Enabling Real-Time High-Resolution Flood Forecasting for the Entire State of Berlin Through ~~RIM2D's~~ Multi-GPU ~~Processing~~Accelerated Physics-Based Modeling

Shahin Khosh Bin Ghomash[1,2], Siqi Deng[3], and Heiko Apel[1]

[1]Section Hydrology, GFZ German Research Centre for Geosciences, Potsdam (Germany)
[2]Department of Earth System Science, Stanford University, Stanford (United States of America)
[3]Institute of Geo-Hydroinformatics, Hamburg University of Technology, Hamburg, (Germany)

Correspondence: Shahin Khosh Bin Ghomash (shahink@stanford.edu)

Abstract. Urban areas are increasingly experiencing more frequent and intense pluvial flooding due to the combined effects of climate change and rapid urbanization—a trend expected to continue in the coming decades. This highlights the growing need for effective flood forecasting and disaster management systems. While recent advances in GPU computing have made high-resolution hydrodynamic modeling feasible at the urban scale, operational use remains limited, particularly for large domains where single-GPU processing falls short in terms of memory and performance.

This study demonstrates the capabilities of the hydrodynamic model RIM2D (Rapid Inundation Model 2D), enhanced with multi-GPU processing, to perform high-resolution pluvial flood simulations across large urban domains such as the whole state of Berlin (891.8 km^2) within operationally relevant timeframes. We evaluate RIM2D's performance across spatial resolutions of 2, 5, and 10 meters using GPU configurations ranging from 1 to 8 units. Two flood scenarios are analyzed: the real-world pluvial flood of June 2017 and a standardized 100-year return period (HQ100) event used for official hazard mapping. Results show that RIM2D can deliver detailed flood extents, flow characteristics, and impact estimates for the 48-hour 2017 event in 8 minutes at 10 m resolution, 34 minutes at 5 m, and approximately 5.5 hours at 2 m using 8 A100 GPUs—fast enough to be integrated into real-time early warning systems~~, even at fine spatial resolutions~~. Multi-GPU processing proves essential not only for enabling high-resolution simulations (e.g., dx = 2 m or finer), but also for making simulations at resolutions finer than 5 m computationally feasible for flood forecasting and early warning applications. Additionally, we find that beyond 4 GPUs, runtime improvements become marginal for 5 and 10 m resolutions, and similarly, more than 6 GPUs offer limited benefit at dx= 2 m resolution, illustrating the balance between computational nodes of the used GPUs and number of raster cells of the model. Moreover, simulations at a finer dx = 1 m resolution demand more than 8 GPUs to be run. Overall, this work demonstrates that large-scale, high-resolution flood simulations can now be executed rapidly enough to support operational early warning and impact-based forecasting. With models like RIM2D and the continued advancement of GPU hardware, the integration of detailed, real-time flood forecasting into urban flood risk management is both technically feasible and urgently needed.

1 Introduction

Pluvial flooding in urban areas is an increasing concern for cities worldwide, with its frequency and severity projected to rise due to accelerating climate change and rapid urban growth (Zhou et al., 2019; Rentschler et al., 2023). In the face of this, timely and reliable flood data is critical to support informed decisions by emergency planners, local officials, and at-risk communities (Kreibich et al., 2021). Forecasting systems designed to provide early warnings are key to reducing casualties and minimizing damage by enabling preemptive actions (Merz et al., 2020; Šakić Trogrlić et al., 2022). At the core of these systems are flood forecasting models, which offer the capability to simulate and anticipate flood patterns under a wide range of conditions, with computational hydrodynamic models playing a key role in supporting such analyses.

Urban environments such as Berlin can exhibit highly complex and variable topographies, shaped by dense infrastructure, heterogeneous land use, and abrupt elevation changes over short distances (Wang et al., 2018). These characteristics can significantly influence the spatial distribution of runoff dynamics and flow connectivity (Khosh Bin Ghomash et al., 2019), necessitating high-resolution flood models to accurately simulate water movement through such intricate terrain. Coarse-resolution models often fail to capture these critical terrain features, leading to underestimation or misrepresentation of localized flood hazards (Khosh Bin Ghomash et al., 2024). However, achieving the necessary spatial detail across an entire urban region poses substantial computational challenges. Traditional hydrodynamic models, typically CPU-based, struggle to run simulations for such large domains at sufficiently high resolutions due to their computational intensity. Moreover, these models are often impractical for real-time flood forecasting and early warning systems, which require rapid processing to provide timely alerts. In Berlin, this challenge is reflected in the absence of a unified official pluvial flood hazard map for the entire city; instead, the city administration offers separate pluvial hazard maps for smaller, individual catchment areas within the city area. Previous studies that developed pluvial hydrodynamic models for the city have similarly focused on limited areas within the broader urban domain (e.g., (Tügel et al., 2025)). Furthermore, there is currently no unified, high-resolution pluvial flood model available for forecasting and early warning across the entire city. This limitation is not unique to Berlin; for example, in other German federal states such as North Rhine-Westphalia (NRW), operational pluvial forecasting relies on CPU-based 1D-2D models at 1-5 m resolution for selected urban catchments only, with city-wide high-resolution real-time simulations remaining infeasible due to runtimes of several hours to days (Landesamt f"ur Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen (LANUV), 2023).

Physically-based hydrodynamic models, often based on the two-dimensional Shallow Water Equations (SWE), have long been used in flood modeling, with applications demonstrated in various contexts e.g. (Pasculli et al., 2021; Apel et al., 2024; Khosh Bin Ghomash et al., 2022, 2025a). These models show potential for enhancing the accuracy and efficiency of early flood forecasting systems (Apel et al., 2022; Costabile et al., 2023; Khosh Bin Ghomash et al., 2024). However, their integration into early warning systems has been hindered by challenges in computational capacity, data assimilation, and real-time decision-making (De Almeida and Bates, 2013). Recent advancements in high-performance parallel-computing (HPC) systems have transformed this landscape, enabling detailed physically-based flood modeling with reduced computation times (Caviedes-Voullième et al., 2023; Khosh Bin Ghomash et al., 2024). Despite advancements in high-performance computing (HPC) and the adaptation of some hydrodynamic models to run on graphical processing units (GPUs) for significantly faster runtimes

compared to CPU-based simulations (Caviedes-Voullième et al., 2023), large domains—such as the Berlin case—often exceed the memory and processing capacity of a single GPU. As a result, harnessing multiple GPUs is essential not only to handle the computational load but also to ensure that simulations can run efficiently and within practical timeframes, especially for time-critical applications like flood forecasting. RIM2D, a multi-GPU-accelerated 2D hydraulic simulation model written in CUDA FORTRAN, leverages massive parallelization on HPC systems with CUDA-enabled NVIDIA GPUs. It solves a simplified version of the Shallow Water Equations (Bates et al., 2010) and has been successfully applied to flood simulations across diverse scenarios e.g. (Apel et al., 2024; Khosh Bin Ghomash et al., 2025a, b).

