-14}
P6-Q64	2.110×10^{-12}	1.732×10^{-12}	2.852×10^{-13}
P6-Q32	2.927×10^{-12}	2.399×10^{-12}	2.532×10^{-13}

3.8.4. Peak Timing and the Forty-Eight-Hour Spatial Comparison

P1-Q64/P6-Q64 RMSEs peaked 34 min after rainfall under RoG and 104 min after rainfall under RoS (Table 9). At 10 min output resolution, the RoG peak coincided with the reference storage peak. The RoS peak followed the monitored surface-to-sewer exchange peak at 190.833 min by 9.167 min. These temporal relationships describe the different rainfall–sewer–surface responses at the retained sampling resolution.

Table 9. Peak times and lags for the six-hour P1-Q64 versus P6-Q64 RMSE comparison (minutes from the simulation start).

Process	RoG Peak	RoG RMSE Lag	RoS Peak	RoS RMSE Lag
Rainfall	96.000	34.000	96.000	104.000
10 mm wet area	110.000	20.000	130.000	70.000
Surface storage	130.000	0.000	130.000	70.000
Sewer overflow	128.000	2.000	126.333	73.667
Selected sewer-to-surface	132.500	−2.500	126.333	73.667
Selected surface-to-sewer	132.333	−2.333	190.833	9.167

For the 48 h P1-Q32/P1-Q64 comparison, the largest absolute 10 mm wet-area difference was 686.948 m^2 at 2660 min, 20 min after peak RMSE. Final IoU was 0.9790/0.9960/0.9976 at 1/10/100 mm, showing the dependence of spatial agreement

on depth threshold. Rainfall and wet-area peaks at 2150/2180 min preceded the peak exchange-precision RMSE at 2640 min. These timings extend the drainage-response interpretation in Section 3.6.

3.8.5. Six-Hour P4 State-Storage Intervention

On the original twelve-configuration common wet mask, promoting P4 state storage to FP64 reduced RMSE from 5.194 to 4.109 mm at 200 min and from 2.822 to 1.582 mm at 360 min (Table 10). The common-mask 2.822 mm value and pairwise-mask 2.867 mm value in Section 3.1.1 use different masks. RMSE decreased at 28 of 33 nonempty outputs and increased at five; the first three outputs had empty masks. The maximum shifted from 200 to 120 min and decreased from 5.194 to 4.219 mm. This intervention demonstrates state-storage sensitivity and a time-dependent response.

Table 10. Original P4 and FP64-state diagnostic results for the six-hour RoS case.

Metric	Original P4-Q64	P4-STATE64-Q64
RMSE at 200 min (mm)	5.194	4.109
RMSE at 360 min (mm)	2.822	1.582
Maximum output-time RMSE (mm)	5.194	4.219
Time of maximum RMSE (min)	200	120
Corrected diagnostic residual (m^3)	2.969886	1.540606×10^{-9}
Boundary-inclusive uncorrected residual (m^3)	2.768700	−0.163193

The corrected residual decreased from 2.969886 to 1.540606×10^{-9} m^3; the FP64-state run retained a boundary-inclusive uncorrected residual of −0.163193 m^3. Together with the depth comparison, these quantities characterize storage effects on numerical trajectories and budget closure for the tested cell-based solver.

3.8.6. Component Execution Times

Table 11a reports the host-time partition from the separate 60 min profiling runs. Dynamic call intervals occupied 46.380–65.457% under RoG and 16.997–25.777% under RoS. Input and initialization accounted for 3.331–6.145% under RoG and 15.879–18.763% under RoS; SWMM calls accounted for 0.075–0.493% across the four conditions. Each host-call interval includes submission and any blocking inside that call. Every instrumented run had zero stack errors, and the exclusive intervals summed exactly to model wall time at nanosecond resolution.

Values are mean seconds (mean percentage of model wall time), with $n = 3$ trace runs per condition. Percentages were averaged within repeats and sum to 100% before rounding. NR denotes an uninvoked routine. The separately marked state-update routines had zero calls in these conditions. Other surface work includes remaining rainfall, management and state work outside the marked calls; other coupling work includes submission and remaining CPU coupling work. Output intervals cover complete routine calls; unclassified time is the remaining measured host interval. This 60 min profile is separate from the six-hour cases in Table 7 and the $n = 10$ performance statistics.

