

An integrated GPU-accelerated modeling framework for high-resolution simulations of rural and urban flash floods

Andreas Buttinger-Kreuzhuber^{a,b,*}, Artem Konev^a, Zsolt Horváth^{a,b}, Daniel Cornel^a, Ingo Schwerdtorf^c, Günter Blöschl^b, Jürgen Waser^a

^a VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH, Donau-City-Strasse 11, 1220 Vienna, Austria

^b Institute of Hydraulic Engineering and Water Resources Management, Vienna University of Technology, Karlsplatz 13/222, 1040 Vienna, Austria

^c Stadtentwässerungsbetriebe Köln, AöR, Ostmerheimer Straße 555, 51109 Cologne, Germany

ARTICLE INFO

Keywords:

Spatially distributed rainfall–runoff model
GPU acceleration
Dual drainage
Surface–sewer coupling
Flash flood hazard modeling

ABSTRACT

This paper presents an integrated modeling framework aiming at accurate predictions of flood hazard from heavy rainfalls. The accuracy of such predictions generally depends on the complexity and resolution of the employed model components. We propose an integration of complementary models in one framework that facilitates GPUs to improve accuracy and simulation time. The spatially distributed runoff model integrates surface flow routing based on the full shallow water equations, infiltration based on the Green–Ampt equation, and interception. In urban areas, the runoff model is coupled with the Storm Water Management Model (SWMM). The integrated model is validated and tested on laboratory, rural and urban scenarios with regards to accuracy and computational efficiency. The GPU acceleration yields speedups of 1000 times compared to a CPU implementation and enables the coupled simulation of flash floods at 1 m resolution for an urban area of 200 km² in realtime.

1. Introduction

Floods are increasing in many parts of the world due to climate and land use change (Chen et al., 2018; Blöschl et al., 2019) causing disproportionately high damage in urban regions (Jongman, 2018). To mitigate future flood damage, detailed models that assist in assessing the flood hazard spatially are crucial (Rosenzweig et al., 2021). In contrast to lumped models, spatially distributed models allow for an explicit representation of spatial variations and inhomogeneities in input data, such as topography, vegetation, soil characteristics, and urban features. Notwithstanding scale issues (Grayson and Blöschl, 2001), there is a lot of value in spatially distributed high-resolution modeling for management purposes. However, higher resolutions lead to slower simulations. In addition, ensemble simulations that quantify the uncertainty of the predictions and provide insights into the effects of parameter variation increase the computational burden even further. Thus, the challenge is to advance the capabilities of numerical modeling while balancing simulation performance and model accuracy.

To accurately represent the topography, the resolution of the simulation grid should be chosen accordingly. As a rule of thumb, terrain features should be covered by at least 3 cells to be represented explicitly (Gallegos et al., 2009; Fewtrell et al., 2011; Horváth et al., 2020). For urban areas, a resolution of 2 m or less is considered

necessary to accurately represent buildings, curbstones, and other features (Fewtrell et al., 2011; Dottori et al., 2013; Xing et al., 2018). Moreover, simulations at submeter resolutions are useful for assessing the benefits of small-scale alteration of street topography for flood risk management purposes at low costs (de Almeida et al., 2016). The simulation grid is typically either a structured Cartesian grid (Costabile et al., 2013; Horváth et al., 2015; Buttinger-Kreuzhuber et al., 2019), or an unstructured triangular mesh (Bermúdez et al., 1998; Hou et al., 2015; Fernández-Pato et al., 2016). Unstructured meshes are able to incorporate complex geometries, however they require time-consuming mesh generation. In contrast, structured grids lack the pre-processing step at the expense of poor resolution of topographic features not aligned with the grid. Thus, they usually require a higher resolution in comparison to unstructured meshes to enable the same level of accuracy.

The shallow water equations (SWEs) are typically used for describing surface flow. Due to the Courant–Friedrichs–Lewy (CFL) condition, a high spatial resolution requires a fine temporal discretization. Therefore, the total amount of computational work increases and in turn slows down simulation runs. To accelerate the simulations, one possibility is to simplify the shallow water model, for example, using diffusive

* Corresponding author at: VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH, Donau-City-Strasse 11, 1220 Vienna, Austria.
E-mail address: buttinger@vrvis.at (A. Buttinger-Kreuzhuber).

wave (sometimes also called zero-inertia) or kinematic wave approximations (Neal et al., 2012; Le et al., 2015; Fry and Maxwell, 2018; Yang et al., 2020). For urban regions, the full or dynamic SWEs in combination with shock capturing schemes are able to reproduce observed hydraulic behavior and velocities more accurately than simplified models (Kvočka et al., 2015; Costabile et al., 2020). Cozzolino et al. (2019) conclude that the preferred model for floodplain simulations should be the full 2D SWEs as simplified models often suffer from a poor representation of receding flows and bed discontinuities. Still, issues such as wetting and drying over complex terrain pose a numerical challenge and constitute an active area of research (Chen and Noelle, 2017; Xia et al., 2017; Buttinger-Kreuzhuber et al., 2019). If not treated properly, numerical instabilities occur and lead to slow simulations. The full SWE offer a model to simulate both complex open channel hydrodynamics and overland flow processes (Costabile et al., 2013; Fernández-Pato et al., 2016), in particular at high resolutions (Caviedes-Voullième et al., 2020). Recent studies (Costabile et al., 2017; Aricò and Nasello, 2018; Caviedes-Voullième et al., 2020) point out that solvers for the full SWEs might in fact require less computational time than their zero-inertia counterparts. Popular scheme choices for the full SWEs include discontinuous-Galerkin (DG) (Kesserwani et al., 2008; Vater et al., 2017; Ayog et al., 2021) and finite volume (FV) (Audusse et al., 2004; Horváth et al., 2015; Hou et al., 2014; Buttinger-Kreuzhuber et al., 2019; Dong and Li, 2021) methods. Second- or higher-order FV schemes are more accurate than their first-order counterparts, which are prone to numerical diffusion (Audusse and Bristeau, 2005; Noelle et al., 2006, 2007; Li et al., 2014; Navas-Montilla and Murillo, 2015; Hou et al., 2015). Most second-order schemes follow a monotonic upstream-centered scheme for conservation laws (MUSCL) approach (Van Leer, 1979), resulting in a shock-capturing scheme with reduced numerical diffusion yet without unphysical oscillations. However, compared to first-order schemes, the higher accuracy of second-order schemes comes at the price of higher runtimes due to a reduced CFL constant and second-order time integration. Thus, in large-scale flood modeling first-order schemes are still commonly used (Xia et al., 2019; Echeverri-bar et al., 2019; Morales-Hernández et al., 2021). In complex flash flood scenarios, it is not immediately clear whether the application of a second-order scheme pays off in terms of accuracy and required computational work.

A promising way to achieve computational speedups is the execution in a massively parallel fashion on supercomputers (Noh et al., 2018; Kuffour et al., 2020) or on graphics processing units (GPUs) (Las-tra et al., 2009; Brodtkorb et al., 2012; Vacondio et al., 2014; Lacasta et al., 2015; Le et al., 2015; Horváth et al., 2016; Xing et al., 2018; Xia et al., 2019; Morales-Hernández et al., 2021). To achieve a high computational performance on GPUs, regular grids are a convenient choice due to the structured arrangement of cores (Morales-Hernández et al., 2020). Unstructured meshes should be reordered for efficient memory access patterns allowing for coalescent transactions (Lacasta et al., 2015). Cutting-edge flash flood models are on the verge of handling resolutions of 5 m for large regions of up to 2500 km², or, 100 million cells (Xia et al., 2019). Traditionally, cities are split into multiple smaller simulation regions that tend to underestimate inundation (Xing et al., 2018). Thus, high-resolution simulations at large scales, e.g. spanning entire cities, are needed.

A variety of infiltration models exist, including the empirical Soil Conservation Service (SCS) curve number method (Chow et al., 1988; Aureli et al., 2020), the empirical Horton model (Fernández-Pato et al., 2016; Fernández-Pato and García-Navarro, 2018), the semi-empirical Green-Ampt model (Fiedler and Ramirez, 2000; Simons et al., 2013; Delestre et al., 2017; Fernández-Pato et al., 2016), and more complex models such as Richards equation (Maxwell, 2013; Le et al., 2015; Kuffour et al., 2020) although capturing macropore flow (Zehe et al., 2007) remains a challenge. The model's ability to capture the local effects of green infrastructure (GI), such as green roofs, rain

gardens, or bioswales, is important in urban flood resilience planning (Berland et al., 2017; Fry and Maxwell, 2018; Rosenzweig et al., 2021). For urban flood hazard modeling, the flow in sewer systems and its interaction with the overland flow may be relevant. The Storm Water Management Model (SWMM) is an established tool for routing stormwater in sewer systems. It is developed by the Environmental Protection Agency (EPA) as an open-source software package (Rossman, 2017). A widely used approach to bidirectionally couple urban drainage networks to overland flow are dual drainage models (Leandro and Martins, 2016; Yang et al., 2020; Li et al., 2020; Rosenzweig et al., 2021). The interaction terms are commonly based on the water level differences between the sewer nodes and the surface water (Djordjević et al., 2005; Chen et al., 2016; Rubinato et al., 2017; Fernández-Pato and García-Navarro, 2018). Previous coupled models featuring a dynamic sewer network simulation, e.g. Leandro and Martins (2016), Fernández-Pato and García-Navarro (2018), Noh et al. (2018), Yang et al. (2020), were run on central processing units (CPUs) and not exploiting recent leaps in model acceleration facilitated by GPUs. As the coupled models runtimes are typically governed by the surface simulation (Noh et al., 2018), fast runoff models are crucial.

In this paper, we present a coupled modeling framework for fast simulations in urban and rural settings. The framework includes several components considered relevant in rainfall-runoff modeling and flash flood hazard assessment, that is, spatially distributed interception and infiltration, an accurate representation of overland flow, and subsurface flow in sewer networks. We go beyond current modeling practice by using both a spatially distributed rainfall-runoff model and a fully bidirectional coupling of the sewer network accounting for drains and overflows at large scales and very high resolutions. We propose a novel, hybrid coupling approach of a GPU-accelerated runoff simulation with an established CPU sewer network simulation. We validate and test the framework in laboratory, rural and urban scenarios. We answer the question whether first-order or second-order schemes in the surface flow discretization of the full 2D SWEs should be favored in terms of the workload-accuracy tradeoff. Moreover, we highlight the influence of resolution and of the individual model components. Finally, we address the extent of computational acceleration on a modern GPU for high-resolution simulations of entire cities.

2. Methods

2.1. Surface flow model

The full shallow water equations (SWEs) are used to describe the surface flow and may be written in vector form as

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) + \partial_y \mathbf{G}(\mathbf{U}) = \mathbf{S}_b(\mathbf{U}, b) + \mathbf{S}_f(\mathbf{U}), \quad (1)$$

where $\mathbf{U} = [h, hu, hv]^T$ is the vector of conserved variables, h represents the water height, hu is the discharge along the x -axis, and hv is the discharge along the y -axis. \mathbf{F} and \mathbf{G} are the flux functions,

$$\mathbf{F} = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ huv \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} hv \\ huv \\ hv^2 + \frac{1}{2}gh^2 \end{bmatrix}. \quad (2)$$

The bed slope term \mathbf{S}_b ,

$$\mathbf{S}_b = \begin{bmatrix} 0 \\ -gh\partial_x b \\ -gh\partial_y b \end{bmatrix}, \quad (3)$$

models the fluid's acceleration due to the gravitational forces. The friction term \mathbf{S}_f ,

$$\mathbf{S}_f = \begin{bmatrix} 0 \\ -gn^2h^{-1/3}u\sqrt{u^2+v^2} \\ -gn^2h^{-1/3}v\sqrt{u^2+v^2} \end{bmatrix}, \quad (4)$$

accounts for the bed friction. Here, u and v are the average flow velocities in x and y directions, respectively, g is the gravitational constant, b is the bed level (assumed to be time-independent), and n is the Manning friction coefficient.

To integrate the interception and infiltration processes of the runoff model and the sewer model with the surface flow, coupling terms S_r and S_s , respectively, are added on the right hand side of Eq. (1). The coupling term for the sewers S_s accounts for the specific sewer exchange discharge q_e (m/s). The source terms S_r and S_s are given by

$$S_r = \begin{bmatrix} r_e(t) \\ 0 \\ 0 \end{bmatrix}, \text{ and } S_s = \begin{bmatrix} q_e(\mathbf{U}) \\ 0 \\ 0 \end{bmatrix}. \quad (5)$$

The runoff rate r_e (m/s) is the difference between the effective precipitation rate and the effective infiltration rate. Its precise application is defined in Section 2.4.

2.2. Spatio-temporal discretization of the surface flow

For the spatial discretization of the SWEs, the finite volume method (FVM) was chosen on a uniform Cartesian grid. Although all runoff components can also be discretized on unstructured grids, structured grids require less pre-processing steps and enable fast memory access patterns on the GPU (Morales-Hernández et al., 2020). Moreover there is no need to store or optimize mesh connectivity. The FVM discretizes the conserved variables \mathbf{U} as cell averages yielding a system of ordinary differential equations for the cell averages $\mathbf{U}_{j,k}(t)$.

For the simulation of the overland flow, we employ either a first-order accurate or a second-order accurate scheme. The first-order scheme of Chen and Noelle (2017), to which we refer as CN scheme, enables a better handling of flow states across bed discontinuities than the original hydrostatic reconstruction (HR) scheme proposed by Audusse et al. (2004). A second-order accurate extension of the first-order CN scheme is presented in Buttinger-Kreuzhuber et al. (2019), to which we refer as BH scheme. The second-order accuracy in space is achieved through a minmod limiter. The minmod parameter is set to 1 in order to ensure robust and fast simulations (Horváth et al., 2020). At wet-dry boundaries only the velocities are set to zero below a cut-off water depth threshold. In the simulations this threshold is set to 0.1 mm. The surface flow is discretized in time by the explicit Euler's method for the first-order CN scheme with a Courant–Friedrichs–Lewy (CFL) constant of 0.5 to guarantee numerical stability and non-negativity of the water depths. The second-order BH scheme is integrated in time with Heun's method and the CFL constant is set to 0.25. The CFL condition restricts the time step Δt ,

$$\Delta t \leq \text{CFL} \frac{\Delta x}{\sigma}, \quad (6)$$

where Δx is the uniform grid resolution and σ is the maximum of the absolute values of the numerical wave speeds. The numerical wave speeds at the cell interfaces are computed from the eigenvalues of the Jacobian of the flux functions \mathbf{F} and \mathbf{G} (Buttinger-Kreuzhuber et al., 2019). Both schemes are mass conserving and preserve lake-at-rest steady states. The friction source term S_f is evaluated in a semi-implicit manner by splitting it into a coefficient-wise product of an implicitly evaluated state and an explicitly evaluated friction term \tilde{S}_f (Brodtkorb et al., 2012; Buttinger-Kreuzhuber et al., 2019).