This study investigates whether high-performance computing (HPC)-enabled shallow water solvers—specifically RIM2D with multi-GPU acceleration—can achieve sufficient accuracy and lead time to support early flood warning systems over large urban domains such as the state of Berlin. Specifically, we evaluate whether such models can be run with sufficiently short computation times to support real-time, impact-based flood forecasting and whether multi-GPU configurations are essential to achieving this. The June 2017 pluvial flood event in Berlin, located in northeastern Germany, serves as the case study. We evaluate RIM2D's performance in simulating the flood under various spatial resolutions and GPU configurations. Model validation is conducted using volunteered geographic information (VGI) (Drews et al., 2023). In addition to the historical event, we simulate an 100-year return period (HQ100) rainfall scenario, used for generating pluvial flood hazard maps for the city. The resulting inundation patterns and flow properties are compared against the official hazard maps provided by the city to further assess the model's capability to support regulatory and planning applications. The analysis for both scenarios includes a detailed assessment of model runtimes and their implications for real-time forecasting, as well as a demonstration of how spatially explicit impact forecasts can enhance flood risk management by providing more actionable information across different impact dimensions.

2 Methods

2.1 Hydrodynamic ~~model~~Model: RIM2D

RIM2D is a two-dimensional hydrodynamic model developed by the Hydrology Section of the Helmholtz Centre for Geosciences (GFZ) in Potsdam, Germany. It employs a raster-based approach and is specifically designed to solve the local inertia form of the Shallow Water Equations, following the methodology proposed by Bates et al. (2010), which has demonstrated strong performance in flood inundation simulations (Falter et al., 2014; Neal et al., 2011; Apel et al., 2022, 2024; Khosh Bin Ghomash et al., 2025a). The model uses a regular grid of square cells, offering computational efficiency while preserving spatial uniformity across large areas.

The local inertia approximation enhances accuracy by including a momentum change term, distinguishing it from simpler models like the diffusive wave approach (De Almeida and Bates, 2013; Caviedes-Voullième et al., 2020). This term reflects how flow momentum evolves over time, linking each time step to the previous one through inertia. By omitting convective acceleration terms, the formulation simplifies flux calculations by decoupling the x- and y-directions:

$$\mathbf{F} = \begin{bmatrix} q_x \\ \frac{1}{2}gh^2 \\ 0 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 0 \\ q_y \\ \frac{1}{2}gh^2 \end{bmatrix}.$$

These fluxes are part of the shallow water equations:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} + \frac{\partial \mathbf{G}}{\partial y} = \mathbf{S}_b + \mathbf{S}_f,$$

with the state vector and source terms defined as:

$$\mathbf{U} = \begin{bmatrix} h \\ q_x \\ q_y \end{bmatrix}, \quad \mathbf{S}_b = \begin{bmatrix} 0 \\ -gh\frac{\partial z}{\partial x} \\ -gh\frac{\partial z}{\partial y} \end{bmatrix}, \quad \mathbf{S}_f = \begin{bmatrix} 0 \\ -\sigma_x \\ -\sigma_y \end{bmatrix}.$$

Here, h is the water depth [L], q_x and q_y are momentum components $[L^2/T]$, z is bed elevation [L], and σ_x, σ_y represent friction slopes computed using Manning's equation. g denotes gravitational acceleration $[L/T^2]$. This formulation balances simplicity and physical realism, capturing inertial behavior more accurately than diffusive models, while remaining computationally lighter than full dynamic solvers.

Although the explicit numerical scheme from Bates et al. (2010) can become unstable in near- or super-critical flow regimes or with fine grid resolutions (De Almeida and Bates, 2013), RIM2D incorporates the stabilization method introduced by de Almeida et al. (2012), based on numerical diffusion. An additional automatic stability control has been introduced to keep the model stable even if the numerical diffusion is not able to capture model instabilities, which can occur in case of super-critical flow conditions (Martins et al., 2015). This makes the model robust across diverse hydrodynamic conditions and spatial resolutions.

RIM2D is implemented in Fortran and optimized for GPU acceleration using NVIDIA CUDA Fortran, enabling efficient simulation of high-resolution or large-domain models. It supports multi-GPU computing and is distributed as open-source software under the European Union Public License version 1.2.

In order to utilize multiple GPUs, the 2D problem domain is divided into N equal parts using a 1D domain decomposition along the longer spatial dimension (typically x- or y-direction, depending on domain aspect ratio) and evenly distributed across the ~~GPUs.~~ N GPUs. This results in each GPU handling a horizontal or vertical 'slab' of the full 2D grid. Ghost zones have been introduced to manage interactions between neighbored cells that cross GPU boundaries. These zones need to be exchanged within each iteration. In order to improve the scalability of the multi-GPU implementation, various optimizations have been applied. These include the overlapping of communication and computation whenever possible - particularly during ghost zone exchange - as well as the utilization of specialized reduction kernels tailored to the problem domain and optimized for multi-GPU usage. These optimizations significantly reduce the overhead caused by data exchange and device synchronization.

However, the scaling might be impacted by load imbalances in the computation domain. The multi-GPU implementation is currently limited to a single compute node, with each GPU being controlled by a dedicated CPU thread. Additional threads are utilized to allow asynchronous reading and writing of input and output data during the simulation.