Values are mean seconds (mean wall-time percentage), with $n = 3$. GPU durations sum activity records; synchronization durations use the union of CUDA synchronization API intervals on the host thread. These activities can overlap each other and the host partition in Table 11a. The fused exchange kernel includes discharge evaluation, distribution and FP64 ledger updates. CUDA API intervals cover synchronization inside the measured

routines; in Table 11a WAIT covers explicitly marked OpenACC calls. Copies are directly observed explicit memcpy activities.

Table 11. Component execution times for the 60 min design storm. (**a**) Exclusive host-call intervals. (**b**) GPU activities and CUDA synchronization API intervals.

(**a**)

Host Interval	RoG P1-Q64	RoG P6-Q64	RoS P1-Q64	RoS P6-Q64
Input and initialization	30.875 (6.145)	31.132 (3.331)	31.093 (18.763)	31.110 (15.879)
SWMM solver calls	0.711 (0.142)	0.705 (0.075)	0.816 (0.493)	0.795 (0.406)
Surface dynamic calls	233.053 (46.380)	611.819 (65.457)	28.167 (16.997)	50.501 (25.777)
Surface CFL calls	67.104 (13.354)	112.001 (11.983)	13.181 (7.952)	15.349 (7.834)
Active-list calls	26.418 (5.258)	30.267 (3.238)	27.105 (16.354)	31.613 (16.136)
Rainfall routine calls	3.857 (0.768)	3.618 (0.387)	NR	NR
Other surface work	87.037 (17.321)	88.838 (9.505)	47.137 (28.445)	47.153 (24.068)
Other coupling work	15.878 (3.160)	15.872 (1.698)	8.257 (4.983)	7.959 (4.063)
H2D API calls	0.073 (0.014)	0.072 (0.008)	0.079 (0.047)	0.074 (0.038)
D2H API calls	2.894 (0.576)	5.629 (0.602)	2.125 (1.283)	3.967 (2.025)
Explicit OpenACC WAIT	33.888 (6.744)	34.055 (3.643)	7.136 (4.305)	6.771 (3.456)
Output routine calls	0.027 (0.005)	0.027 (0.003)	0.028 (0.017)	0.027 (0.014)
Shutdown	0.027 (0.005)	0.027 (0.003)	0.036 (0.022)	0.036 (0.018)
Unclassified	0.642 (0.128)	0.627 (0.067)	0.561 (0.339)	0.561 (0.286)
Model wall time	502.485 (100.000)	934.692 (100.000)	165.718 (100.000)	195.917 (100.000)

(**b**)

Separate Measurement	RoG P1-Q64	RoG P6-Q64	RoS P1-Q64	RoS P6-Q64
Fused exchange kernel	7.536 (1.500)	7.428 (0.795)	3.888 (2.346)	3.847 (1.964)
GPU H2D copies	0.216 (0.043)	0.233 (0.025)	0.180 (0.109)	0.181 (0.092)
GPU D2H copies	2.717 (0.541)	5.447 (0.583)	1.979 (1.194)	3.828 (1.954)
CUDA sync API union	338.804 (67.426)	766.385 (81.993)	64.042 (38.642)	93.233 (47.588)

3.8.7. Q64 Spatial and Exchange Diagnostics

Table 12a summarizes the additional P1-Q64/P6-Q64 diagnostics for the 80 mm design storm. The final RMSE was 1.201 mm for RoG and 4.408 mm for RoS. Final 10 mm inundation IoU was 0.9989/0.9948 and net-exchange WAPE was 1.079/1.104%. These diagnostics characterize the depth and exchange differences between the two precision settings under the conditions in Section 2.4.5.

RMSE uses the area-weighted pairwise $\geq$1 mm wet mask. Peak RMSE and largest absolute depth difference cover the six scheduled snapshots; final metrics refer to 60 min. Both rows compare P1-Q64 with P6-Q64 within the implementation specified in Section 2.4.5. Net-exchange WAPE uses 528 mapped nodes.