2.3. Runoff model

The spatially distributed runoff simulation integrates the surface flow routing component with an interception and an infiltration component determining the effective surface runoff. First, part of the rain is stored in the canopy of vegetation through interception. Second, infiltration occurs as surface water percolates into permeable soils. The remaining water effectively materializing during a rain event runs off the surface as overland flow.

The rainfall intensity is given by a time- and space-dependent precipitation rate p . The integrated interception component reduces the effective precipitation and accounts for micro-topographic depressions and losses due to vegetation. The interception storage capacity I_S is typically roughly around 1 mm (Robinson and Ward, 2017), except for rural forests, where values as high as 5 mm were found (Xiao et al., 1998). The cumulative interception $I(t)$ until time t is modeled with a constant non-negative rate i until a predefined storage capacity I_S is reached. Thus, the spatially distributed effective precipitation rate is given by

$$p_e(t) = \begin{cases} p(t) - i & \text{if } I(t) < I_S, \\ p(t) & \text{else.} \end{cases} \quad (7)$$

The infiltration process is modeled by the Green–Ampt equation. The cumulative infiltration F up to time t is

$$\int_0^{F(t)} \frac{F}{F + (\psi + h) \Delta\theta} dF = \int_0^t K_s dt, \quad (8)$$

where K_s is the saturated hydraulic conductivity. The difference $\Delta\theta$ between the initial water content and the saturated water content of the soil is usually called effective porosity. The suction head ψ represents the capillary attraction of the water towards the soil voids. Solving Eq. (8) for the infiltration rate f , the time derivative of F , we obtain

$$f(t) = \frac{dF}{dt} = K_s \left[\frac{(\psi + h) \Delta\theta}{F(t)} + 1 \right]. \quad (9)$$

The proposed dynamic infiltration model accounts for the surface water pressure via the surface water height h . A shortcoming of the presented Green–Ampt model is the inability to account for multi-layered soils, limited storage capacity, soil water redistribution in dry phases and macropores. Extensions to overcome these limitations have been proposed (Corradini et al., 2000; Gowdsh and Muñoz Carpena, 2009; Mohammadzadeh-Habili and Heidarpour, 2015; Leandro et al., 2016).

2.4. Temporal discretization of the runoff model

The Green–Ampt (GA) model is discretized in time with the implicit Euler method solving Eq. (9) at every cell for every time step. The infiltration depths F at time step t^{n+1} is given by

$$F^{n+1} = \frac{1}{2} (F^n + K_s \Delta t) + \frac{1}{2} \sqrt{(F^n + K_s \Delta t)^2 + 4 K_s \Delta t \Delta\theta (h_0^n + \Psi)}. \quad (10)$$

We note that, even though the infiltration rate is undefined for $F = 0$, the implicit Euler method yields a well-defined infiltration depth close to zero, in contrast to the explicit Euler method. If the infiltration depth increment ΔF , defined by $\Delta F^n = F^{n+1} - F^n$, exceeds the available surface water depth, it is restricted to ensure a nonnegative surface water depth. The effective infiltration rate is then given by

$$f_e^n = \frac{1}{\Delta t^n} \min(\Delta F^n, h^n), \quad (11)$$

where Δt^n is the CFL-limited timestep of the overland flow for timestep n . For the runoff model, it is enough to perform a simple integration for both the effective precipitation and infiltration rate. We combine the precipitation and infiltration increment into a single effective runoff increment Δr_e^n ,

$$\Delta r_e^n = \Delta t^n (p_e^n - f_e^n). \quad (12)$$

The overland flow and runoff models are tightly coupled, every step in the surface flow simulation is synchronized with the computation and application of the runoff update.

2.5. GPU implementation of the runoff model

The spatial discretization of the surface flow with the FVM enables straightforward parallelization on structured grids, as only neighboring cells need to be considered when computing the next time step. The GPU implementation of the runoff model uses the CUDA platform of NVIDIA. In CUDA, parallel tasks are organized into thread blocks and typically each parallel task (each so-called thread) is associated with one cell. A block size of 16 by 16 cells ensures a high utilization of the GPU (Horváth et al., 2016). The number and positions of the neighboring cells required to compute the fluxes of a particular cell are called the computation stencil. For the first-order CN scheme, only the direct neighbors are accessed. For the second-order BH scheme, the computation of the minmod-limited gradient requires a neighborhood of two cells to be accessed in the four axis-aligned directions. Halo cells are required to exchange the data between adjacent blocks. Effectively, fluxes are computed only for the inner block of 14 by 14 cells for the first-order CN scheme or for 12 by 12 cells for the second-order BH scheme due to the different computational stencils and hence the different number of halo cells required. The GPU implementation handles data dependencies with on-chip memory for fast processing of the computational stencil, e.g. shuffles introduced with the Kepler microarchitecture. For implementation details, the reader is referred to Horváth et al. (2016).

GPU implementations of the SWEs operate either on single-precision or on double-precision floating-point variables. For single-precision state variables the memory burden is lower and floating-point operations are faster, as GPUs were originally optimized for single-precision computations and therefore speed-ups of up to eight times over double-precision are reported (Morales-Hernández et al., 2020). To alleviate the computational burden in the proposed runoff model, both the surface state and the infiltration state are stored in single-precision floating-point variables. In previous publications (Buttinger-Kreuzhuber et al., 2019; Horváth et al., 2020), we verified that the employed surface flow model relying on single-precision state variables captures lake-at-rest steady states, displays the expected convergence order and is capable of handling wet/dry transitions. Still, truncation errors might accumulate over time, which may have an effect on the mass balance (Liang and Smith, 2015; Dazzi et al., 2021). Therefore, we verify that the runoff model's mass balance is within machine precision in the presented test cases.

On a modern GPU with 24 GB of video memory the domain size in our computational model is limited to around 175 million active (wet) cells for the second-order BH scheme with dynamic runoff and single-precision floating-point variables. For the first-order CN scheme, the domain size is limited to around 225 million active cells as the first-order time integration does not require the storage of an intermediate state.

2.6. Sewer network model

In the Storm Water Management Model (SWMM), a sewer network is represented by a set of nodes connected by links (Rossman, 2017). Links transmit pipe discharges Q from one node to another. A so-called node assembly consists of the node and all connected half-links, these are the halves of the link that are connected to the node. At each node assembly, the change in the hydraulic head is modeled by the continuity equation. The pipe flow is governed by the transient 1D SWE, i.e.

$$\partial_t Q = 2U \partial_x A + U^2 \partial_x A - gA \partial_x H - gAS_f, \quad (13)$$

where U is the pipe flow velocity, A is the cross-sectional flow area, H is the hydraulic head, and S_f is the friction slope in the pipes. In the SWMM, Eq. (13) is solved with a finite difference scheme. Thus, at each link, momentum and continuity are conserved, in contrast to the nodes, where only continuity is conserved. The continuous state variables

in the time differences are approximated with their average values over the conduit length. SWMM 5.1 uses an implicit backwards Euler method for the time discretization, which is solved iteratively with Picard's method. We use the Preissmann slot model, implemented in the latest version of SWMM 5.1.013, which is integrated in our coupled model setup. Unfortunately, SWMM does not exactly preserve water volumes. Continuity errors are typically within a few percent (Rossman, 2006) and may be reduced by artificially introducing a finer spatial discretization (Pachaly et al., 2020). SWMM 5.1 is written in C/C++ and is easily incorporated into existing C++ software such as the proposed modeling framework.

2.7. Bidirectional surface–sewer coupling

The surface–sewer discharge exchange term depends on the water level at the surface w , the hydraulic head at the manhole H and the bed surface elevation b . In the following, A_m and D_m are the manhole's area and diameter, respectively, d_i is the distance between the surface and the invert level of the pipes entering the node. The invert level refers to the lowest elevation admitting water flow. Following Djordjević et al. (2005), Chen et al. (2016), Rubinato et al. (2017), Fernández-Pato and García-Navarro (2018), we distinguish between four cases:

- (1) inflow into a non-pressurized node,
- (2) inflow into a pressurized node, where the surface flow depth is small when compared to the node width,
- (3) inflow into a pressurized node, where the surface flow depth is large when compared to the node width,
- (4) outflow onto the floodplain.

If the head in the pipe network is lower than the surface elevation, i.e. $H < d_i$, the discharge exchange Q_e (m^3/s) is given by the free weir equation

$$Q_e = -\frac{2}{3} c_{d,w} \pi D_m (2g)^{1/2} h^{3/2}. \quad (14)$$

If the head in the pipe network exceeds the surface elevation, i.e. $H > d_i$, the discharge exchange is either given by the submerged weir equation,

$$Q_e = -c_{d,sw} \pi D_m (2g)^{1/2} h (h + d_i - H)^{1/2}, \quad (15)$$

as long as $h < A_m / \pi D_m$. If $h \geq A_m / \pi D_m$ the node is considered fully submerged, the submerged orifice equation

$$Q_e = -c_{d,o} A_m (2g)^{1/2} (h + d_i - H)^{1/2}, \quad (16)$$

is considered a more appropriate description. For example, for circular manholes, the orifice equation is applied if $h > D_m/4$. The discharge coefficients for the free weir, the submerged weir, and the orifice equations are set to $c_{d,w} = 0.56$, $c_{d,sw} = 0.11$, and $c_{d,o} = 0.2$, respectively (Rubinato et al., 2017).

If the head in the pipe system exceeds the water level of the surface flow, an orifice equation is used (Djordjević et al., 2005). Assuming that the surface velocity is negligible, the discharge exchange is given by

$$Q_e = c_{d,o} A_m (2g)^{1/2} (H - h - d_i)^{1/2}. \quad (17)$$

This equation also holds for dry surface cells, i.e. when $h = 0$. With these four cases, all exchange flow conditions are properly handled. A negative exchange discharge Q_e indicates flow into the sewer network from the surface, a positive value indicates sewer overflow.

Sewer overflow also occurs if water flow from the roofs of the surrounding buildings exceeds the sewer inflow capacity. Consequently, water spills over at this node. In this case, the roof water is directly added to the sewer overflow as excess discharges and not reduced by the surface–sewer exchange equations, Eqs. (14)–(17).

The surface–sewer coupling takes place only at the cells where manholes and inlets are connected to the surface. To this end, the geometry of the manholes and inlets is rasterized on the simulation

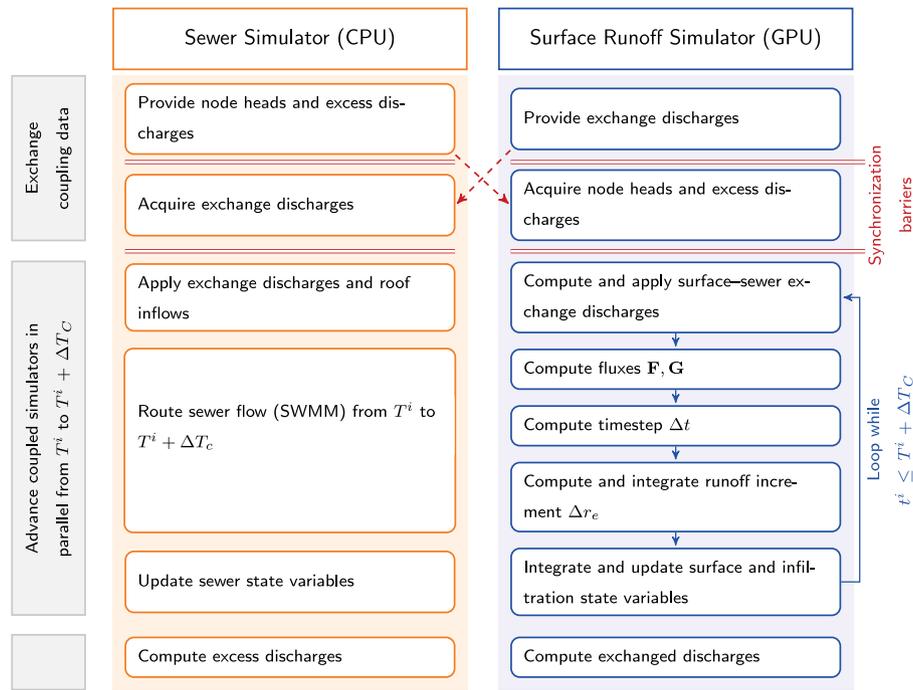


Fig. 1. In a single time step of the coupled surface-sewer model, three major substeps (indicated in gray at the left) with associated routines are performed in each simulator. The sewer and the surface simulator are executed in parallel, but they need to be synchronized for data exchange. These synchronization barriers (indicated with double red lines) are needed for data exchange between the two simulators (indicated with the dashed red arrows). With this less tightly coupled approach, each simulator is able to advance independently of the other.

grid. Effectively, at each cell the perimeters and areas of all intersecting manholes and inlets are collected. Furthermore, we also collect the contributions of each cell to each sewer node. We are able to include cases where multiple manholes intersect the same cell, as well as cases where multiple cells contribute to the same node. For each rasterized cell, we keep track of the corresponding node head by averaging over all nodes connected to the specified cell. The sewer node exchange discharge Eqs. (14)–(17) are solved on a per-cell basis, where we account for the relative contributions,

$$Q_e = \sum A_{j,k} q_{e,j,k}. \quad (18)$$

The specific sewer exchange term for the surface model is limited by the water availability in the case of sewer inflow, i.e.,

$$-q_{e,j,k} = \min \left(\frac{h_{j,k}}{\Delta t}, -q_{e,j,k} \right). \quad (19)$$

In time, the sewer network model is interleaved with the surface and runoff model with an a priori defined coupling timestep ΔT_C . At multiples of the coupling timestep, information between the runoff model and the sewer model is exchanged. The coupling time step size also represents an upper limit to the time step sizes of the individual simulators. If not stated otherwise, it is set to 1 s. As a rule of thumb, the coupling time step size should be roughly of the same order as the time step sizes of the individual solvers. A time step of the coupled model is illustrated in Fig. 1 and subdivided into the following steps:

- (1) Exchange sewer-surface coupling data, i.e. provide node heads and excess discharges to the surface flow simulation, and provide exchange discharges Q_e to the sewer model.
- (2) Advance the simulators in parallel from time T^i to $T^i + \Delta T_C$.
- (3) Compute the exchanged discharges in this coupling time step in the respective simulator.