RIM2D is specifically designed for urban flood simulation considering drainage by the sewer system and infiltration. A capacity-based approach (Apel et al., 2024) is implemented to consider the effect of urban drainage on inundation, which is highly relevant for getting realistic inundation simulations. Spatially uniform or distributed sewer capacities and infiltration rates can be provided in mm/h. The sewer capacities are applied to grid cells classified as sealed (impervious), wheres the infiltration rates are applied to non-sealed (pervious) grid cells. Each grid cell can be be completely sealed/non-sealed or contain fractions of sealed/non-sealed areas. For infiltration saturated infiltration is ~~assume~~assumed, i.e. initially different infiltration rates at the start of the wetting of the ground are neglected.

2.2 Case Study

This study focuses on Berlin, the capital of Germany, which has a population of approximately ~~3.6 million~~ 3.89 million (as of end of 2024). The city is located in northeastern Germany along the River Spree. Berlin spans approximately 891.8 km^2 and has an average elevation of 34 meters above sea level. The city is characterized by a complex urban landscape with relatively flat topography (Seleem et al., 2022). The city experiences a moderate continental climate (Köppen classification Cfb) (Peel et al., 2007), characterized by warm, wet summers and cold, relatively dry winters, with distinct seasonal variations, with an average annual precipitation of around 570 mm (Berghäuser et al., 2021). Intense summer precipitation has caused several instances of urban pluvial flooding in recent decades, such as the event on June 29–30, 2017, which recorded up to 196.9 mm precipitation during a 24 hour period (Caldas-Alvarez et al., 2022). Such rainfall events, combined with Berlin's extensive urban infrastructure, heightens its susceptibility to pluvial flooding, where heavy rainfall overwhelms drainage systems, leading to surface water accumulation. For instance, the June 2017 event led to heavy runoff that accumulated in low-lying areas, resulting in widespread flooding throughout the city (Berghäuser et al., 2021). This was worsened by the city's flat topography, which limits storm-water drainage.

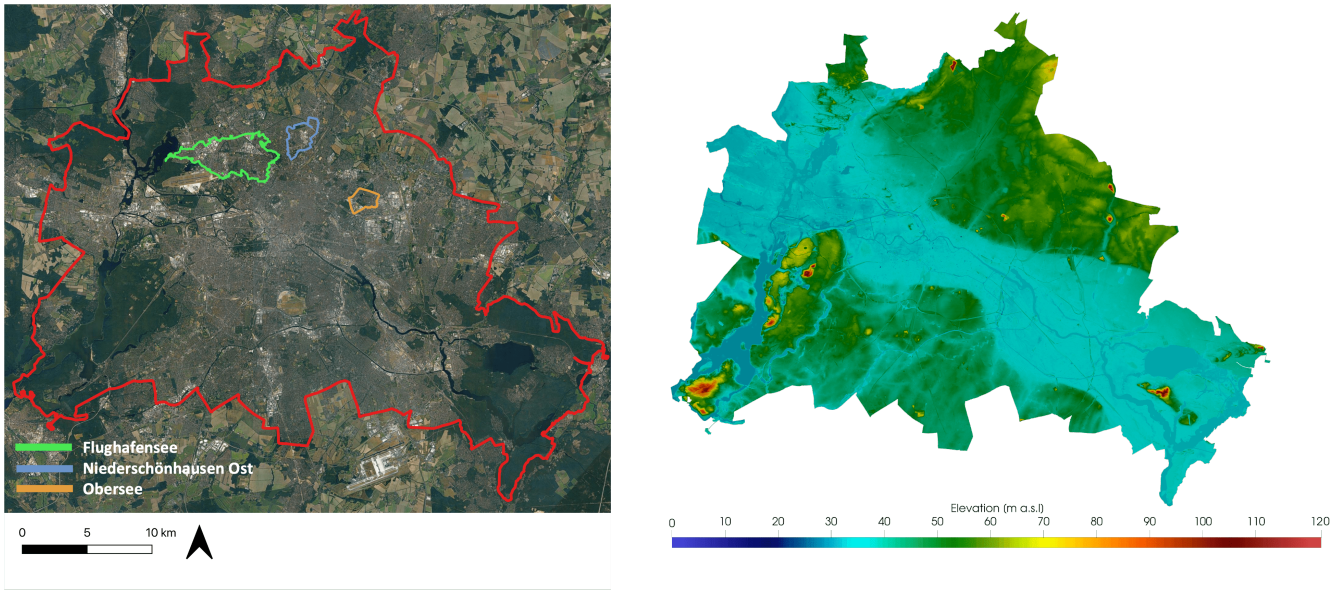

Figure 1. The red line in the left figure marks the administrative boundary of Berlin (Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen Berlin, 2022) and also the boundary of our simulations domain. The catchments Obersee (orange line), Flughafensee (green line) and Niederschönhausen Ost (blue line), defined by the city administration for the creation of pluvial hazard maps are also displayed. The right figure depicts the topography of the city. Satellite imagery © Google Earth 2025.

2.3 Data and ~~model set-up~~Model Set-up

This study employs the DGM1 digital elevation model (DEM) of the city Berlin, obtained from Geoportal Berlin/ATKIS®DGM (Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen Berlin, 2022), as the basis for the model setup. To further evaluate the performance of RIM2D at various spatial resolutions and to reduce the computational demand of the simulations, the DGM1 dataset was resampled to grid resolutions of dx = 2, 5, and 10 meters using the averaging technique. This resampling produced raster grids containing approximately 418.9 million cells at ~~a resolution of~~ dx = 2 ~~meters~~m, 67 million cells at dx = 5 ~~meters~~m, and 16.7 million cells at dx = 10 ~~meters across the study area.~~ m. These numbers correspond to the full rectangular computational domain (red line in Figure 1 left) that encloses the administrative boundary of Berlin plus a small buffer. Only cells inside the actual city mask are physically active during the simulations. Building outlines obtained from OpenStreetMap were used to remove building footprints from all three DEMs. Consequently, the building surfaces acted as closed reflective boundaries in the simulations.