Masks use P6-Q64 reference depth from the additional spatial diagnostics. RMSE and MAE are area-weighted; the 99th percentile is the left inverse of cumulative element area. All masks are nonempty.

Across the six P1/P6 snapshot pairs, the largest absolute differences were 116.605 mm under RoG at 60 min and 122.608 mm under RoS at 40 min. At the RoG maximum, P6/P1 depths were 0.009978/0.126583 m in element 5,077,550. At the final RoS local maximum, they were 0.330694/0.215303 m in element 3,600,600, a 115.390 mm difference. The 10 mm wet-area disagreement occupied 862.943/1906.077 m^2 at the final RoG/RoS time. Table 12b reports the deeper-reference masks; RoS RMSE was 4.988/5.074 mm at $\geq$0.10/$\geq$0.30 m, with 99th-percentile absolute differences of 22.706/22.794 mm.

P1-minus-P6 final surface storage in the additional diagnostics was -8.237 m^3 under RoG and -122.743 m^3 under RoS. Maximum absolute node net-exchange differences were 55.634 m^3 at node 6203-1 and 205.170 m^3 at node 10414-01. The saved coupling-controller sequences matched within each rainfall method, while P1/P6 solver-step counts were 103,254/103,262 for RoG and 53,380/53,392 for RoS. Final full-diagnostic surface-budget residuals were $-0.786164/-0.009165$ m^3 for P1 under RoG/RoS and $-1.217 \times 10^{-8}/-1.263 \times 10^{-9}$ m^3 for P6. The separately accounted RoG dry-cut volumes were 3180.113/3179.815 m^3 for P1/P6.

Table 12. P1-Q64 relative to P6-Q64 in the additional 80 mm/60 min spatial diagnostics. (**a**) Domain-level and exchange metrics. (**b**) Metrics under reference-depth masks at 60 min.

(**a**)

Comparison	Rainfall	Final RMSE (mm)	Peak RMSE (mm)	Largest $\lvert\Delta h\rvert$ (mm)	Final 10 mm IoU	Net-Exchange WAPE (%)
P1−P6	RoG	1.201	1.201	116.605	0.9989	1.079
P1−P6	RoS	4.408	4.408	122.608	0.9948	1.104

(**b**)

Rainfall	Reference Depth	Elements	Area (m^2)	RMSE (mm)	MAE (mm)	99th Abs. Difference (mm)
RoG	≥0.10 m	749,106	324,627.2	2.279	0.446	8.805
RoG	≥0.30 m	185,734	80,428.1	2.039	0.283	2.665
RoS	≥0.10 m	520,673	225,405.9	4.988	2.094	22.706
RoS	≥0.30 m	129,157	55,906.6	5.074	2.373	22.794

4. Discussion

4.1. Reference-Relative Numerical Differences and Water Balance Across Component-Wise Precision Configurations

The precision rankings varied across metrics and rainfall pathways. P2–P4 retained more FP64 operation groups than P1, yet yielded different rankings under RoG and RoS. The state-storage intervention connects this nonmonotonic behavior to stored-state sensitivity and subsequent trajectories: RMSE decreased at 28 of 33 nonempty outputs and increased at five. Rounding, CFL steps, wetting/drying and exchange thresholds provide pathways for precision changes to propagate [14,26,30], consistent with selective-precision and compensation studies [16,17]. Section 2.2 describes the reconstruction and intermediate-operation differences relevant to P4/P5, while the controlled intervention evaluates the combined response to P4 storage promotion.

RoG wets the surface directly, whereas RoS supplies the surface through sewer surcharge. The later RoS RMSE peak and spatial patterns follow these different pathways. Differences also persisted in reference-depth masks ≥0.10/≥0.30 m. The local wet/dry disagreement and depth-conditioned results in Section 3.8.7 show why locations and depth bands are useful alongside domain RMSE when selecting precision.

The corrected 2D water-balance residuals for P5/P6 were close to zero. For P4, the recorded volume changes caused by storage rounding and the storage intervention directly linked storage precision to budget closure. Depth agreement and budget closure assess complementary properties. Section 3.8.3 supplies rainfall and surface-inflow normalizations, with clipping and other corrections retained as explicit accounting terms.