Each simulator performs multiple routines at each of these steps (see Fig. 1). When coupling data need to be exchanged, the simulators are required to wait for the other simulator at synchronization barriers. In

the proposed C++ framework, the synchronization barriers are implemented with the functionality offered by `std::thread` and `boost::barrier`. Synchronization barriers are set only at the beginning of a coupled simulation time step, but not in the individual simulator’s advance methods. The loop for the time steps in the advance step is executed independently of the other simulator. To advance from T^i to $T^i + \Delta T_C$, each simulator only needs the minimal amount of time steps required for its own numerical stability.

For the sewer simulation, the exchange discharges are obtained from the surface simulation and applied to the sewer network. Additional inflows from external sources, e.g. roofs, are applied and compared with the sewer network’s inflow capacities. These excess discharges contribute to node overflow. SWMM routes the water flow in the sewer network. Then, the sewer network state (node depths and heads, link discharges and volumes, inflow volumes, and excess volumes) is updated. The excess volumes accumulated during one coupling time step in the sewer network simulation are provided as excess discharges to the surface runoff simulation in the next coupling time step. More precisely, we compute the excess discharges prescribed in the surface simulation at time step T_{i+1} from the excess volumes accumulated in the sewer network simulation from T_i to $T_{i+1} = T_i + \Delta T_C$. As a consequence of the parallel execution, during the coupling time step the state of the surface variables remains fixed for the sewer simulation and vice versa.

For the simulation of the surface flow and runoff, node heads from the sewer simulation are necessary for the computation of the surface-sewer exchange discharges. Once acquired, the surface runoff simulation advances independently of the sewer simulation. The routines in the surface runoff advance step correspond to typical routines in FVMs for overland flow simulation. These routines are performed on the GPU and looped until the next coupling time step $T^i + \Delta T_C$ is reached. The surface state variables (water depth and level, surface flow discharges) and the infiltration state variables are updated after timesteps Δt^n according to Section 2.4. After the loop, the applied exchange discharges during the coupling time step are computed and are provided to the sewer simulation in the next coupling time step.

The coupled simulation time steps are executed until the simulation end time is reached. As the sewer simulation and the surface runoff simulation are executed in parallel, the execution time of the coupled surface–sewer model is determined by the execution time of the slower simulator and not by the sum of the execution times as it would be the case for a sequential coupled simulation. Usually, the one-dimensional sewer network simulation is faster than the two-dimensional surface flow simulation (Noh et al., 2018).

3. Results and discussion

We demonstrate the capabilities of the coupled model on laboratory and real-world test cases. The scenarios include a small-scale rainfall–runoff plot experiment, a rural catchment, and a full-scale urban test case including sewer coupling. More specifically, we simulate the Thiès plot experiment (Section 3.1), the HOAL catchment at Petzenkirchen, Austria (Section 3.3), and the city of Cologne, Germany (Section 3.4). Furthermore, we validate the sewer–surface coupling approach on a laboratory experiment presented in Rubinato et al. (2017) in Section 3.2.

The numerical simulations were performed on a desktop PC equipped with 10 Intel i9-9820X cores at 3.3 GHz and 128 GB RAM. The GPU utilized for the test cases was an NVIDIA Titan RTX in Tesla Compute Cluster (TCC) mode. It features 4608 CUDA cores and has 24 GB memory. In the following, the term runtime describes the cumulative execution time of the simulation measured via wall clock timing. The runtime neither includes the initialization process, such as reading input data, nor postprocessing steps, such as writing results to the disk. However, the GPU runtime includes data transfer between the GPU and the CPU during simulation.

3.1. Thiès plot experiment

We validate the model with measurements performed at in Thiès, Senegal, by Tatard et al. (2008). The experiment was carried out on a $10 \times 4 \text{ m}^2$ plot. The plot has an average slope of 1%, and the resolution of the digital terrain model (DTM) is 0.1 m. Rainfall was simulated with a constant rate of 70 mm/h for a duration of 1 h on the sandy soil. In the reference data set of Mügler et al. (2011), measurements of mean flow velocities are available at 62 locations across the plot (Fig. 2a). Following Simons et al. (2013), Manning’s roughness coefficient was set to a constant value of $0.014 \text{ m}^{1/3}/\text{s}$ throughout the entire plot. We compare the results from the first-order CN and second-order BH scheme for the steady state after 1 h. The simulated water depths show a slightly clearer depiction of the flow paths in the second-order scheme, compare Fig. 2b–c. The simulated velocities are shown as arrows in Fig. 2b–c for the CN and BH scheme.

In Fig. 3a, we compare the simulated velocities with the measured velocities. Second-order schemes are computationally more involved than first-order schemes, but are supposed to yield superior results due to the improved accuracy. The root mean square error (RMSE) of the velocities is defined by

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_i (v_s^i - v_o^i)^2}, \quad (20)$$

where N is the total number of all observed velocities v_o^i . The RMSEs of the velocities are consistently lower for the second-order scheme for all resolutions from 0.05 to 0.25 m, as is shown in Fig. 3b, and the achieved RMSE of 0.026 m/s is in line to results in the literature (Tatard et al., 2008; Mügler et al., 2011; Simons et al., 2013; Caviedes-Voullième et al., 2020). However, we emphasize that a proper discretization of the source term is important for the simulation of runoff processes, even more so in the case of first-order accurate schemes. The superiority of the first-order CN scheme above the popular HR scheme (Audusse et al., 2004) is noticeable from the velocity errors in Fig. 3a and b. The simulated velocities of the HR scheme are consistently lower than

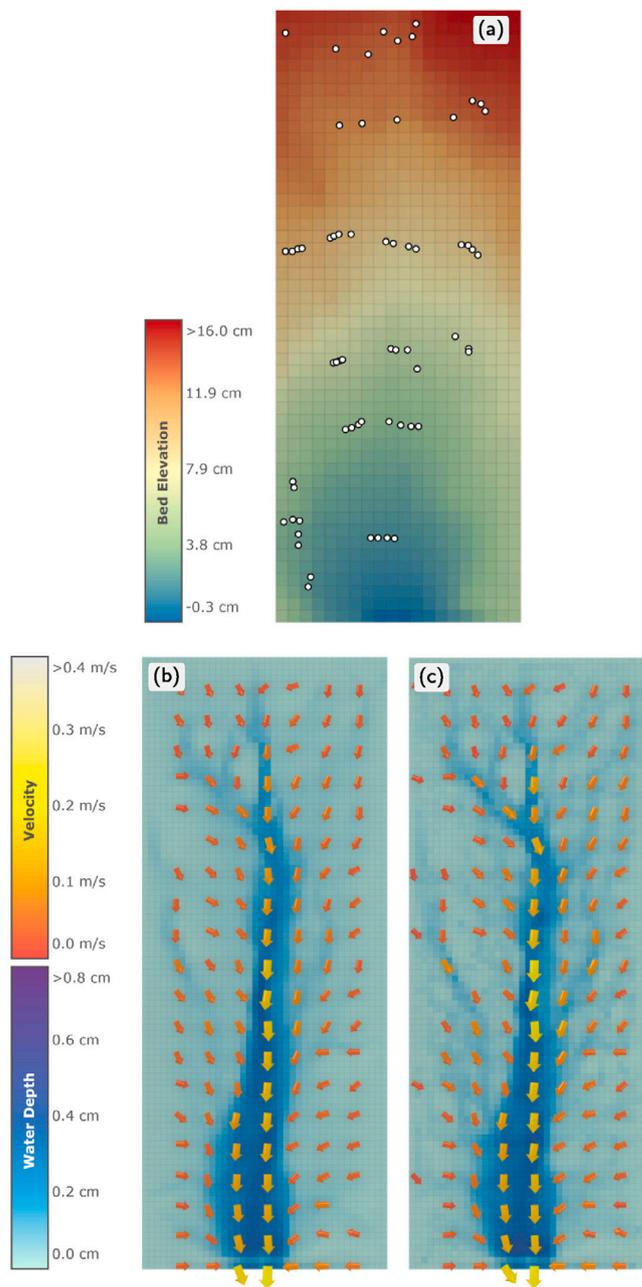


Fig. 2. Thiès experiment on a plot of $10 \times 4 \text{ m}^2$. (a) Terrain with velocity measurement locations (white dots). (b) Simulated water depths after 1 h and velocities (colored arrows) of the first-order accurate CN scheme and a resolution of 0.1 m. (c) Simulated water depths after 1 h and velocities (colored arrows) of the second-order accurate BH scheme and a resolution of 0.1 m. The flow patterns of the first-order and second-order scheme are comparable for this fine resolution.

for the CN scheme, as the HR scheme is not able to fully account for the bed slope in the case of shallow flow (Delestre et al., 2012). Switching to second-order accuracy in the HR scheme fixes this issue, albeit at the cost of a higher computational workload. The velocities of the second-order HR scheme are only slightly deviating from the BH scheme, therefore they were excluded from the plots in Fig. 3.

In line with numerical theory, the RMSE decreases with grid refinement in general as shown in Fig. 3b. However, the RMSEs of the second-order scheme do not exhibit such a clear trend as the RMSEs of the first-order scheme. For the second-order scheme, convergence with regard to mesh refinement is limited for cell sizes below 0.1 m due to the resolution of the underlying DTM (0.1 m). Typically, for smooth

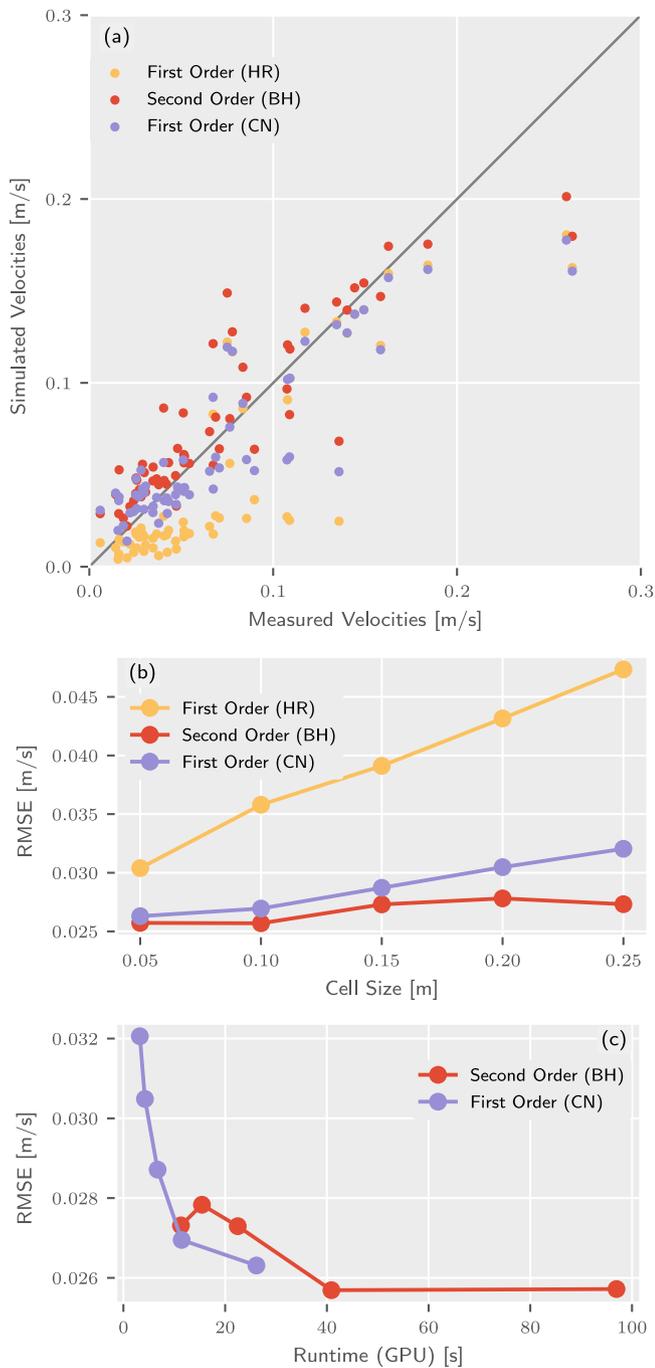


Fig. 3. Thiès. (a) Simulated velocities over measured velocities at steady-state conditions for a resolution of 0.1 m. The second-order scheme develops higher velocities than the first-order scheme. (b) Root mean square errors (RMSEs) over cell sizes. For the first-order scheme, the error decreases with resolution. For the second-order scheme the trend is not as pronounced. (c) RMSEs over GPU runtimes. For the same amount of computational time spent, the first-order CN scheme produces better results than the BH scheme.

solution surfaces, e.g. in river floods, the second-order accurate scheme is expected to be more efficient when considering the accuracy versus runtime tradeoff (Horváth et al., 2020).

When the corresponding runtimes of the CN and the BH schemes are compared in Fig. 3c, the first-order CN scheme produces better results for the same amount of computational time spent. In terms of the tradeoff between computational workload and accuracy, this suggests the use of finer grids together with first-order schemes for the

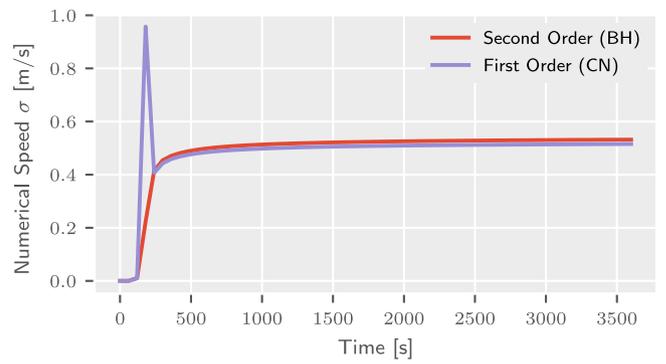


Fig. 4. Thiès. Maximum numerical speeds over time for the first-order accurate CN and the second-order accurate BH scheme. No high numerical speeds develop over time.