Manning roughness values for the domain were assigned based on the 2020 land cover classification of Germany, which was derived from Sentinel-2 data (Riembauer et al., 2021). This classification uses atmospherically corrected Sentinel-2 imagery processed with the MAJA algorithm (Lonjou et al., 2016) and made available by the EOC Geoservice of the German Aerospace Centre (DLR). It incorporates training data from reference sources such as OpenStreetMap, alongside the Sentinel-2 imagery.

This land cover dataset was chosen for the study due to its relatively high spatial resolution of 10 meters. Manning values of 0.04, 0.03, 0.1, 0.025, 0.035, and 0.035 ~~m-1/3s~~ $[m^{-1/3}s]$ were assigned to the mapped land cover classes of vegetation, water bodies, forest, built-up areas, bare soil, and agricultural land, respectively.

Infiltration rates for the domain surface were determined by combining two distinct datasets. First, the soil type within the study area was classified using the dataset Umweltatlas Berlin/Bodenkundliche Kennwerte 2020, which is available for download at the cities online database, Geoportal Berlin (Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen Berlin, 2022). Saturated hydraulic conductivity values for each soil category of the first 1 m of the soil were assigned based on a thorough literature review. Next, the Imperviousness Density percentage raster, sourced from the Earth Observation component of the European Union's Copernicus program at a resolution of dx = 10 meters (European Environment Agency, 2020), was used to calculate the surface perviousness across the study domain. The final infiltration rates integrated into the model were derived by multiplying the saturated hydraulic conductivity of each cell, based on the Bodenkundliche Kennwerte 2020 product, by the surface perviousness percentage of that cell. RIM2D also allows for the integration of urban sewer drainage systems using a capacity-based approach, similar to the method applied for infiltration. This approach has been shown to effectively simulate drainage volumes, thereby reliably capturing the influence of sewer drainage on surface inundation (Apel et al., 2024).

The sewer capacity was estimated based on the assumption that sewer system dimensions in Germany are typically based on design rainfalls following a 2-year event of a duration of 15 minutes (Apel et al., 2024). The design rainfall values were taken from the KOSTRA database (Deutscher Wetterdienst (DWD), 2017) of the German Weather Service DWD. KOSTRA provides gridded data over whole Germany of precipitation and specific volume of precipitation of different return periods and durations, at a resolution of 5 km. For the 2-year return period and 15 minute duration the gridded values vary slightly within the city area of Berlin in the range of 11.9 mm to 13.2, with a mean of 12.5 mm (139 $l*s^{-1}*ha^{-1}$ specific volume). For this study the mean value was used to estimate a uniform sewer capacity (i.e. sewer design value) for Berlin based on the guide of the German Water Association DWA-A 118 (118E, 2006), considering mean slope of the city area and mean degree of surface sealing. This evaluates ~~in~~ a mean sewer system capacity of 21 mm/h, which was multiplied with the surface sealing in each grid cell to calculate the sewer drainage loss in the flood simulation. We also add clarification that the present simulations assume an initially empty sewer system, as no information on antecedent pipe filling was available for this event.

Although the RIM2D hydrodynamic model has been validated across a variety of flood scenarios, including both pluvial (Apel et al., 2024; Khosh Bin Ghomash et al., 2025a) and fluvial events (Khosh Bin Ghomash et al., 2024), this study intentionally conducts simulations in a blind manner, without parameter calibration (e.g., roughness coefficients). This choice is driven by the limited availability of validation data for the simulated event, which are chronically hard to obtain for these short term-high intensity pluvial flood events. Moreover, the focus of this study is to demonstrate the feasibility of using models like RIM2D for operational flood forecasting purposes—recognizing that model refinement and calibration is advisable for the actual use of the model in an actual operational forecast. However, due the chronical lack of calibration/validation data for this kind of events, relying on available spatial data and standard parameterizations is often the only viable option. While calibration is preferable, the common absence of calibration data and the occasional need for rapid model deployment in operational

settings make using uncalibrated models a common practice. Consequently, this study presents uncalibrated simulation results without engaging in extensive model calibration efforts.

To define the precipitation ~~boundary~~ input for simulating the June 2017 flood event in Berlin, rainfall radar data from the German Weather Service (DWD), known as "RADOLAN," was employed. RADOLAN is an operational, open-access product provided by the DWD and can be obtained through its Climate Data Center (CDC). The dataset offers nationwide coverage of Germany at a spatial resolution of 1,000 × 1,000 meters with hourly temporal resolution. For the model we use the hourly product which results in 48 hourly RADOLAN images corresponding to June 29 and 30, 2017—spanning the spatial extent of the study area—were used as ~~input boundary conditions~~rainfall input.

2.4 Model Validation

Observation data critical to understanding urban pluvial flooding—such as water level measurements—are often unavailable or, at best, extremely limited (Drews et al., 2023). This scarcity is largely due to the short-lived and highly localized nature of intense rainfall and the resulting pluvial floods, which significantly reduces the likelihood of capturing these events through operational remote sensing platforms. Even when such events are observed, challenges remain: optical sensors are hindered by cloud cover, while radar sensors can be affected by reflections from buildings. Furthermore, the infrequency of extreme rainfall events presents an additional challenge for generating reliable pluvial flood maps for model validation. As a result, observational data from actual pluvial flood events are typically both scarce and of limited quality, making the validation of advanced urban flood models particularly difficult. In fact, many locations that are theoretically prone to pluvial flooding lack any historical records altogether (Cea et al., 2025).

Due to the absence of systematic observations from the 2017 flood event in Berlin, validation of the model relied solely on volunteered geographic information (VGI), including photographs and videos captured by local residents during the event. Prior studies, such as Assumpção et al. (2018) and See (2019), suggest that although VGI datasets are often limited in scope, they can nonetheless offer valuable insights for the validation of urban-scale pluvial flood models. In this study, VGI sources at 19 different spots in the city are used to assess flood extents and inundation depths simulated by RIM2D for the 2017 event. To enable a more quantitative assessment, the photos and videos that provided the clearest and most measurable indications of water depth were selected. Reference objects such as street poles, building corners, and curb stones visible in the imagery were then manually measured on-site in the city. Based on these measurements, water depths were estimated and subsequently compared to the simulated water depths produced by RIM2D at the corresponding locations.