4.2. Changes in Wall-Clock Time Across Precision Configurations

GPU-based shallow-water studies have examined solver throughput and parallel scalability. TRITON reported growing communication/I/O shares with additional GPUs [9]; SERGHEI-SWE compared performance portability across CPU/GPU systems [11]. Li et al. [15] combined FP32 kernels and FP64 accumulation. Our whole-run timer captures sewer routing, surface flow, exchange and reporting within the boundary defined in Section 2.4.2, allowing component improvements to be assessed at the coupled-model level.

The P6-Q64/P1-Q64 median ratios for the surface solver's Dynamic calls were 2.650 for RoG and 1.850 for RoS, compared with 1.883 and 1.205 for whole-run time and 1.941 and 1.252 for surface total. Initialization and other work account for the smaller whole-run savings. The component timers provide the ratios reported here. Section 3.8.6 reports separate host-time measurements partitioned into non-overlapping intervals. Q64/Q32 timing contrasts in Section 3.5 assess the implemented exchange-precision settings with their FP64 support calculations.

Q32 changes selected head and integration operations, while available-volume reductions, capacity limits, distribution and cumulative accounting retain FP64 calculations. In the separate Q64 profiles, accumulated GPU time for the fused exchange kernel, including discharge evaluation, distribution and ledger updates, was equivalent to 0.795–2.346% of model wall time (Table 11b). Surface-solver calls, initialization and other work also contributed to whole-run time (Table 11a). The limited scope of the FP32 operations and the time spent on other work are consistent with the small and configuration-dependent effect of Q32 on total runtime.

The larger dynamic share under RoG and larger initialization share under RoS help explain their different whole-run savings. Table 11b adds fused exchange-kernel, explicit-copy and synchronization activity, while Table 11a supplies exclusive host-call shares. When memory transfers limit execution, FP32 storage can reduce the number of bytes read and written. When arithmetic execution limits runtime, the potential benefit depends on the GPU's FP32 and FP64 processing rates. Kernel bottleneck classification would require hardware counters beyond these activity traces.

4.3. Practical Interpretation and Scope of Transfer

Precision selection for repeated hydraulic scenarios requires consideration of local depth, sewer–surface exchange and computational cost. The hydraulic analyses describe the numerical response of the precision settings, while the separate one-hour performance experiment compares runtime under common execution conditions. For repeated local scenario screening, candidate configurations can first be compared with P6-Q64 at locations relevant to flood-management decisions, using local depth, inundation classification and exchange volume. The runtime comparison then indicates the computational cost of each setting.

P1-Q64 had the shortest median runtime among the tested configurations in the one-hour performance experiment. In the hydraulic analyses, P2/P3 retained selected FP64 operations with metric-dependent rankings; P4 showed storage and trajectory sensitivity; and P5 achieved close corrected budget closure with a distinct depth response. P6-Q64 supplied the internal FP64 numerical reference. The depth-conditioned and local-difference results help identify the locations and depth bands that require attention when choosing a configuration.

For transfer to another catchment, terrain, drainage network, rainfall pathway and hardware should be included in the precision assessment. The fixed node-area proxy should also be replaced or assessed against local inlet geometry when inlet-scale performance is important. Reference [22] reports hydrograph NSE values of 0.9836–0.9969 and inundation-

trace coincidences of 81.1% for RoG and 78.3% for RoS. These measures assess flow and inundation extent, whereas precision RMSE quantifies differences between simulated depths. A direct comparison of their magnitudes would require corresponding model–observation depth errors.

5. Conclusions

We examined precision effects in the surface solver and sewer–surface exchange through hydraulic comparisons and a separate repeated performance experiment. The hydraulic analyses covered depth, inundation extent, manhole exchange, surface budgets and targeted numerical diagnostics. The one-hour performance experiment compared component and whole-run timing under common execution settings. The main findings are as follows.

First, the surface-flow precision configuration with the smallest difference relative to the P6-Q64 reference depended on the evaluation metric and rainfall-input method. In the six-hour hydraulic experiments, the maximum corrected 2D surface-domain accounting residual, normalized by exchange volume, was 9.52×10^{-3}%, but the relative ordering of the water-depth RMSE and manhole-level exchange-volume WAPE differed. Therefore, selecting a surface-flow precision configuration requires the joint consideration of water-depth differences, inundation extent, exchange volume, and the water balance of the 2D surface domain.