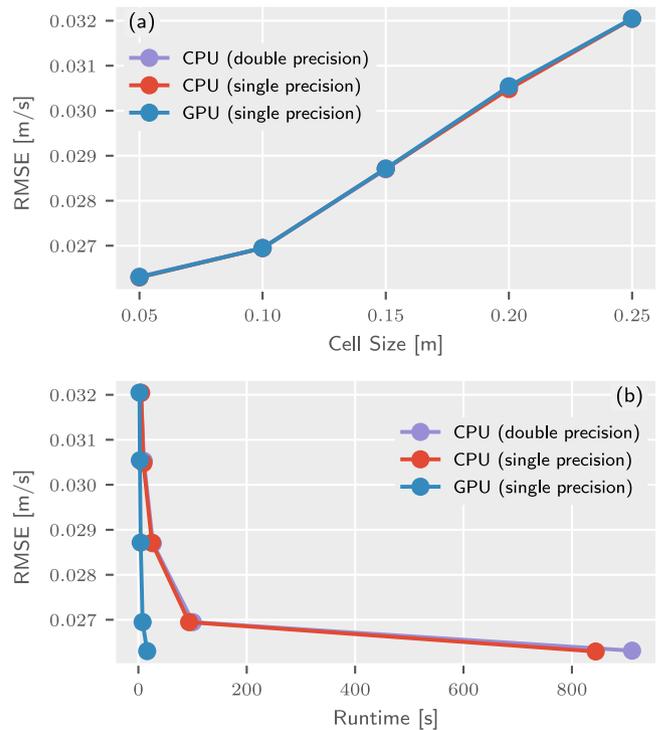


Fig. 5. Thiès. Comparison of a single-precision GPU, a single-precision CPU, and a double-precision CPU implementation of the CN scheme in terms of the velocity root mean square errors (RMSEs) over cell sizes (a) and over the corresponding runtimes (b). There are no significant differences in terms of accuracy.

surface flow in rainfall–runoff simulations. Summarizing, we conclude that fast first-order schemes, which properly resolve the source term, are sufficiently accurate for rainfall–runoff simulations.

Both the first-order CN and the second-order BH scheme are robust, as no unphysical high numerical speeds develop over time (Fig. 4). The time step Δt is inversely proportional to the maximum numerical speed σ by the CFL condition, see (6). Thus, if the maximum numerical speeds are small, large time steps are possible. There is a small spike at the beginning of the wetting, but its magnitude is reasonable. In Fig. 5, we compare the effect of using single-precision against double-precision floating-point numbers for the first-order CN scheme on the CPU and the GPU. In terms of the velocity RMSEs, the differences are noticeable for a cell size of 20 cm, but they are still very small (<0.1 mm/s). The mass balance error for the 10 cm runoff model is below 0.0001%. The rain volume is exactly 2.8 m^3 , 1.074432 m^3 are infiltrated, 1.697895 m^3 are flowing out at the open southern boundary during the simulation, and 0.027672 m^3 remain at the surface at the end of the simulation. In

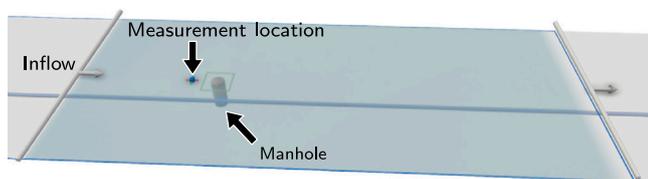


Fig. 6. Geometric setup of the surface-sewer coupling test case.

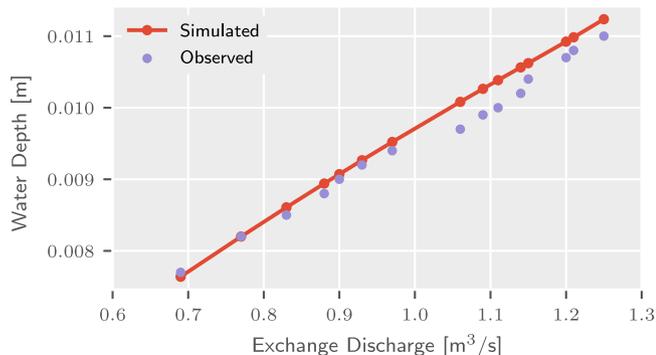


Fig. 7. Surface to sewer coupling, i.e. water is flowing from the surface into the sewer. Simulated (red) and observed water depths (violet) over exchange discharge for a range of prescribed surface flow discharges.

terms of the runtimes, the parallel single-precision GPU implementation is more than 30 times faster than the sequential single-precision CPU implementation for a resolution of 5 cm (Fig. 5b). In this case, due to the low number of cells in the domain, that is 16000 cells, the GPU cannot fully exploit its parallel capabilities. The modest speedup in this small-scale experiment is therefore not representative of large real-world test cases.

3.2. Surface-sewer coupling

In this section, we present simulation results to verify the surface-sewer coupling against experimental data provided by Rubinato et al. (2017). The setup consists of a flume and a single manhole connected to a pipe. The surface bed is 4 m wide and 8 m long with a slope of 1 m per 1 km. At the upper end, a hydrograph with a constant discharge of 11 L/s is specified. At the outlet, critical flow conditions are imposed. The manhole is located 2.5 m downstream of the inlet and has a diameter of 0.24 m. The invert level of the pipe is 0.478 m below the flume bed. Manning’s roughness coefficient is set to 0.009 s/m^{1/3} for both the pipe and the flume as both are PVC. In the experiment, the pipe pressure and the surface level was measured 0.34 m and 0.35 m away from the manhole (Fig. 6).

First, we tested steady state inflow from the surface into the sewer system for various prescribed surface discharges ranging from 5 L/s to 11 L/s. The simulation reaches a steady state after 300 s. The simulation is able to accurately reproduce the measured water depths and exchange flows (Fig. 7). With a mean absolute error (MAE) of 0.19 mm and relative differences ranging from 0.1% to 4.3%, the accuracy is excellent.

In the second test case, we simulated overflow from the sewer onto the wet flume. The surface inflow is fixed at 11 L/s and the pipe inflow ranges from 2.2 L/s to 7.6 L/s. Again, the simulated water depths agree well with the measured water depths with a MAE of 0.72 mm. The relative differences range from 3% to 4.8% in this case. Unfortunately, SWMM cannot extract the pressure head at an arbitrary location along the pipe, so instead we extracted the pressure head directly at the manhole. The pressure head at the node is supposed to be lower than

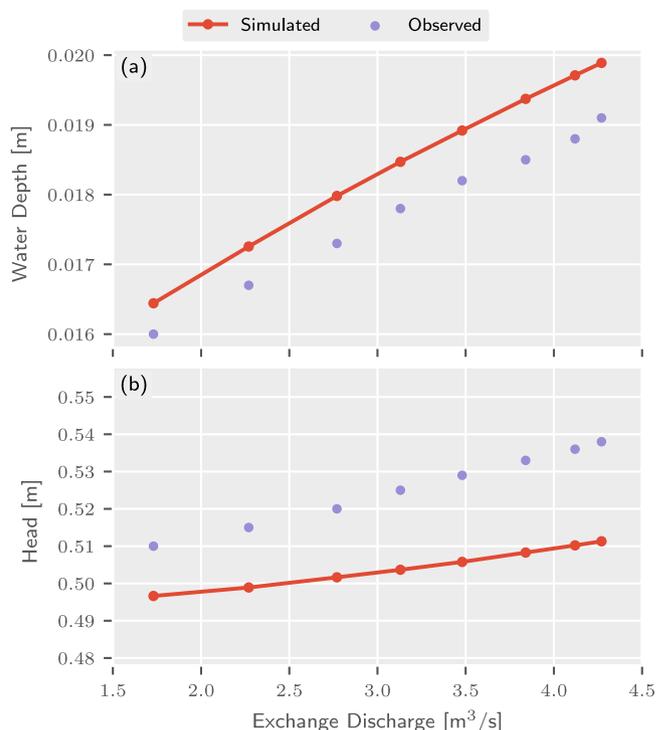


Fig. 8. Sewer to surface coupling, i.e. water is flowing from the sewer onto the surface. Simulated (red) and measured (violet) water depths (a) and pressure heads (b) over the exchange discharges. The water depths show good agreement. The simulated pressure head is extracted directly at the node and not 0.35 m away from the node as in the experiment thus resulting in lower simulated pipe pressure heads.

in the pipe, as energy is dissipated in the transition from the pipe into the manhole. This discrepancy is visible in Fig. 8b.

Overall, the coupled simulation correctly exchanges flows from the surface to the sewer system and vice versa as simulated and observed values agree. The small differences are comparable to results in the literature (Rubinato et al., 2017; Fernández-Pato and García-Navarro, 2018).

3.3. HOAL Petzenkirchen

This scenario analyzes a rainfall event in June 2013 in the Hydrological Open Air Laboratory (HOAL) catchment in Petzenkirchen, Lower Austria. The HOAL catchment is used to test hydrological hypotheses under natural conditions. The catchment is 0.66 km² in size and is mainly covered by arable land (87%) and grassland (10%) (Blöschl et al., 2016). A high resolution (0.5 m) DTM of 2012 was used. The topographic elevation of the catchment ranges from 255 to 325 m.a.s.l. Manning’s roughness coefficient is set to 0.1 s/m^{1/3} for the riparian forest, to 0.05 s/m^{1/3} for grassland and arable land, and to 0.03 s/m^{1/3} everywhere else.

The saturated conductivities range from 1 to 32 mm/h and are set according to literature values (Rawls et al., 1983; Carsel and Parrish, 1988; Smith et al., 2002) and measured values (Picciafuoco et al., 2019). Streets and the river bed are assumed to be impermeable. We specify an interception storage capacity of 5 mm for the riparian forests around the outflow (colored in light green in Fig. 9a) and of 2 mm for the arable land and grassland.

The investigated heavy rain event starts on June 23 at 21:00 and is simulated for 1.5 days. There are two distinct blocks of intense rainfall. The first occurs after 1 h, and the second after 28 h. The main flow paths of the surface runoff for the second rain block are clearly visible (Fig. 9a).

Table 1

HOAL Petzenkirchen. Total runtimes of the sequential CPU implementation and the GPU implementation of the first-order CN scheme at the end of the simulated 39 h. For this small-scale case, the GPU is exploiting its computational power only at the highest resolutions, for lower resolutions it is not fully occupied and computational resources are left unused.

Resolution Δx [m]	Number of cells	Runtime [h]		Runtime per cell [ms]		Speedup
		CPU	GPU	CPU	GPU	
1.0	748095	146.14	0.081	703.3	0.39	1813.3
2.0	187717	15.64	0.018	300.0	0.34	888.9
4.0	47276	2.18	0.008	166.1	0.61	272.8
8.0	11997	0.42	0.005	126.3	1.62	78.0

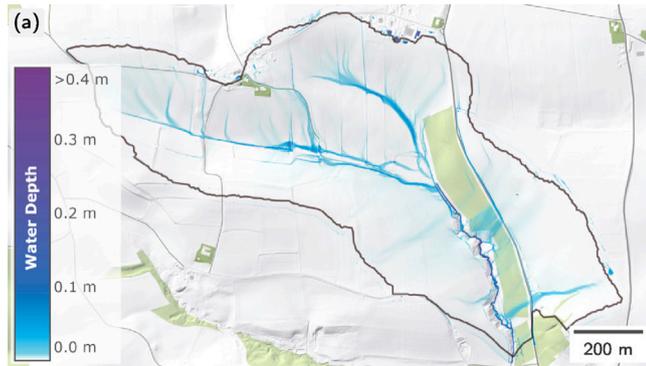


Fig. 9. HOAL Petzenkirchen. (a) Simulated water depths at the peak of the second rain block after 27.8 h.

For the 1 m simulation, 0.75 million cells are wet at the peak of the second rain block, which corresponds to the simulated region of 0.75 km² (including a small buffer zone around the catchment). Regarding the runtimes of the sequential CPU and the parallel GPU implementation, speedups of three orders of magnitude are achieved (Table 1). The speedup increases with higher resolutions and higher workloads as the GPU is not fully utilized for a low resolution. For a fully occupied GPU, doubling the resolution causes a theoretical increase of the amount of work by eight times as the number of cells quadruples and due to the CFL condition twice the number of time steps are required. For the 1 m simulation, 90% of the computation time on the GPU is spent in the following routines: the reconstruction and flux computation (32%), the time integration and time step reduction (30%), and the computation and integration of the runoff (28%). In the latter routines, it is not the amount of floating-point operations but the memory transfers that prevent the GPU from achieving faster runtimes. On the CPU, the distribution is slightly different with 60%, 11%, and 9%, respectively, due to faster memory access rates.

One drawback of the GPU implementation is the comparably longer development time. Parallel CPU implementations are possible (Neal et al., 2010; Noh et al., 2018; Morales-Hernández et al., 2021), but even if the implementation achieves full parallelization speedups, over thousand CPU cores would be needed to match the computational advantage of the GPU. From an economical and ecological perspective the GPU simulation still performs better with regards to power consumption than a parallel CPU simulation running on a supercomputer. The fast GPU simulation opens up new possibilities for this small catchment, such as calibration tasks within reasonable time spans.

3.4. Urban flooding in Cologne

We study two urban scenarios in the city of Cologne, Germany. First, we present a dual-drainage model at the central part of the city, at the eastern bank of the Rhine river, to which we refer as Cologne Center-East. Second, we present a city-scale simulation encompassing the entire city of Cologne with an area of 23.73 × 27.14 km² without sewer coupling.

In the first scenario, the region simulated with the coupled models lies at the eastern bank of the Rhine river and covers an area of 5.41 × 9.86 km². The terrain model is obtained from light detection and ranging (LIDAR) data where solid urban features such as buildings or bridges were removed in a pre-processing step, so that the DTM represents a so-called bare earth DTM. The resolution of the DTM is 1 m with a typical vertical accuracy around a decimeter (Kraus, 2011; Dottori et al., 2013). The terrain is relatively flat, mostly ranging between 40 and 60 m.a.s.l. (Fig. 10a). The simulation domain exhibits modest average slopes of 0.3 m/km along the Rhine river from south to north, and of around 1 m/km from east to west.