As an additional validation step, the study incorporates the Berlin city's official flood hazard maps, which are based on a 100-year return period rainfall event, corresponding to a uniform rainfall intensity of 100 mm/h sustained for one hour across the entire domain. The same ~~boundary condition~~ rainfall input is applied to the RIM2D model for consistency. The Berlin city's official flood hazard maps, derived from expert-validated hydrodynamic simulations, provide an additional reference point for evaluating the model's performance. Specifically, we compare flood extents and water depths from RIM2D outputs with those from the official city hazard maps. Berlin's official flood hazard maps were developed using a coupled 1D sewer and 2D surface runoff model and were subjected to a comprehensive plausibility review by the city authorities. This validation

process included a detailed assessment of simulated flood pathways and accumulation areas to identify any unrealistic flow behavior or implausible water pooling. In cases where discrepancies were identified, field inspections were carried out, and missing hydraulic features—such as barriers, channels, or culverts—were subsequently incorporated into the model. However, it is also noted in the documentation that these models were not subject to any calibration process. It should also be noted, that due to computational limitations and model constraints, the Berlin city authorities were only able to generate hazard maps for selected small catchment areas rather than for the entire city. In this study, we focus on three of those areas—Obersee, Flughafensee, and Niederschönhausen Ost—as shown in Figure 1. Since RIM2D is not constrained by such computational limits and provides results for the entire Berlin area, its output rasters were clipped to match the extents of these three regions for comparison.

To compare the inundated areas from RIM2D simulations with the official city flood maps, a comprehensive set of performance metrics was employed. These metrics were calculated by assessing the maximum inundation extents from RIM2D against the corresponding official maps. Each cell in the domain was classified according to Table 1, with RIM2D outputs treated as simulated data and the official maps considered as observed data. Based on this classification, a confusion map was generated for each comparison, providing counts for true positives, false negatives, false positives, and true negatives. These values were then used to compute several domain-wide flood performance metrics, as defined in Table 2. The metrics used in this study are based on methodologies presented by Wing et al. (2017) and Bernini and Franchini (2013).

		Simulated	
		Wet	Dry
Observed	Wet	True Positive (TP)	False Negative (FN)
	Dry	False Positive (FP)	True Negative (TN)

Table 1. Inundation confusion matrix (Wing et al., 2017; Bernini and Franchini, 2013). Each cell in the domain is compared between the simulation and observation and classified accordingly.

To evaluate the simulated water depths from RIM2D against those from the official city hazard maps, we employ two indicators: Bias and Root Mean Square Error (RMSE). The Bias is computed using the following equation:

$$\text{Bias} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)$$

and RMSE is calculated as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$$

In both equations:

- n is the total number of observations.
- $\hat{y}_i$ represents the simulated values.
- y_i represents the observed values.

Metric	Equation	Poor	Perfect	Description
Critical Success Index (CSI)	$\frac{TP}{TP+FP+FN}$	0	1	Ratio of correctly predicted wet cells to the total number of wet and missed wet cells.
Hit Rate (HR)	$\frac{TP}{TP+FN}$	0	1	Proportion of observed wet cells correctly identified by the simulation.
False Alarm (FA)	$\frac{FP}{TP+FP}$	1	0	Proportion of predicted wet cells that were not actually wet.
Error Bias (EB)	$\frac{FP}{FN}$	0 or ∞	1	Ratio of overpredicted wet cells to underpredicted wet cells.
Bias Percentage Indicator (BPI)	$100\left(\frac{TP+FP}{TP+FN}-1\right)$	-100 or 100	0	Relative percentage error in the total extent of the predicted flood area.

Table 2. Flood inundation performance metrics (Wing et al., 2017; Bernini and Franchini, 2013) used for evaluating the RIM2D simulations against the official flood maps.

3 Results and ~~discussion~~Discussion

3.1 Computational performance and runtime

We begin by assessing computational performance, as simulation runtime and resource efficiency are key factors in determining the feasibility of using this technology for flood forecasting and early warning systems. Figure 2 presents (a) the absolute simulation runtimes and (b) the ratio of simulated period to simulation runtime, for simulations at spatial resolutions of dx = 10, 5, and 2 meters over a 48-hour period during the June 29–30, 2017 flood event. Each resolution was tested using 1, 2, 4, 6, and 8 GPUs to demonstrate the performance benefits of RIM2D's multi-GPU processing capability for large-scale domains. All GPUs used in this study were NVIDIA A100 units.

As expected, simulation runtimes increase substantially with finer resolutions. For instance, at dx = 2 m simulations for the June 2017 event (48 hours simulation) require runtimes in the range of several hours, while at dx = 10 m simulations complete within minutes. Across all resolutions, increasing the number of GPUs significantly reduces runtimes, with the impact most pronounced at finer resolutions. At dx = 2 m, runtime drops from 794 minutes with 2 GPUs to 330 minutes (5.5 hours) with 8 GPUs. At dx = 5 m, runtime decreases from 116 minutes with 1 GPU to 63 minutes with 2 GPUs, and further down to 34 minutes with 8 GPUs. For dx = 10 m, runtimes are reduced from 16 minutes (1 GPU) to 8 minutes (8 GPUs). These results highlight the importance of multi-GPU configurations, particularly at higher resolutions, to achieve runtimes that are viable for real-time flood forecasting and early warning. It is also worth noting that for high-resolution simulations (dx = 2 m) over large domains such as the ~~city~~ state of Berlin, multi-GPU support is essential—not only for runtime efficiency but also due to the memory limitations of single-GPU setups. At this resolution, a single GPU is simply incapable of handling a simulation with such a large domain, which was also the case in this study. It should also be noted that simulations at a resolution of

dx = 1 m—corresponding to approximately 2 billion cells across the domain—would require more than 8 GPUs, a level of computational resources that was not available for this study.