Second, the separate one-hour performance experiment (n = 10) gave P6-Q64/P1-Q64 median whole-run ratios of 1.883 for RoG and 1.205 for RoS, with exploratory 95% intervals of 1.876–1.887 and 1.199–1.212 obtained by repeated sampling.

Third, changing exchange arithmetic from Q64 to Q32 produced within-configuration depth and net-exchange differences in the hydraulic experiments. Final six-hour Q32/Q64 RMSEs were 0.108–0.345 mm for RoG and 0.716–4.571 mm for RoS, with net-exchange WAPE of 0.674–2.277%. The 48 h P1 RoS comparison reached 11.347 mm. In the separate performance experiment, the small, configuration-dependent runtime changes under Q32 were consistent with its limited FP32 arithmetic scope and the contribution of other model work to total time.

The P4 storage intervention reduced the final common-mask RMSE from 2.822 to 1.582 mm and improved 28 of 33 nonempty outputs, demonstrating state-storage sensitivity. The results identified P1-Q64 as an efficient candidate for the tested catchment. Precision selection should consider local depth, exchange volume, surface-budget closure and runtime under the intended rainfall, output workload and catchment conditions.

Author Contributions: Conceptualization, S.-B.S. and H.-J.K.; methodology, S.-B.S. and H.-J.K.; validation, S.-B.S.; writing—original draft, S.-B.S.; writing—review and editing, S.-B.S. and H.-J.K.; visualization, S.-B.S.; project administration, H.-J.K.; funding acquisition, H.-J.K. All authors have read and agreed to the published version of the manuscript.

Funding: The research for this paper was carried out under the KICT Research Program (Project no. 20260161–001, Development of Digital Urban Flood Control Technology for the Realization of Flood Safety City), funded by the Ministry of Science and ICT.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors gratefully acknowledge the support of the Korea Institute of Civil Engineering and Building Technology (KICT) and the Ministry of Science and ICT through the KICT Research Program.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

1D/2D	One-dimensional/two-dimensional
KICT-UF	KICT Urban Flood model
EPA-SWMM	U.S. Environmental Protection Agency Storm Water Management Model
CPU/GPU	Central processing unit/graphics processing unit
FP32/FP64	32-bit single precision/64-bit double precision
RoG	Rain-on-grid
RoS	Rain-on-subcatchment
HLLC	Harten–Lax–van Leer–Contact
CFL	Courant–Friedrichs–Lewy
RMSE	Root mean square error
IoU	Intersection over union
WAPE	Weighted absolute percentage error
I/O	Input/output

Appendix A Runtime Statistics

Tables A1 and A2 summarize ten measured runtimes per condition. Values are reported in seconds to three decimal places. SD is the sample standard deviation and IQR is Q3 − Q1.

Table A1. RoG wall-time statistics for the 60 min performance experiment (n = 10 per condition).

Configuration	Mean (s)	SD (s)	Median (s)	IQR (s)
P1-Q64	490.336	2.815	490.522	2.651
P1-Q32	491.030	1.951	490.995	2.888
P2-Q64	604.896	1.644	605.063	1.998
P2-Q32	605.809	2.305	605.762	2.313
P3-Q64	924.483	1.831	924.790	2.685
P3-Q32	926.124	4.708	925.526	3.430
P4-Q64	745.251	2.103	745.073	2.608
P4-Q32	745.441	3.798	745.590	3.697
P5-Q64	780.779	2.460	780.940	2.466
P5-Q32	780.888	2.874	780.861	3.457
P6-Q64	922.769	2.316	923.628	3.215
P6-Q32	922.964	2.173	922.594	2.681

Table A2. RoS wall-time statistics for the 60 min performance experiment (n = 10 per condition).