Buildings and land use data are extracted from the official ALKIS data set of 2021 (Caffier et al., 2017). Buildings cover 13% of the area, they are impermeable for the surface flow and water from roofs is routed to sewer nodes in the coupled model, thus building cells remain dry during simulation. Roughness coefficients are mapped from the land use, a detailed overview of the spatial distribution is shown in Fig. 10b. The interception parameters are assumed to correlate with land use. Woods and gardens are assigned a storage capacity of 5 mm, for public recreational areas and residential areas it is set to 2 mm. Rivers, streets and parking lots are assumed to not retain rain, thus their storage capacity is set to zero. The infiltration parameters are derived from a soil map, compare the saturated hydraulic conductivity in Fig. 10c. Streets and squares as well as rivers and lakes are considered impermeable.

We preprocessed sewer data for SWMM for the eastern bank of the Rhine river, therefore, cells west of the river are excluded from the simulation. The active simulation region is thus restricted to 39 km². The sewer network consists of 6392 junction nodes and 7206 conduits linking them with a total length of 245 km. Moreover there are 16 pumps, 16 outfalls and 18 weirs in the simulation domain, which are included in the model. The sewer network and the invalidated region are shown in Fig. 11a in yellow and pink, respectively. Exchange between the sewer network and the surface is assumed to occur at the nodes with the parameters specified in Section 2.7. Each node has a maximum inflow capacity of 0.1 m³/s. Rain that falls on buildings is directly routed to an assigned sewer junction node, if the roof water discharge exceeds the capacity, it spills over at the nodes. We simulate a hypothetical uniform one-hour rainfall of 53 mm/h corresponding to approximately a 100-year event according to the KOSTRA 2010R data set (Junghänel et al., 2017). With a cell size of 1 × 1 m², the grid has nearly 40 million cells valid for simulation. The water depths are aggregated in time during the first-order 1 m simulation resulting in a maximum water depth for each cell at the end of the simulation (Fig. 11b).

To illustrate the effects of the sewers, the resolution, and the order of accuracy of the surface flow scheme, we focus on the region marked with a red frame in Fig. 11b. The water depths of the coupled simulation for the specified region are aggregated in time resulting in maximum water depths (Fig. 12a). This heightfield serves as the reference for the difference heightfields, where positive values indicate higher maximum water depths in the reference simulation than in the corresponding alternative simulations. The differences between the simulation with sewer coupling and the one without are spatially restricted to the vicinity of the sewer network (Fig. 12b). For the

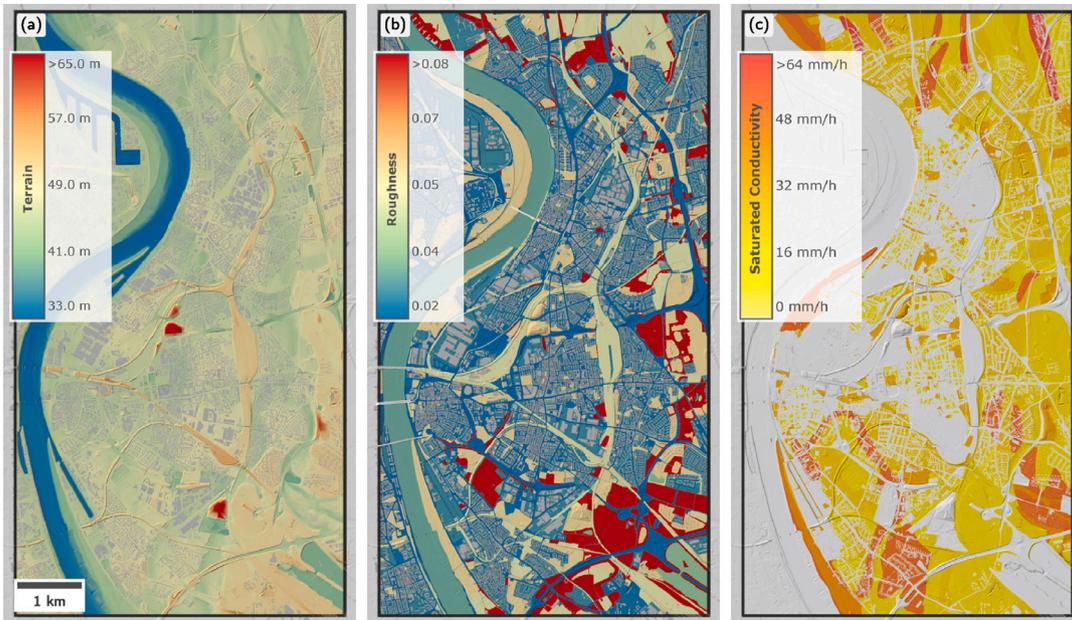


Fig. 10. Cologne Center-East coupled simulation. (a) The map of the terrain elevations shows only modest slopes across the simulation domain. The simulated region is heavily urbanized as indicated by the large amount of buildings, displayed in gray. (b) Manning's roughness coefficients ($s/m^{1/3}$) are derived from land use. (c) The spatial distribution of the saturated hydraulic conductivity is based on soil types. Streets and the river bed are set as impermeable zones.

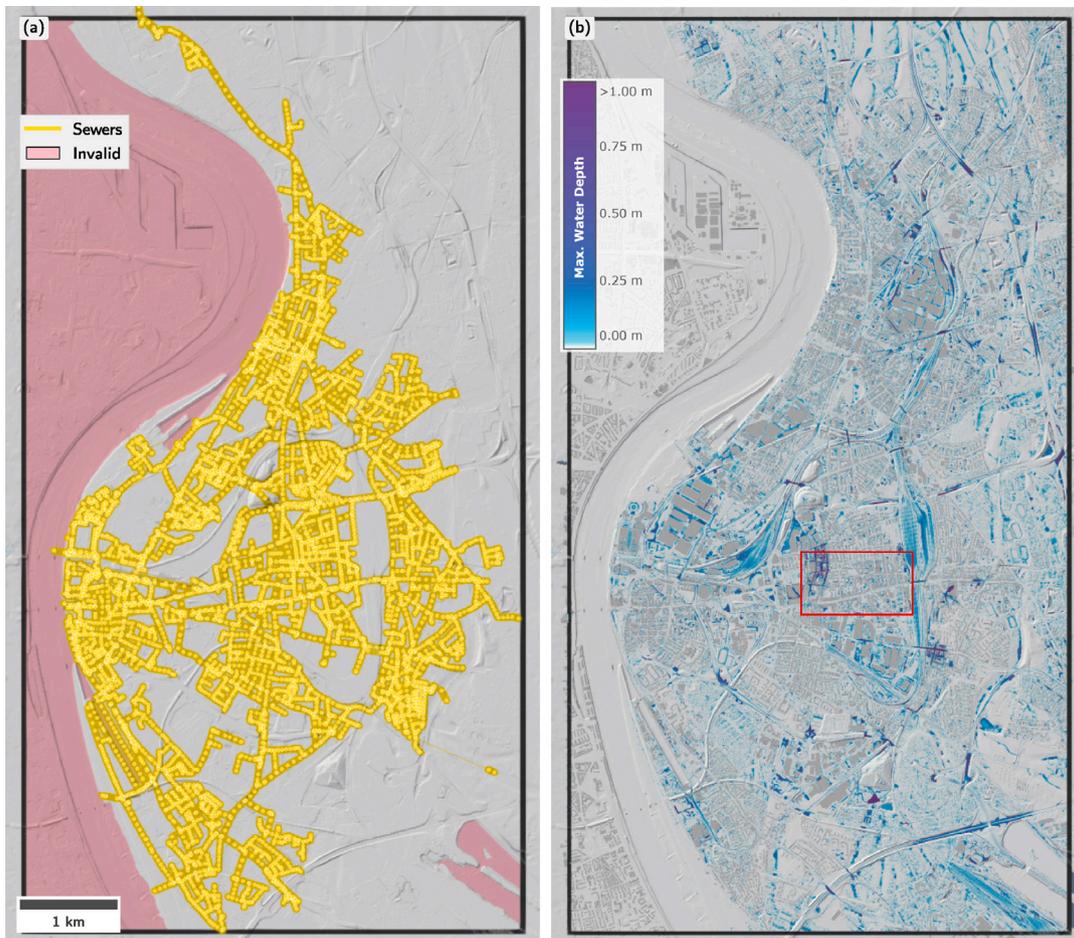


Fig. 11. Cologne Center-East coupled simulation. (a) The sewer network (yellow) and regions invalid for simulation (pink). (b) Maximum water depths of the 1 m simulation at the final simulation time $T = 2$ h. Fig. 12 focuses on the region marked with the red frame.

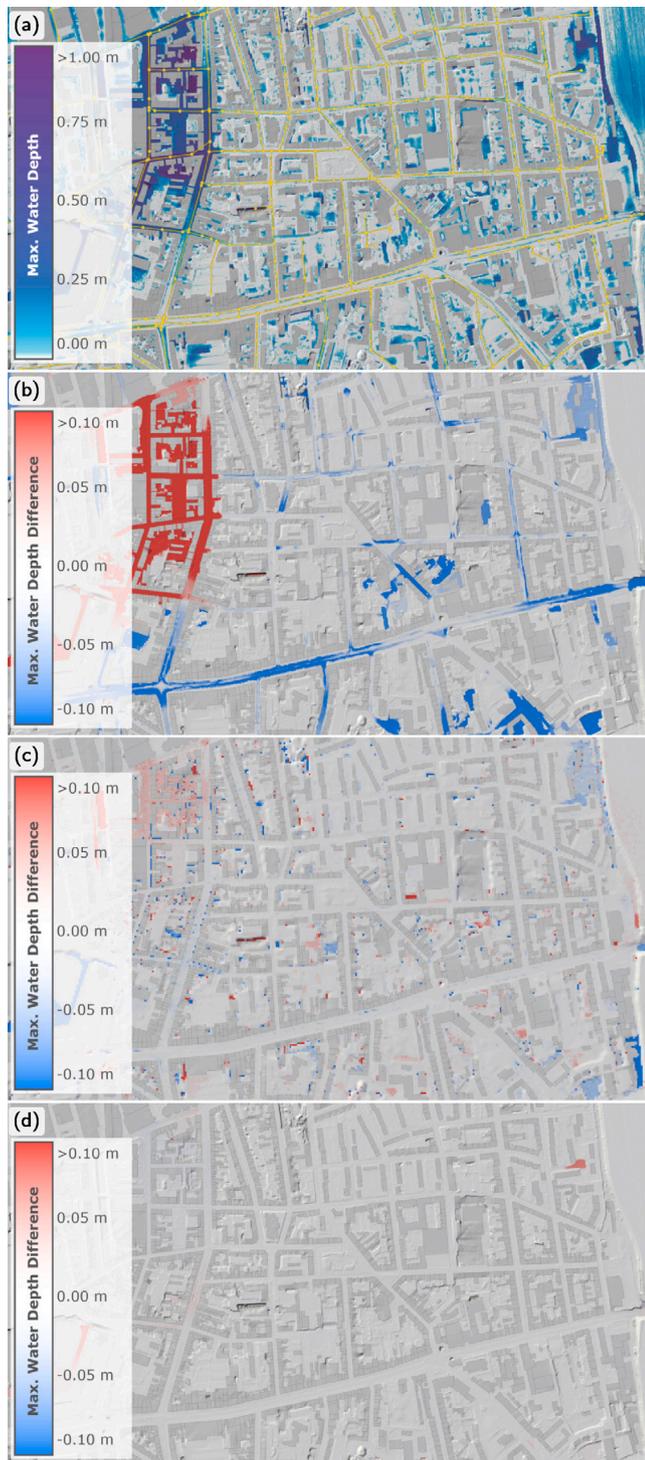


Fig. 12. Cologne Center-East detail. (a) Maximum water depths occurring during the coupled simulation with the first-order scheme and a resolution of 1 m. Difference heightfield of the maximum water depths between (b) the 1 m simulations with and without sewers, (c) the 1 m and the 4 m simulations, and (d) the first-order CN and the second-order BH accurate scheme. Regions with positive values (red) indicate higher maximum water depths in the coupled first-order simulation at 1 m shown in (a) than in the corresponding alternative simulations (b-d).

simulation without sewer coupling, we assume that water falling onto roofs can be drained by the sewer network. In regions with positive values (red), the maximum water depths of the coupled simulation

Table 2

Cologne Center-East coupled simulation. Water balance errors over a range of different coupling time step sizes ΔT_c and maximum water depth differences between the tightly coupled simulation and the proposed coupled model.

ΔT_c [s]	Continuity errors [m ³]			MAE [mm]	MSE [mm]
	Coupled	SWMM	Runoff		
0.5	858.6	69.8	15.1	0.10	0.04
1.0	7920.2	8588.2	15.0	0.38	0.32
2.0	19305.4	20464.9	14.9	0.98	0.89
4.0	20168.7	22034.5	14.9	1.35	1.22
8.0	24777.0	26448.8	14.8	1.46	1.35

(Fig. 12a) are higher than the maximum water depths of the simulation without sewer coupling. In terms of the MAE, the differences amount to 7.37 mm. The mean signed error (MSE), where the results of the runoff simulation without sewers are subtracted from those of the coupled simulation, amounts to -6.60 mm. This indicates that surface water levels do not rise as high in the coupled simulation due to sewer drainage. In fact, more water is drained from the streets than what is spilling onto the streets as excess roof water, which exceeds node inflow capacities. The sewer simulation also induces a water redistribution and causes minor floodings at a few streets due to sewer overflows. Overall, the sewer simulation drains around 250000 m³ of surface water.

In Fig. 12c, the difference field resulting from the subtraction of the maximum water depths computed by a 4 m simulation from the 1 m simulation is displayed. In regions with negative values (blue), the maximum water depths of the 4 m simulation are higher than the corresponding maximum water depths of the 1 m simulation. The difference between the maximum water depths of the 4 m and the 1 m grids are spatially concentrated at certain locations and appear mostly where the DTM shows strong variations at the scale of the employed cell size. For example, major differences occur in the vicinity of underpasses and garage entrances, or at the edge of elevated plateaus. At the edges of buildings, differences appear as boundary cells are rasterized as wall cells in one grid but not in the other. Overall, the MAE between the 4 m and the 1 m simulations is 8.12 mm. For the MAE computation the results of the 1 m grid are downsampled onto the 4 m grid.

In Fig. 12d, we compare the first-order accurate CN and the second-order accurate BH scheme. In general, the differences in the maximum water depths between the two surface flow discretizations are marginal. Noticeable deviations are concentrated on a few spots and occur where relatively high velocities up to 1 m/s develop. This happens for example at sloped entrances to inner courtyards. In the entire simulation domain, the MAE between the two schemes is 1.2 mm and the MSE amounts to -0.1 mm. As we consider differences between the maximum water depths occurring during the event, a non-nil MSE does not indicate a volume error but rather indicates that the surface water travels a greater distance. In fact, the negative MSE reveals that maximum water depths are slightly higher in the second-order BH scheme. In this comparison, the errors are considerably smaller than in the previous comparisons. The relatively small differences have to be considered in light of the mild terrain slopes.