Another noteworthy observation from the figures is that, for the dx = 5 m and 10 m configurations, increasing the number of GPUs beyond 4 yields only marginal reductions in runtime. For example, at dx = 5 m, increasing from 2 to 4 GPUs reduces the runtime from 63 to 39 minutes, whereas increasing from 4 to 6 GPUs results in a much smaller improvement—from 39 to 36 minutes. This indicates that, for simulations at dx = 5 and 10 m over the Berlin domain, using more than 4 GPUs offers limited additional benefit. A similar pattern is observed at dx = 2 m, where increasing beyond 6 GPUs provides only minor runtime gains, suggesting that more than 6 GPUs is unnecessary at this resolution. Scaling efficiency decreases beyond 4 GPUs for 5 m and 10 m resolutions (and beyond 6 GPUs at dx = 2 m) primarily due to increasing inter-GPU communication overhead (ghost-zone exchange) and load imbalances caused by heterogeneous wetting/drying patterns and building masks across sub-domains. We are actively working on further improvements such as including dynamic load balancing or adaptive ghost-zone compression to push scaling efficiency higher in future RIM2D releases.

Figure 2b illustrates how simulation runtimes compare to the real-time duration of the flood event. At a resolution of dx = 10 m, with over 16 million cells, RIM2D runs 178 times faster than real time using a single GPU, and up to 347 times faster with 8 GPUs. At dx = 5 m (over 67 million cells), the model runs 24 times faster than real time with 1 GPU, increasing to 84 times faster with 8 GPUs. Even at the finest tested resolution (dx = 2 m, with over 418 million cells), RIM2D achieves runtimes 3.6 times faster than real time using 2 GPUs, and up to 8.7 times faster with 8 GPUs. These results highlight RIM2D's capability to deliver real-time or faster-than-real-time performance, even at high spatial resolutions, making it highly suitable for integration into early warning systems. Its computational efficiency supports operational flood forecasting while maintaining appropriate lead times. In particular, the multi-GPU architecture of RIM2D enables high-resolution, impact-based flood simulations that allow for timely and informed emergency responses. Additionally, the estimation of secondary impact metrics—such as affected population, economic losses, risk of floating vehicles, disruption of transport networks, or exposure of critical infrastructure—can be completed in very short time periods-typically within seconds (Khosh Bin Ghomash et al., 2025a). This ensures that no significant computational burden is added during post-processing. In a real-time operational setting, updating a forecast would only require modifying the rainfall ~~boundary conditions~~ input based on the most recent radar or forecast data. The simulation would then produce updated flood and impact outputs accordingly.

The presented simulations of the June 2017 pluvial flood event in Berlin demonstrate the strong potential of integrating RIM2D into early warning systems. The obtained runtimes allow an integration into operational flood forecast systems. For example, driving RIM2D with numerical weather predictions (NWP) models, such as ICON-D2-EPS, which can provide high-resolution rainfall forecasts with lead times up to 17 hours (Najafi et al., 2024), would definitively be feasible. Additionally, the German Weather Service (DWD) provides RADALON radar-based nowcasts with lead times up to 2 hours—offering higher spatial accuracy and reliability, albeit with shorter lead times. But even when driven by short-term nowcast products, RIM2D's rapid runtimes ensure flood impact predictions can be generated and disseminated well before peak inundation occurs. This allows for timely evacuation warnings and infrastructure protection efforts, reinforcing RIM2D's role as a powerful tool in impact-based flood early warning systems. It is also important to highlight that no specific model or hardware optimizations

were applied to maximize performance in this scenario. For operational use, such optimizations would be essential and could further improve computational efficiency.

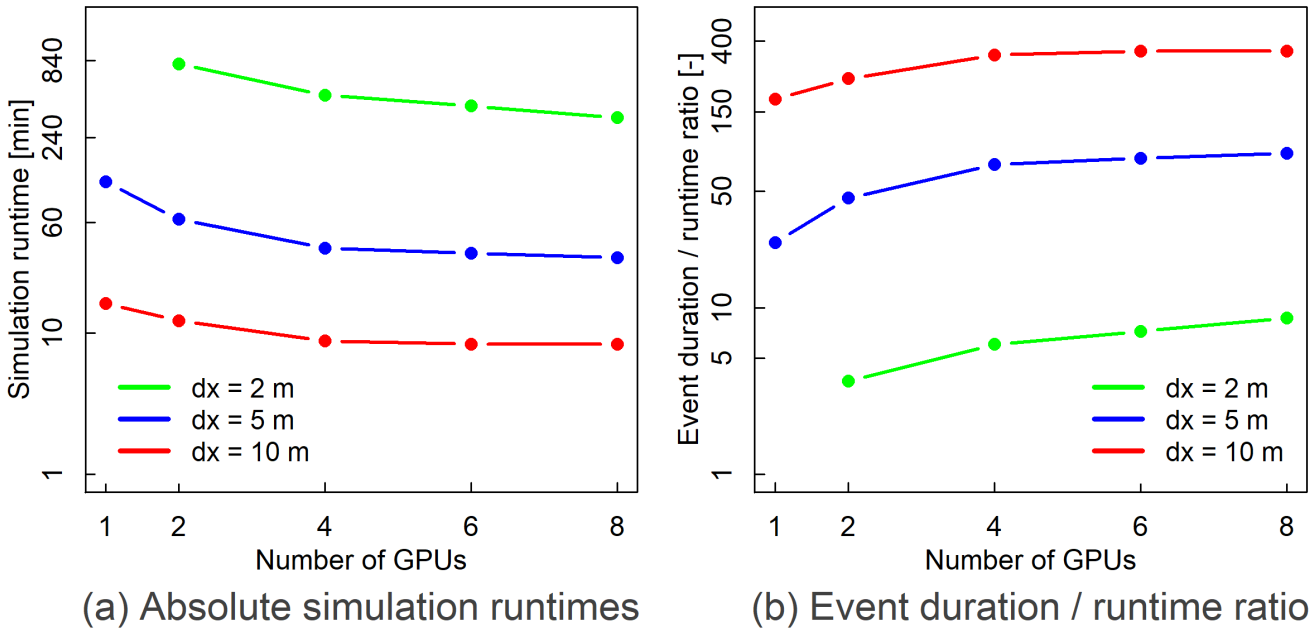

Figure 2. (a) Absolute simulation runtimes and (b) the ratio of simulation duration to runtime for dx = 10, 5, and 2 m resolutions during the June 2017 flood event (48-hour simulation), using 1, 2, 4, 6, and 8 GPU configurations.