Configuration	Mean (s)	SD (s)	Median (s)	IQR (s)
P1-Q64	156.839	1.094	156.916	2.046
P1-Q32	157.284	0.888	157.553	1.253
P2-Q64	167.084	0.685	166.950	0.901
P2-Q32	167.975	1.161	167.845	1.606
P3-Q64	185.869	0.734	186.058	0.757
P3-Q32	186.191	1.097	186.142	1.782

Table A2. *Cont.*

Configuration	Mean (s)	SD (s)	Median (s)	IQR (s)
P4-Q64	172.887	0.929	172.894	1.288
P4-Q32	173.321	0.548	173.459	0.846
P5-Q64	181.955	1.212	181.833	2.106
P5-Q32	182.232	1.546	181.701	1.072
P6-Q64	190.651	4.026	189.124	1.756
P6-Q32	191.233	3.542	189.787	2.066

References

1. Leandro, J.; Chen, A.S.; Djordjević, S.; Savić, D.A. Comparison of 1D/1D and 1D/2D coupled (sewer/surface) hydraulic models for urban flood simulation. *J. Hydraul. Eng.* **2009**, *135*, 495–504. [CrossRef]
2. Seyoum, S.D.; Vojinović, Z.; Price, R.K.; Weesakul, S. Coupled 1D and noninertia 2D flood inundation model for simulation of urban flooding. *J. Hydraul. Eng.* **2012**, *138*, 23–34. [CrossRef]
3. Fan, Y.; Ao, T.; Yu, H.; Huang, G.; Li, X. A coupled 1D–2D hydrodynamic model for urban flood inundation. *Adv. Meteorol.* **2017**, *2017*, 2819308. [CrossRef]
4. Guo, K.; Guan, M.; Yu, D. Urban surface water flood modelling—A comprehensive review of current models and future challenges. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 2843–2860. [CrossRef]
5. Brodtkorb, A.R.; Sætra, M.L.; Altinakar, M. Efficient shallow water simulations on GPUs: Implementation, visualization, verification, and validation. *Comput. Fluids* **2012**, *55*, 1–12. [CrossRef]
6. Smith, L.S.; Liang, Q. Towards a generalised GPU/CPU shallow-flow modelling tool. *Comput. Fluids* **2013**, *88*, 334–343. [CrossRef]
7. Vacondio, R.; Dal Palù, A.; Mignosa, P. GPU-enhanced finite volume shallow water solver for fast flood simulations. *Environ. Model. Softw.* **2014**, *57*, 60–75. [CrossRef]
8. Noh, S.J.; Lee, J.-H.; Lee, S.; Kawaike, K.; Seo, D.-J. Hyper-resolution 1D–2D urban flood modelling using LiDAR data and hybrid parallelization. *Environ. Model. Softw.* **2018**, *103*, 131–145. [CrossRef]
9. Morales-Hernández, M.; Sharif, M.B.; Kalyanapu, A.; Ghafoor, S.K.; Dullo, T.T.; Gangrade, S.; Kao, S.-C.; Norman, M.R.; Evans, K.J. TRITON: A multi-GPU open source 2D hydrodynamic flood model. *Environ. Model. Softw.* **2021**, *141*, 105034. [CrossRef]
10. Buttinger-Kreuzhuber, A.; Konev, A.; Horváth, Z.; Cornel, D.; Schwerdorf, I.; Blöschl, G.; Waser, J. An integrated GPU-accelerated modeling framework for high-resolution simulations of rural and urban flash floods. *Environ. Model. Softw.* **2022**, *156*, 105480. [CrossRef]
11. Caviedes-Voullième, D.; Morales-Hernández, M.; Norman, M.R.; Özgen-Xian, I. SERGHEI (SERGHEI-SWE) v1.0: A performance-portable high-performance parallel-computing shallow-water solver for hydrology and environmental hydraulics. *Geosci. Model Dev.* **2023**, *16*, 977–1008. [CrossRef]
12. Kim, B.; Ryu, H.; Lee, S.; Lee, J.-H.; Noh, S.J. Portable multi-GPU urban flood modelling via MPI-OpenACC: Evaluating acceleration, scaling, and workload-per-GPU. *Environ. Model. Softw.* **2026**, *200*, 106957. [CrossRef]
13. Zhu, W.; Cao, Z.; Luo, P.; Wang, S.; Xv, C.; Ji, Y. An urban flood inundation model accelerated by the parallel acceleration technology. *Environ. Model. Softw.* **2025**, *189*, 106441. [CrossRef]
14. Klöwer, M.; Düben, P.D.; Palmer, T.N. Number formats, error mitigation, and scope for 16-bit arithmetics in weather and climate modeling analyzed with a shallow water model. *J. Adv. Model. Earth Syst.* **2020**, *12*, e2020MS002246. [CrossRef] [PubMed]
15. Li, H.; Hathaway, J.; Canestrelli, A. Watershed-scale hydrodynamic modeling of Hurricane Helene flooding in Southern Appalachians. *J. Hydrol. Reg. Stud.* **2026**, *64*, 103167. [CrossRef]
16. Ackmann, J.; Dueben, P.D.; Palmer, T.; Smolarkiewicz, P.K. Mixed-precision for linear solvers in global geophysical flows. *J. Adv. Model. Earth Syst.* **2022**, *14*, e2022MS003148. [CrossRef] [PubMed]
17. Lai, J.; Wang, L.; Yang, Y.; Wu, Q.; Chen, M. Evaluating effectiveness of round-off error compensation with three methods in shallow-water models. *Ocean Model.* **2025**, *198*, 102617. [CrossRef]
18. Li, D.; Hou, J.; Zhang, Y.; Guo, M.; Zhang, D. Influence of time step synchronization on urban rainfall-runoff simulation in a hybrid CPU/GPU 1D–2D coupled model. *Water Resour. Manag.* **2022**, *36*, 3417–3433. [CrossRef]
19. Fappiano, F.; Maurer, M.; Leitão, J.P. Coupling sewers to the surface: Systematic approaches to correcting data discrepancies for 1D–2D drainage modelling. *J. Hydrol.* **2024**, *645*, 132239. [CrossRef]
20. Kim, H.-J.; Kim, H.-J. Development of the KICT-UF model for urban flood analysis. *Water Future* **2018**, *51*, 30–35. (In Korean)
21. Sim, S.-B.; Kim, H.-J. An urban flood model development coupling the 1D and 2D model with fixed-time synchronization. *Water* **2024**, *16*, 2726. [CrossRef]