To verify the validity of the proposed parallelized coupling approach, we perform a comparison of the proposed coupling method with a tightly coupled simulation as in Leandro and Martins (2016). In the tightly coupled simulation, the individual solvers execute one time step after each other, so that each solvers advances exactly one time step with a fixed time step size of 0.1 s. The continuity errors in the tightly coupled simulation amount to 1047.9 m³ due to flow routing errors (942.3 m³) in SWMM. The overall error can be considered acceptable when compared to the total sewer inflow volume (426802.8 m³). Furthermore, we compare the maximum water depths for a cell size of 4 m. Across the entire simulation domain the MSE is 0.32 mm and the MAE is 0.38 mm. Differences between the maximum water depths of the proposed model and the tightly coupled model are predominantly located around overflowing sewer nodes, see Fig. 13b,

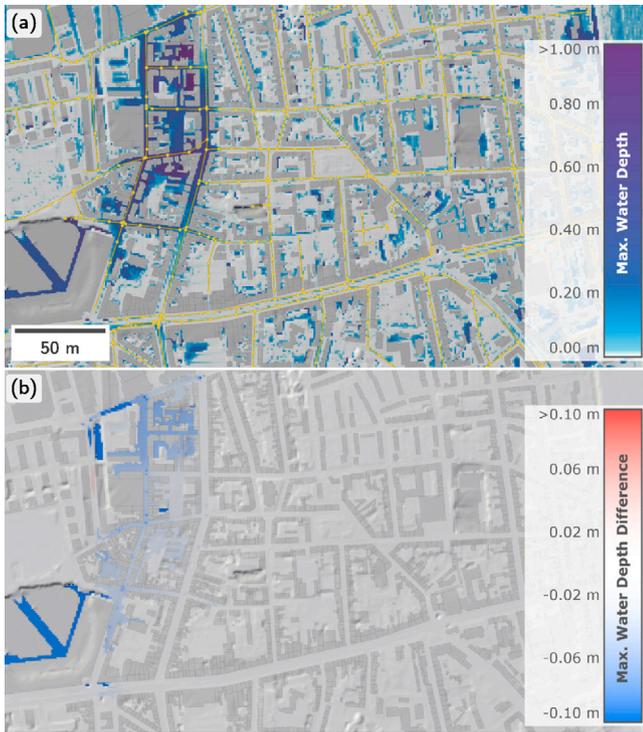


Fig. 13. Cologne Center-East detail. (a) Maximum water depths occurring during the tightly coupled simulation with the first-order scheme and a resolution of 4 m. (b) Difference heightfield of the maximum water depths of the tightly coupled simulation and the proposed model with a parallel coupling time step of 1 s. Regions with negative values (blue) indicate lower maximum water depths in the tightly coupled simulation.

Table 3

Cologne Center-East coupled simulation. Runtimes of the sequential SWMM CPU solver and the GPU runoff simulation for the simulated timespan of 2 h over different resolutions Δx . For resolutions above 2 m, the sewer simulation on the CPU dominates the runtime of the coupled model.

Resolution Δx [m]	Number of cells	Runtime [h]		Runoff/sewers runtime ratio
		Runoff (GPU)	Sewers (CPU)	
1.0	39023122	0.3555	0.2670	1.331
2.0	9755780	0.0545	0.2441	0.223
4.0	2438945	0.0110	0.2408	0.046

which shows the region marked with a red frame in Fig. 11b. Still, the MAE between the two coupling approaches is small. It is lower than the MAE between first- and second-order scheme, and considerably lower than the difference between the simulations on the 1 m and the 4 m grid.

We investigate the effect of the coupling time step size ΔT_C on the accuracy of the simulation results by considering continuity errors in the water volumes as well as differences in the maximum water depths with respect to the tightly coupled simulation, see Table 2. The overall inlet volume amounts to 1935514.8 m³. The water volume on the surface and in the sewers ranges from 698052.3 m³ to 762353.6 m³ and the outlet volumes from 1238814.4 m³ to 1243885.8 m³ for the different coupling time step sizes. The mass balance errors are dominated by the flow routing error of SWMM (Table 2). Overall, the total mass balance errors is below 1 per mille for coupling time step sizes under 2 s, and around 1 per mille for 4 s and 8 s. The runoff simulation's time step Δt varies between 0.18 s and 0.32 s for the 4 m grid. During the coupling time step (1 s) around 4 steps of the surface runoff simulation and around 10 sewer routing steps are performed in average. A coupling time step size of 1 s seems to offer a good compromise between computational performance and accuracy.

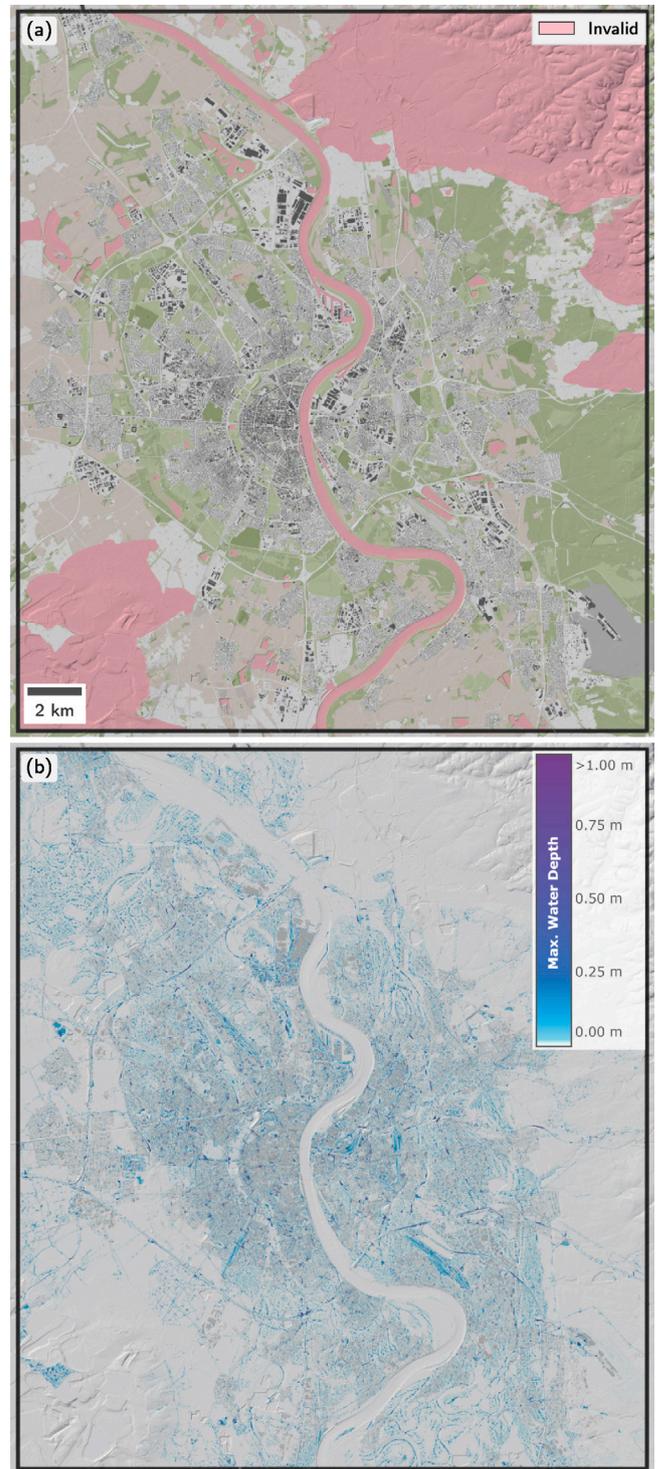


Fig. 14. Cologne runoff simulation. (a) Buildings (gray), land use and invalid cells (pink) for the simulated region covering the entire city. (b) Maximum water depth of the 1.5 m runoff simulation without sewer coupling after 2 h, i.e. at the end of the simulation.

To thoroughly assess the quality of the predictive capabilities of the coupled model, further validation is required. The validation of coupled models on large-scale scenarios is challenging as in most cases the collected data is sparse. Possible strategies to tackle this problem are the collection of crowd-sourced data (Yu et al., 2016; Wang et al., 2018; Xing et al., 2018), imagery from unmanned aerial vehicle sensing (Perks et al., 2016) or insurance claims (Zischg et al., 2018).

A comparison of runtimes of the coupled simulator for resolutions of 1 m, 2 m, and 4 m, in Table 3 shows that the GPU-accelerated 2D surface flow simulation is faster than the sequential 1D CPU sewer simulation for low resolutions. This emphasizes once more the massive gain in computing power for the surface flow simulation due to the GPU acceleration. Usually, solving the continuity and momentum equations for the sewer flow only requires around 0.1% of the total coupled simulation runtime for a sequential implementation (Noh et al., 2018). In the proposed implementation, as the simulations advance in parallel, the runtime per coupling time step is determined by the slower coupling component, which is either the GPU runoff simulation or the CPU sewer simulation. Thus, in order to improve the model performance further, effective parallelization strategies (Burger et al., 2014) of the sewer module are necessary.

For the derivation of pluvial flood hazard maps, we perform benchmark tests regarding large simulation domains with a size of $23.73 \times 27.14 \text{ km}^2$ spanning the entire city of Cologne. We simulate uniform rainfall of 53 mm/h that lasts for 1 h. In order to account for surface flow routing after the rainfall ends, the total simulated duration is extended to 2 h. In this city-scale scenario shown in Fig. 14a, infiltration and interception are considered and are set up analogously to the previously studied smaller domain. The 1.5 m simulation grid has a total number of 220 million cells valid for simulation, of which 17 million cells are rasterized buildings. In this scenario, we do not explicitly simulate the sewer network of the entire city, but we assume that water falling onto roofs can be drained by the sewer network, therefore building cells are excluded from simulation. All other input parameters remain unchanged. In Fig. 14b, we display the maximum water depths that occur during the simulated event. The computational speed is faster than physical time, with a total runtime of 1.62 h for a total simulated duration of 2 h. The simulation uses up to 23.4 GB of memory on the NVIDIA Titan RTX GPU, close to its limit of 24 GB.

In previous studies, urban regions of 40 km^2 were modeled with an efficient hybrid parallelization strategy on an adaptive grid with a minimum cell size of 1 m (Noh et al., 2018). Their model uses runoff coefficients depending on land use instead of a dynamic infiltration model as in this work, but dynamic sewer coupling is also integrated. In Xing et al. (2018), the city of Fuzhou, China, is modeled with a resolution of 2 m and a constant drainage loss at streets neglecting bidirectional sewer interaction and water routing in sewers. The simulation involves 66 million cells and runs almost in realtime on a rack of 8 NVIDIA Tesla K80 GPUs. The proposed computational model is efficient in terms of memory consumption and performance, it is able to simulate over 200 million cells on a single Titan RTX GPU with 24 GB memory, processing almost 10 million cells per GB of GPU memory. It is useful to extend the capabilities of single-GPU implementations to large regions as they typically run faster than comparable multi-GPU implementations requiring inter-GPU communication, which introduces an additional bottleneck (Morales-Hernández et al., 2021). In order to support larger simulation regions and higher spatial resolutions, an indispensable extension to multiple GPUs is possible.

4. Conclusions and perspectives

In this study, we present an integrated modeling framework for the simulation of rainfall–runoff processes and urban flash floods. The introduced modeling framework accounts for all major processes needed for an accurate description of flash floods while still accomplishing very fast runtimes. Infiltration is modeled with the Green–Ampt equations in a fully dynamic and spatially distributed way. An interception module accounts for initial rain abstractions due to vegetation. Instead of applying simple surface flow approximations, we discretize the full 2D shallow water equations (SWEs). In the context of rainfall–runoff simulations, the first-order accurate CN scheme (Chen and Noelle, 2017) was shown to be able to reproduce velocities and discharges accurately. In particular, using higher resolutions in the first-order

scheme proved to be more beneficial than using the second-order accurate BH scheme (Buttinger-Kreuzhuber et al., 2019) when considering the tradeoff between accuracy and required computational work. We remark that an appropriate bed source term discretization of the surface flow is essential for providing correct velocity estimates.

For urban flash floods, the rainfall–runoff simulation is coupled with the sewer network simulation from the Storm Water Management Model (SWMM). An effective approach for the bidirectional coupling of the sewer simulation to the surface runoff simulation is developed, where the two simulators advance in parallel in each coupling timestep. The runoff simulation is validated in the Thiès irrigation experiment and the surface–sewer coupling is validated in a laboratory experiment, both showing good agreement between simulations and observations. Regarding the validation of the integrated model, we point out that more detailed spatial observations are needed in order to assess the model's predictive performance in a more exhaustive way.

We demonstrate the benefits of using graphics processing units (GPUs) as computational devices to speed up the rainfall–runoff simulations. The GPU-accelerated model enables high-resolution simulations for entire cities with simulation domains involving up to 225 million cells on a single GPU with 24 GB of memory. In other words, we enable simulation of regions up to 220 km^2 with a resolution of 1 m in realtime. This removes the need for multiple localized, small-scale simulations. Speed-ups of up to three orders of magnitude are achieved for simulated regions with around 10 million cells, if compared against a serial CPU implementation. The speedup increases with the number of cells, thus large simulation regions profit even more from GPU acceleration.

The efficiency of the approach opens up new possibilities regarding ensemble simulations and high-resolution environmental modeling. The coupling of the surface flow with the spatially distributed infiltration and interception component allows the direct inclusion of green infrastructure by varying parameters accordingly. Detailed results help raise public awareness for flash floods by providing straightforward impact analysis at the scale of individual buildings and enable the analysis of efforts to mitigate the effects of climate change in rural and urban settings.