Figure 3 displays the same performance metrics as Figure 2, but for the 1-hour, 100-year return period flood simulations for the entire state of Berlin. Overall, the observed trends closely mirror those in Figure 2. The key takeaway from this figure is that RIM2D's multi-GPU processing capabilities enable the generation of high-resolution flood hazard maps (even at dx = 2 m) for the entire Berlin area in under 30 minutes. This represents a significant advancement compared to current practices, where flood hazard maps are typically generated only for smaller catchments within ~~the~~ Berlin city due to computational limitations, rather than for the entire urban domain. With advanced tools such as RIM2D now available, capable of delivering high-resolution, ~~city-scale~~ state-scale flood simulations in operationally feasible time frames, it is ~~hughly~~ highly advisable that governmental agencies and city planners begin adopting such tools— not only to enhance the accuracy and coverage of flood hazard mapping but also enable truly operational early warning systems that support faster, data-driven decision-making and more effective risk mitigation.

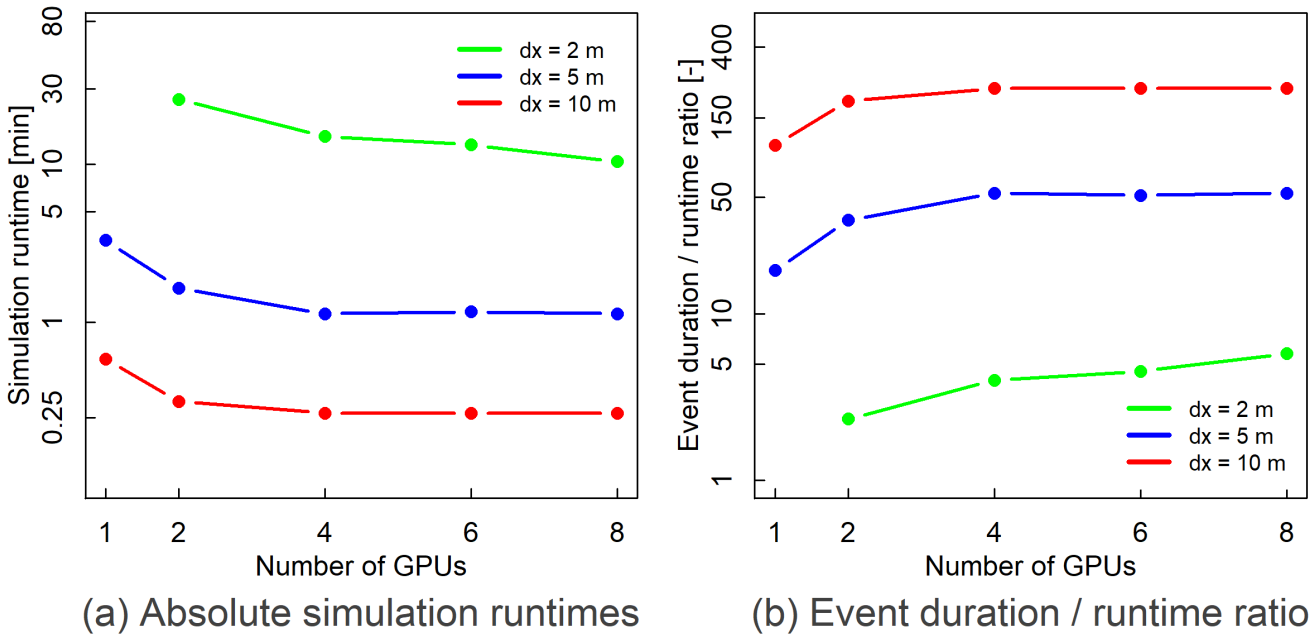

Figure 3. (a) Absolute simulation runtimes and (b) the ratio of simulation duration to runtime for dx = 10, 5, and 2 m resolutions for the HQ100 scenario (1-hour simulation), using 1, 2, 4, 6, and 8 GPU configurations.

3.2 Model Validation

For the June 2017 pluvial flood event in Berlin, there were no systematically collected or officially reported hydrological observations available, such as water levels, flow rates, or inundation extents, which could be used directly for model validation. In the absence of traditional observation data, this study utilized geo-located social media content—specifically, photos and videos shared by residents during and shortly after the event—as an alternative data source for validating the RIM2D simulation. This approach follows a growing body of research demonstrating the utility of Volunteered Geographic Information (VGI) for hydrological validation in urban environments where formal measurements are lacking (e.g., Drews et al., 2023).

A total of 19 locations across ~~the~~ Berlin city were identified where social media posts provided clear visual evidence of flooding. These media were manually geo-referenced and used to evaluate the accuracy of RIM2D in terms of flood extent and, where possible, estimated water depths. To support visual comparison, simulation results from RIM2D were exported and rendered in Google Earth. This visualization technique allowed for a direct spatial overlay between the simulated flood maps and the photographic evidence. Figure 4 presents two representative examples: the upper panel shows flooding along Hofjägerallee, a major street that leads to the Victory Column monument in central Berlin, while the lower panel captures inundation in Yorckstraße, a key transportation corridor in the southern part of the city. These visual comparisons highlight a strong spatial alignment between observed and simulated flood extents. Additional examples are provided in Figure A1 in Appendix A to illustrate the consistency of results across diverse urban contexts.

Figure 4. The red line indicates the administrative boundary of Berlin, while the red dots mark the 19 locations where photos or videos from the June 2017 flood were available and were compared to RIM2D simulation results. Two example comparisons are shown for Berlin Hofjägerallee (top) and Berlin Yorckstraße (bottom), illustrating how the dx = 5 m RIM2D simulation results align with observations from social media. The RIM2D outputs are visualized using Google Earth, with background imagery provided by Google Earth (Satellite imagery: © Google Earth 2025). Social media content used for validation is © X Corp. 2025 and reproduced under fair use for academic purposes.