22. Sim, S.-B.; Kim, H.-J. Quantifying the impact of rainfall application (RoS vs. RoG) in urban flood modelling: Joint validation on laboratory and real basins. *KSCE J. Civ. Environ. Eng. Res.* **2025**, *45*, 711–719. [CrossRef]
23. Rossman, L.A.; Simon, M. *Storm Water Management Model User's Manual*; Version 5.2; EPA/600/R-22/030; U.S. Environmental Protection Agency: Washington, DC, USA, 2022.
24. Hou, J.; Simons, F.; Liang, Q.; Hinkelmann, R. An improved hydrostatic reconstruction method for shallow water model. *J. Hydraul. Res.* **2014**, *52*, 432–439. [CrossRef]
25. Rossman, L.A. *Storm Water Management Model Reference Manual Volume II—Hydraulics*; EPA/600/R-17/111; U.S. Environmental Protection Agency: Washington, DC, USA, 2017.
26. Brodtkorb, A.R.; Holm, H.H. Coastal ocean forecasting on the GPU using a two-dimensional finite-volume scheme. *Tellus A Dyn. Meteorol. Oceanogr.* **2021**, *73*, 1876341. [CrossRef]
27. Ministry of the Interior and Safety. *Guidelines for Setting and Operating Regional Disaster-Prevention Performance Targets Considering Future Climate Change*; Ministry of the Interior and Safety: Sejong, Republic of Korea, 2017. (In Korean)
28. Gyeongsangnam-do. *Jinjeoncheon River Master Plan (Partial Revision)*; Gyeongsangnam-do Notice No. 2024-46; Water Resources Division: Changwon, Republic of Korea, 2024. (In Korean)
29. Beck, Y.; Park, M.; Ahn, J. Comparative study on the regression equation of Huff time distribution for design rainfall. *J. Korea Water Resour. Assoc.* **2025**, *58*, 425–438. [CrossRef]
30. Croci, M.; Rosilho de Souza, G. Mixed-precision explicit stabilized Runge–Kutta methods for single- and multi-scale differential equations. *J. Comput. Phys.* **2022**, *464*, 111349. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.