Software availability

The presented modeling framework Visdom (Waser et al., 2011) is developed in C++. The GPU implementation is written in the CUDA language. The Visdom framework is closed source. Licensing information as well as information on cooperation in a joint research project with Visdom is available on request from VRVis (www.vrvis.at).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors are supported by the VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH and by the Austrian Science Funds (FWF) as part of the Vienna Doctoral Programme on Water Resource Systems (DK W1219-N28). VRVis is funded by BMK, BMDW, Styria, SFG, Tyrol and Vienna Business Agency in the scope of COMET — Competence Centers for Excellent Technologies (879730) which is managed by FFG, Austria. The authors thank D. Hesslow for contributing to the implementation of the runoff model. The authors thank the Bundesamt für Wasserwirtschaft — Institut für Kulturtechnik und Bodenwasserhaushalt, the Stadtentwässerungsbetriebe (StEB) Köln (www.steb-koeln.de), and the open data platform of Nordrhein-Westfalen (<https://open.nrw>) for providing us with data necessary for setting up the Petzenkirchen and Cologne scenarios, respectively.

References

- Aricò, C., Nasello, C., 2018. Comparative analyses between the zero-inertia and fully dynamic models of the shallow water equations for unsteady overland flow propagation. *Water* 10 (1), <http://dx.doi.org/10.3390/w10010044>.
- Audusse, E., Bouchut, F., Bristeau, M.-O., Klein, R., Perthame, B., 2004. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM J. Sci. Comput.* 25 (6), 2050–2065.
- Audusse, E., Bristeau, M.-O., 2005. A well-balanced positivity preserving "second-order" scheme for shallow water flows on unstructured meshes. *J. Comput. Phys.* 206 (1), 311–333. <http://dx.doi.org/10.1016/j.jcp.2004.12.016>.
- Aureli, F., Prost, F., Vacondio, R., Dazzi, S., Ferrari, A., 2020. A GPU-accelerated shallow-water scheme for surface runoff simulations. *Water* 12 (3), 637. <http://dx.doi.org/10.3390/w12030637>.
- Ayog, J.L., Kesserwani, G., Shaw, J., Sharifian, M.K., Bau, D., 2021. Second-order discontinuous Galerkin flood model: Comparison with industry-standard finite volume models. *J. Hydrol.* 594, 125924. <http://dx.doi.org/10.1016/j.jhydrol.2020.125924>.
- Berland, A., Shiflett, S.A., Shuster, W.D., Garmestani, A.S., Goddard, H.C., Herrmann, D.L., Hopton, M.E., 2017. The role of trees in urban stormwater management. *Landscape Urban Plan.* 162, 167–177. <http://dx.doi.org/10.1016/j.landurbplan.2017.02.017>.
- Bermúdez, A., Dervieux, A., Desideri, J.-A., Vázquez, M.E., 1998. Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes. *Comput. Methods Appl. Mech. Engrg.* 155 (1–2), 49–72. [http://dx.doi.org/10.1016/S0045-7825\(97\)85625-3](http://dx.doi.org/10.1016/S0045-7825(97)85625-3).
- Blöschl, G., Blaschke, A.P., Broer, M., Bucher, C., Carr, G., Chen, X., Eder, A., Exner-Kittridge, M., Farnleitner, A., Flores-Orozco, A., Haas, P., Hogan, P., Amiri, A.K., Oismüller, M., Parajka, J., Silasari, R., Stadler, P., Strauss, P., Vreugdenhil, M., Wagner, W., Zessner, M., 2016. The hydrological open air laboratory (HOAL) in petzenkirchen: a hypothesis-driven observatory. *Hydrol. Earth Syst. Sci.* 20 (1), 227–255. <http://dx.doi.org/10.5194/hess-20-227-2016>.
- Blöschl, G., Hall, J., Viglione, A., Perdigão, R.A.P., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G.T., Bilibashi, A., Boháč, M., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G.B., Claps, P., Frolova, N., Ganora, D., Gorbachova, L., Gül, A., Hannaford, J., Harrigan, S., Kireeva, M., Kiss, A., Kjeldsen, T.R., Kohnová, S., Koskela, J.J., Ledvinka, O., Macdonald, N., Mavrou-Guirguinova, M., Mediero, L., Merz, R., Molnar, P., Montanari, A., Murphy, C., Osuch, M., Ovcharuk, V., Radevski, I., Salinas, J.L., Sauquet, E., Šraj, M., Szolgay, J., Volpi, E., Wilson, D., Zaimi, K., Živković, N., 2019. Changing climate both increases and decreases European river floods. *Nature* 573 (7772), 108–111. <http://dx.doi.org/10.1038/s41586-019-1495-6>.
- Brodtkorb, A.R., Sætra, M.L., Altinakar, M.S., 2012. Efficient shallow water simulations on GPUs: Implementation, visualization, verification, and validation. *Comput. & Fluids* 55, 1–12. <http://dx.doi.org/10.1016/j.compfluid.2011.10.012>.
- Burger, G., Sitzenfrei, R., Kleidorfer, M., Rauch, W., 2014. Parallel flow routing in SWMM 5. *Environ. Model. Softw.* 53, 27–34. <http://dx.doi.org/10.1016/j.envsoft.2013.11.002>.
- Buttinger-Kreuzhuber, A., Horváth, Z., Noelle, S., Blöschl, G., Waser, J., 2019. A fast second-order shallow water scheme on two-dimensional structured grids over abrupt topography. *Adv. Water Resour.* 127, 89–108. <http://dx.doi.org/10.1016/j.advwatres.2019.03.010>.
- Caffier, A., Elsner, C., Rath, C., Robens, F., Seidel, J., Will, K., 2017. Offene Geobasisdaten für NRW. *Z. Geodäsie Geoinf. Landmanag.(ZfV)* 142, 133–145.
- Carsel, R.F., Parrish, R.S., 1988. Developing joint probability distributions of soil water retention characteristics. *Water Resour. Res.* 24 (5), 755–769. <http://dx.doi.org/10.1029/wr024i005p00755>.
- Caviedes-Voullième, D., Fernández-Pato, J., Hinz, C., 2020. Performance assessment of 2D zero-inertia and shallow water models for simulating rainfall-runoff processes. *J. Hydrol.* 584, 124663. <http://dx.doi.org/10.1016/j.jhydrol.2020.124663>.
- Chen, A.S., Leandro, J., Djordjević, S., 2016. Modelling sewer discharge via displacement of manhole covers during flood events using 1D/2D SIPSON/P-DWave dual drainage simulations. *Urban Water J.* 13 (8), 830–840. <http://dx.doi.org/10.1080/1573062X.2015.1041991>.
- Chen, Y., Moufouma-Okia, W., Masson-Delmotte, V., Zhai, P., Pirani, A., 2018. Recent progress and emerging topics on weather and climate extremes since the fifth assessment report of the intergovernmental panel on climate change. *Ann. Rev. Environ. Res.* 43 (1), 35–59. <http://dx.doi.org/10.1146/annurev-environ-102017-030052>.
- Chen, G., Noelle, S., 2017. A new hydrostatic reconstruction scheme based on subcell reconstructions. *SIAM J. Numer. Anal.* 55 (2), 758–784.
- Chow, V.T., Maidment, D., Mays, L., 1988. *Applied Hydrology*. In: McGraw-Hill series in water resources and environmental engineering, McGraw-Hill.
- Corradini, C., Melone, F., Smith, R.E., 2000. Modeling local infiltration for a two-layered soil under complex rainfall patterns. *J. Hydrol.* 237 (1), 58–73. [http://dx.doi.org/10.1016/S0022-1694\(00\)00298-5](http://dx.doi.org/10.1016/S0022-1694(00)00298-5).
- Costabile, P., Costanzo, C., De Lorenzo, G., Macchione, F., 2020. Is local flood hazard assessment in urban areas significantly influenced by the physical complexity of the hydrodynamic inundation model? *J. Hydrol.* 580, 124231. <http://dx.doi.org/10.1016/j.jhydrol.2019.124231>.
- Costabile, P., Costanzo, C., Macchione, F., 2013. A storm event watershed model for surface runoff based on 2D fully dynamic wave equations. *Hydrol. Process.* 27 (4), 554–569. <http://dx.doi.org/10.1002/hyp.9237>.
- Costabile, P., Costanzo, C., Macchione, F., 2017. Performances and limitations of the diffusive approximation of the 2-d shallow water equations for flood simulation in urban and rural areas. *Appl. Numer. Math.* 116, 141–156. <http://dx.doi.org/10.1016/j.apnum.2016.07.003>, *New Trends in Numerical Analysis: Theory, Methods, Algorithms and Applications (NETNA 2015)*.
- Cozzolino, L., Cimorelli, L., Della Morte, R., Pugliano, G., Piscopo, V., Pianese, D., 2019. Flood propagation modeling with the local inertia approximation: Theoretical and numerical analysis of its physical limitations. *Adv. Water Resour.* 133, 103422. <http://dx.doi.org/10.1016/j.advwatres.2019.103422>.
- Dazzi, S., Shustikova, I., Domeneghetti, A., Castellarin, A., Vacondio, R., 2021. Comparison of two modelling strategies for 2D large-scale flood simulations. *Environ. Model. Softw.* 146, 105225. <http://dx.doi.org/10.1016/j.envsoft.2021.105225>.
- de Almeida, G., Bates, P., Ozdemir, H., 2016. Modelling urban floods at submetre resolution: challenges or opportunities for flood risk management? *J. Flood Risk Manag.* 11, S855–S865. <http://dx.doi.org/10.1111/jfr3.12276>.
- Delestre, O., Cordier, S., Darboux, F., James, F., 2012. A limitation of the hydrostatic reconstruction technique for shallow water equations. *C. R. Math.* 350 (13–14), 677–681. <http://dx.doi.org/10.1016/j.crma.2012.08.004>.
- Delestre, O., Darboux, F., James, F., Lucas, C., Laguerre, C., Cordier, S., 2017. FullSWOF: Full shallow-water equations for overland flow. *J. Open Source Softw.* 2 (20), 448. <http://dx.doi.org/10.21105/joss.00448>.
- Djordjević, S., Prodanović, D., Maksimović, Č., Ivetić, M., Savić, D., 2005. SIPSON – simulation of interaction between pipe flow and surface overland flow in networks. *Water Sci. Technol.* 52 (5), 275–283. <http://dx.doi.org/10.2166/wst.2005.0143>.
- Dong, J., Li, D.F., 2021. A new second-order modified hydrostatic reconstruction for the shallow water flows with a discontinuous topography. *Appl. Numer. Math.* 161, 408–424. <http://dx.doi.org/10.1016/j.apnum.2020.11.019>.
- Dottori, F., Di Baldassarre, G., Todini, E., 2013. Detailed data is welcome, but with a pinch of salt: Accuracy, precision, and uncertainty in flood inundation modeling. *Water Resour. Res.* 49 (9), 6079–6085. <http://dx.doi.org/10.1002/wrcr.20406>.
- Echeverribar, I., Morales-Hernández, M., Brufau, P., García-Navarro, P., 2019. 2D numerical simulation of unsteady flows for large scale floods prediction in real time. *Adv. Water Resour.* 134, 103444. <http://dx.doi.org/10.1016/j.advwatres.2019.103444>.
- Fernández-Pato, J., Caviedes-Voullième, D., García-Navarro, P., 2016. Rainfall/runoff simulation with 2D full shallow water equations: Sensitivity analysis and calibration of infiltration parameters. *J. Hydrol.* 536, 496–513. <http://dx.doi.org/10.1016/j.jhydrol.2016.03.021>.
- Fernández-Pato, J., García-Navarro, P., 2018. Development of a new simulation tool coupling a 2D finite volume overland flow model and a drainage network model. *Geosciences* 8 (8), 288. <http://dx.doi.org/10.3390/geosciences8080288>.
- Fewtrell, T.J., Duncan, A., Sampson, C.C., Neal, J.C., Bates, P.D., 2011. Benchmarking urban flood models of varying complexity and scale using high resolution terrestrial LiDAR data. *Phys. Chem. Earth, Parts A/B/C* 36 (7), 281–291. <http://dx.doi.org/10.1016/j.pce.2010.12.011>, *Recent Advances in Mapping and Modelling Flood Processes in Lowland Areas*.
- Fiedler, F.R., Ramirez, J.A., 2000. A numerical method for simulating discontinuous shallow flow over an infiltrating surface. *Internat. J. Numer. Methods Fluids* 32 (2), 219–239. [http://dx.doi.org/10.1002/\(SICI\)1097-0363\(20000130\)32:2<219::AID-FLD936>3.0.CO;2-J](http://dx.doi.org/10.1002/(SICI)1097-0363(20000130)32:2<219::AID-FLD936>3.0.CO;2-J).
- Fry, T.J., Maxwell, R.M., 2018. Using a distributed hydrologic model to improve the green infrastructure parameterization used in a lumped model. *Water* 10 (12), <http://dx.doi.org/10.3390/w10121756>.
- Gallegos, H.A., Schubert, J.E., Sanders, B.F., 2009. Two-dimensional, high-resolution modeling of urban dam-break flooding: A case study of Baldwin Hills, California. *Adv. Water Resour.* 32 (8), 1323–1335. <http://dx.doi.org/10.1016/j.advwatres.2009.05.008>.
- Gowdith, L., Muñoz Carpena, R., 2009. An improved Green-Ampt infiltration and redistribution method for uneven multistorm series. *Vadose Zone J.* 8 (2), 470–479. <http://dx.doi.org/10.2136/vzj2008.0049>.
- Grayson, R., Blöschl, G., 2001. *Spatial Patterns in Catchment Hydrology: Observations and Modelling*. CUP Archive.
- Horváth, Z., Buttinger-Kreuzhuber, A., Konev, A., Cornel, D., Komma, J., Blöschl, G., Noelle, S., Waser, J., 2020. Comparison of fast shallow-water schemes on real-world floods. *J. Hydraul. Eng.* 146 (1), 05019005. [http://dx.doi.org/10.1061/\(ASCE\)HY.1943-7900.0001657](http://dx.doi.org/10.1061/(ASCE)HY.1943-7900.0001657).
- Horváth, Z., Perdigão, R.A., Waser, J., Cornel, D., Konev, A., Blöschl, G., 2016. Kepler shuffle for real-world flood simulations on GPUs. *Int. J. High Perform. Comput. Appl.* 30 (4), 379–395. <http://dx.doi.org/10.1177/1094342016630800>.
- Horváth, Z., Waser, J., Perdigão, R.A., Konev, A., Blöschl, G., 2015. A two-dimensional numerical scheme of dry/wet fronts for the saint-venant system of shallow water equations. *Internat. J. Numer. Methods Fluids* 73 (3), 159–182.
- Hou, J., Liang, Q., Zhang, H., Hinkelmann, R., 2014. Multislope MUSCL method applied to solve shallow water equations. *Comput. Math. Appl.* 68 (12), 2012–2027.
- Hou, J., Liang, Q., Zhang, H., Hinkelmann, R., 2015. An efficient unstructured MUSCL scheme for solving the 2D shallow water equations. *Environ. Model. Softw.* 66, 131–152. <http://dx.doi.org/10.1016/j.envsoft.2014.12.007>.

- Jongman, B., 2018. Effective adaptation to rising flood risk. *Nature Commun.* 9 (1), <http://dx.doi.org/10.1038/s41467-018-04396-1>.
- Junghänel, T., Ertel, H., Deutschländer, T., 2017. Bericht zur Revision von KOSTRA-DWD-2010. Technical Report, Deutscher Wetterdienst, Abteilung Hydrometeorologie, URL https://www.dwd.de/DE/leistungen/kostra_dwd_rasterwerte/kostra_dwd_rasterwerte.html (Last visited on 29 April 2021).
- Kesserwani, G., Ghostine, R., Vazquez, J., Ghenaïm, A., Mosé, R., 2008. Application of a second-order Runge–Kutta discontinuous Galerkin scheme for the shallow water equations with source terms. *Internat. J. Numer. Methods Fluids* 56 (7), 805–821. <http://dx.doi.org/10.1002/fld.1550>.
- Kraus, K., 2011. Photogrammetry. De Gruyter, <http://dx.doi.org/10.1515/9783110892871>.
- Kuffour, B.N.O., Engdahl, N.B., Woodward, C.S., Condon, L.E., Kollet, S., Maxwell, R.M., 2020. Simulating coupled surface–subsurface flows with ParFlow v3.5.0: capabilities, applications, and ongoing development of an open-source, massively parallel, integrated hydrologic model. *Geosci. Model Dev.* 13 (3), 1373–1397. <http://dx.doi.org/10.5194/gmd-13-1373-2020>.
- Kvočka, D., Falconer, R.A., Bray, M., 2015. Appropriate model use for predicting elevations and inundation extent for extreme flood events. *Nat. Hazards* 79 (3), 1791–1808. <http://dx.doi.org/10.1007/s11069-015-1926-0>.
- Lacasta, A., Morales-Hernández, M., Murillo, J., García-Navarro, P., 2015. GPU implementation of the 2D shallow water equations for the simulation of rainfall/runoff events. *Environ. Earth Sci.* 74 (11), 7295–7305.
- Lastra, M., Mantas, J.M., Ureña, C., Castro, M.J., García-Rodríguez, J.A., 2009. Simulation of shallow-water systems using graphics processing units. *Math. Comput. Simulation* 80 (3), 598–618.
- Le, P.V., Kumar, P., Valocchi, A.J., Dang, H.-V., 2015. GPU-based high-performance computing for integrated surface–sub-surface flow modeling. *Environ. Model. Softw.* 73, 1–13. <http://dx.doi.org/10.1016/j.envsoft.2015.07.015>.
- Leandro, J., Martins, R., 2016. A methodology for linking 2D overland flow models with the sewer network model SWMM 5.1 based on dynamic link libraries. *Water Sci. Technol.* 73 (12), 3017–3026. <http://dx.doi.org/10.2166/wst.2016.171>.
- Leandro, J., Schumann, A., Pfister, A., 2016. A step towards considering the spatial heterogeneity of urban key features in urban hydrology flood modelling. *J. Hydrol.* 535, 356–365. <http://dx.doi.org/10.1016/j.jhydrol.2016.01.060>.
- Li, G., Caleffi, V., Gao, J., 2014. High-order well-balanced central WENO scheme for pre-balanced shallow water equations. *Comput. & Fluids* 99, 182–189. <http://dx.doi.org/10.1016/j.compfluid.2014.04.022>.
- Li, Q., Liang, Q., Xia, X., 2020. A novel 1D-2D coupled model for hydrodynamic simulation of flows in drainage networks. *Adv. Water Resour.* 137, 103519. <http://dx.doi.org/10.1016/j.advwatres.2020.103519>.
- Liang, Q., Smith, L.S., 2015. A high-performance integrated hydrodynamic modelling system for urban flood simulations. *J. Hydroinform.* 17 (4), 518–533. <http://dx.doi.org/10.2166/hydro.2015.029>.
- Maxwell, R.M., 2013. A terrain-following grid transform and preconditioner for parallel, large-scale, integrated hydrologic modeling. *Adv. Water Resour.* 53, 109–117. <http://dx.doi.org/10.1016/j.advwatres.2012.10.001>.
- Mohammadzadeh-Habili, J., Heidarpour, M., 2015. Application of the Green–Ampt model for infiltration into layered soils. *J. Hydrol.* 527, 824–832. <http://dx.doi.org/10.1016/j.jhydrol.2015.05.052>.
- Morales-Hernández, M., Sharif, M.B., Gangrade, S., Dullo, T.T., Kao, S.-C., Kalyanapu, A., Ghafoor, S.K., Evans, K.J., Madadi-Kandjani, E., Hodges, B.R., 2020. High-performance computing in water resources hydrodynamics. *J. Hydroinform.* 22 (5), 1217–1235. <http://dx.doi.org/10.2166/hydro.2020.163>.
- Morales-Hernández, M., Sharif, M.B., Kalyanapu, A., Ghafoor, S., Dullo, T., Gangrade, S., Kao, S.-C., Norman, M., Evans, K., 2021. TRITON: A multi-GPU open source 2D hydrodynamic flood model. *Environ. Model. Softw.* 141, 105034. <http://dx.doi.org/10.1016/j.envsoft.2021.105034>.
- Mügler, C., Planchon, O., Patin, J., Weill, S., Silvera, N., Richard, P., Mouche, E., 2011. Comparison of roughness models to simulate overland flow and tracer transport experiments under simulated rainfall at plot scale. *J. Hydrol.* 402 (1), 25–40. <http://dx.doi.org/10.1016/j.jhydrol.2011.02.032>.
- Navas-Montilla, A., Murillo, J., 2015. Energy balanced numerical schemes with very high order. The augmented Roe flux ADER scheme. Application to the shallow water equations. *J. Comput. Phys.* 290, 188–218. <http://dx.doi.org/10.1016/j.jcp.2015.03.002>.
- Neal, J.C., Fewtrell, T.J., Bates, P.D., Wright, N.G., 2010. A comparison of three parallelisation methods for 2D flood inundation models. *Environ. Model. Softw.* 25 (4), 398–411. <http://dx.doi.org/10.1016/j.envsoft.2009.11.007>.
- Neal, J.C., Schumann, G., Bates, P.D., 2012. A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resour. Res.* 48 (11), <http://dx.doi.org/10.1029/2012WR012514>.
- Noelle, S., Pankratz, N., Puppo, G., Natvig, J.R., 2006. Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows. *J. Comput. Phys.* 213 (2), 474–499.
- Noelle, S., Xing, Y., Shu, C.-W., 2007. High-order well-balanced finite volume WENO schemes for shallow water equation with moving water. *J. Comput. Phys.* 226 (1), 29–58.
- Noh, S.J., Lee, J.-H., Lee, S., Kawaike, K., Seo, D.-J., 2018. Hyper-resolution 1D-2D urban flood modelling using LiDAR data and hybrid parallelization. *Environ. Model. Softw.* 103, 131–145. <http://dx.doi.org/10.1016/j.envsoft.2018.02.008>.
- Pachaly, R.L., Vasconcelos, J.G., Allasia, D.G., Tassi, R., Bocchi, J.P.P., 2020. Comparing SWMM 5.1 calculation alternatives to represent unsteady stormwater Sewer flows. *J. Hydraul. Eng.* 146 (7), 04020046. [http://dx.doi.org/10.1061/\(ASCE\)HY.1943-7900.0001762](http://dx.doi.org/10.1061/(ASCE)HY.1943-7900.0001762).
- Perks, M.T., Russell, A.J., Large, A.R.G., 2016. Technical note: Advances in flash flood monitoring using unmanned aerial vehicles (UAVs). *Hydrol. Earth Syst. Sci.* 20 (10), 4005–4015. <http://dx.doi.org/10.5194/hess-20-4005-2016>.
- Picciafuoco, T., Morbidelli, R., Flammini, A., Saltalippi, C., Corradini, C., Strauss, P., Blöschl, G., 2019. On the estimation of spatially representative plot scale saturated hydraulic conductivity in an agricultural setting. *J. Hydrol.* 570, 106–117. <http://dx.doi.org/10.1016/j.jhydrol.2018.12.044>.
- Rawls, W.J., Brakensiek, D.L., Miller, N., 1983. Green–Ampt infiltration parameters from soils data. *J. Hydraul. Eng.* 109 (1), 62–70. [http://dx.doi.org/10.1061/\(asce\)0733-9429\(1983\)109:1\(62\)](http://dx.doi.org/10.1061/(asce)0733-9429(1983)109:1(62)).
- Robinson, M., Ward, R., 2017. *Hydrology: Principles and Processes*. IWA Publishing, <http://dx.doi.org/10.2166/9781780407296>.
- Rosenzweig, B.R., Herreros Cantis, P., Kim, Y., Cohn, A., Grove, K., Brock, J., Yesuf, J., Mistry, P., Welty, C., McPhearson, T., Sauer, J., Chang, H., 2021. The value of urban flood modeling. *Earth's Fut.* 9 (1), e2020EF001739. <http://dx.doi.org/10.1029/2020EF001739>.
- Rossman, L.A., 2006. *Storm Water Management Model Quality Assurance Report: Dynamic Wave Flow Routing*. Technical Report, U.S. Environmental Protection Agency, Cincinnati, OH, EPA/600/R-06/097.
- Rossman, L.A., 2017. In: EPA (Ed.), *Storm Water Management Model*. U.S. Environmental Protection Agency, Cincinnati, OH, EPA/600/R-17/111.
- Rubinato, M., Martins, R., Kesserwani, G., Leandro, J., Djordjević, S., Shucksmith, J., 2017. Experimental calibration and validation of sewer/surface flow exchange equations in steady and unsteady flow conditions. *J. Hydrol.* 552, 421–432. <http://dx.doi.org/10.1016/j.jhydrol.2017.06.024>.
- Simons, F., Busse, T., Hou, J., Özgen, I., Hinkelmann, R., 2013. A model for overland flow and associated processes within the hydroinformatics modelling system. *J. Hydroinform.* 16 (2), 375–391. <http://dx.doi.org/10.2166/hydro.2013.173>.
- Smith, E., Smettem, R.J., Broadbridge, P., Woolhiser, D.A., 2002. Infiltration theory for hydrologic applications. American Geophysical Union, <http://dx.doi.org/10.1029/wm015>.
- Tatard, L., Planchon, O., Wainwright, J., Nord, G., Favis-Mortlock, D., Silvera, N., Ribolzi, O., Esteves, M., Huang, C.H., 2008. Measurement and modelling of high-resolution flow-velocity data under simulated rainfall on a low-slope sandy soil. *J. Hydrol.* 348 (1–2), 1–12. <http://dx.doi.org/10.1016/j.jhydrol.2007.07.016>.
- Vacondio, R., Dal Palù, A., Mignosa, P., 2014. GPU-enhanced finite volume shallow water solver for fast flood simulations. *Environ. Model. Softw.* 57, 60–75.
- Van Leer, B., 1979. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *J. Comput. Phys.* 32 (1), 101–136.
- Vater, S., Beisiegel, N., Behrens, J., 2017. Comparison of wetting and drying between a RKDG2 method and classical FV based second-order hydrostatic reconstruction. In: Cancès, C., Omnes, P. (Eds.), *Finite Volumes for Complex Applications VIII — Hyperbolic, Elliptic and Parabolic Problems*. Springer International Publishing, Cham, pp. 237–245.
- Wang, Y., Chen, A.S., Fu, G., Djordjević, S., Zhang, C., Savić, D.A., 2018. An integrated framework for high-resolution urban flood modelling considering multiple information sources and urban features. *Environ. Model. Softw.* 107, 85–95. <http://dx.doi.org/10.1016/j.envsoft.2018.06.010>.
- Waser, J., Ribičić, H., Fuchs, R., Hirsch, C., Schindler, B., Blöschl, G., Gröller, M.E., 2011. Nodes on ropes: A comprehensive data and control flow for steering ensemble simulations. *IEEE Trans. Vis. Comput. Graphics* 17 (12), 1872–1881.
- Xia, X., Liang, Q., Ming, X., 2019. A full-scale fluvial flood modelling framework based on a high-performance integrated hydrodynamic modelling system (HIPIMS). *Adv. Water Resour.* 132, 103392. <http://dx.doi.org/10.1016/j.advwatres.2019.103392>.
- Xia, X., Liang, Q., Ming, X., Hou, J., 2017. An efficient and stable hydrodynamic model with novel source term discretization schemes for overland flow and flood simulations. *Water Resour. Res.* 53 (5), 3730–3759. <http://dx.doi.org/10.1002/2016WR020055>.
- Xiao, Q., McPherson, E.G., Simpson, J.R., Ustin, S.L., 1998. Rainfall interception by Sacramento's urban forest. *J. Arboric.* 24 (4), 235–244.
- King, Y., Liang, Q., Wang, G., Ming, X., Xia, X., 2018. City-scale hydrodynamic modelling of urban flash floods: the issues of scale and resolution. *Nat. Hazards* 96 (1), 473–496. <http://dx.doi.org/10.1007/s11069-018-3553-z>.
- Yang, Y., Sun, L., Li, R., Yin, J., Yu, D., 2020. Linking a storm water management model to a novel two-dimensional model for urban pluvial flood modeling. *Int. J. Disaster Risk Sci.* 11 (4), 508–518. <http://dx.doi.org/10.1007/s13753-020-00278-7>.
- Yu, D., Yin, J., Liu, M., 2016. Validating city-scale surface water flood modelling using crowd-sourced data. *Environ. Res. Lett.* 11 (12), 124011. <http://dx.doi.org/10.1088/1748-9326/11/12/124011>.
- Zehe, E., Elsenbeer, H., Lindenmaier, F., Schulz, K., Blöschl, G., 2007. Patterns of predictability in hydrological threshold systems. *Water Resour. Res.* 43 (7), <http://dx.doi.org/10.1029/2006WR005589>.
- Zischg, A.P., Mosimann, M., Bernet, D.B., Röthlisberger, V., 2018. Validation of 2D flood models with insurance claims. *J. Hydrol.* 557, 350–361. <http://dx.doi.org/10.1016/j.jhydrol.2017.12.042>.