In addition to qualitative comparisons of flood extent, a semi-quantitative validation of water depth was carried out at 16 of the 19 locations. These sites were selected based on the presence of identifiable reference objects—such as lampposts, building facades, street curbs, and bollards—that could be used to estimate inundation depth in the shared photos and videos. Physical measurements of these objects were subsequently performed in the field to enable approximate depth estimation at each location. These depth estimates were then compared to RIM2D simulation outputs at corresponding grid points. The results, presented in Figure 5, show that while RIM2D tended to slightly underestimate water depths in most cases, the differences were generally small and within an acceptable range for uncalibrated urban flood simulations. This indicates a strong baseline performance of the model despite the lack of calibration.

It is important to note that the current simulation setup did not incorporate detailed data on the city's drainage network, underground infrastructure, or elements such as underpasses and bridges, which can significantly influence local flood dynamics. The one bigger overestimation of inundation depth by RIM2D of about 40 cm shown in Figure 5 points towards a neglected hydraulic feature. Such detailed data were not available for this study but are typically accessible to municipal authorities. Their inclusion in future modeling efforts is expected to further enhance the accuracy of simulations, particularly in areas where hydraulic connectivity and localized flow pathways are governed by sub-surface or built infrastructure.

Despite these limitations, the validation exercise indicates that RIM2D is capable of reproducing both the spatial extent and relative depth of urban flooding with a high degree of reliability, even without calibration. The close agreement between simulated and observed outcomes—both in terms of flooded areas and estimated water depths—demonstrates the model's

potential as a practical and effective tool for real-time urban flood forecasting. It also reinforces the viability of VGI-based validation approaches, especially in urban contexts where traditional hydrometric networks are sparse or absent.

Taken together, these results suggest that RIM2D, when run with high-performance computing capabilities such as multi-GPU acceleration, can support impact-based flood forecasting and early warning in large and complex urban environments like Berlin. The findings underscore the importance of integrating HPC-enabled hydrodynamic modeling with non-traditional data sources to improve situational awareness and emergency response in the face of increasingly frequent and intense pluvial flood events.

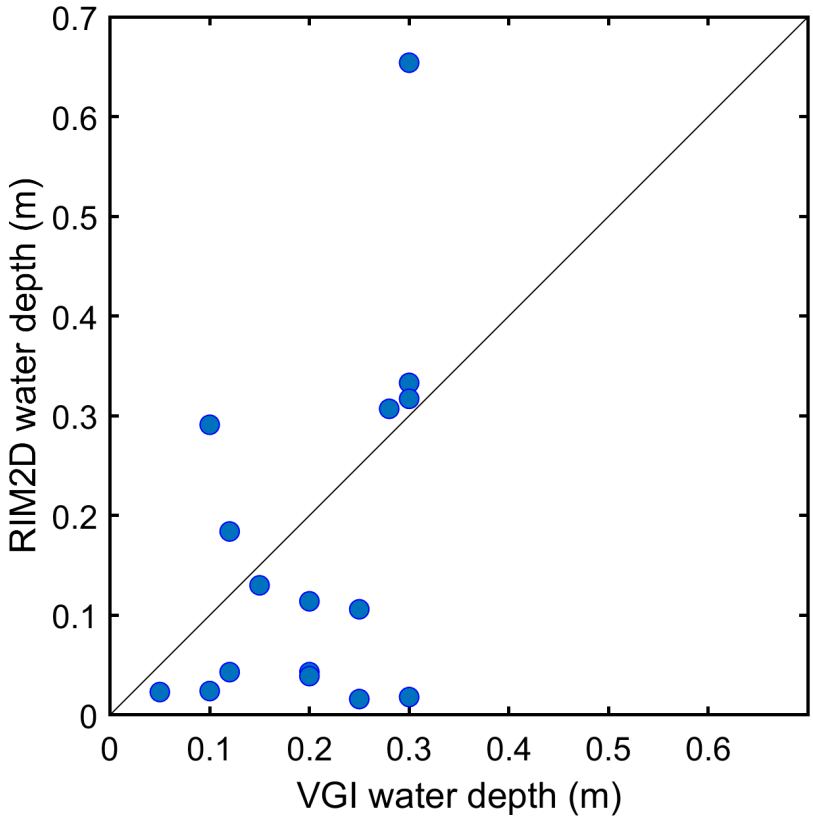

Figure 5. Validation of the dx = 5m June 2017 flood RIM2D simulation water depths at 16 VGI validation locations in Berlin using measured water marks. The figure shows RIM2D simulated water depths compared to estimated VGI water depths derived from field measurements.

We now turn to the simulation of the HQ100 event using RIM2D and its comparison with the city's official flood hazard maps. Figure 6 presents the comparison of RIM2D-simulated flood extents against the official maps at 2, 5, and 10-meter resolutions across the Obersee, Flughafensee, and Niederschönhausen Ost catchments. Evaluation metrics include the Critical Success Index (CSI), Hit Rate (HR), False Alarm Ratio (1–FA), Error Bias (EB), and Bias Percentage Indicator (BPI). Complementarily, Figure 7 presents the comparison of water depth predictions using Root Mean Square Error (RMSE) and Bias, offering insight into the accuracy of flow dynamics reproduced by RIM2D.

Overall, RIM2D demonstrates strong agreement with the official maps, achieving high scores across all performance metrics for both flood extent and depth. A clear resolution-dependent trend is observed, with finer-resolution inputs yielding improved predictive skills in comparison to the city hazard maps, although the resolution-based improvements show to be minor in almost all cases. Notably, RIM2D achieved this level of alignment without any site-specific calibration, whereas the official maps have undergone extensive plausibility assessments by the ~~city~~ state administration which involved a detailed qualitative and quantitative review of simulated flow paths and ponding areas to detect implausible hydraulic behavior. Moreover as mentioned before, the current RIM2D simulation setup did not account for detailed information on the city's drainage system (i.e. spatially distributed sewer capacities in RIM2D), underground infrastructure (car parks, tunnels, etc.), or features like underpasses and bridges—factors that can strongly influence local flood behavior. While this data was unavailable for the present study, it is generally accessible through municipal sources and incorporating such details in future modeling efforts would likely improve simulation accuracy, especially in areas where subsurface or built structures control flow dynamics.

What sets RIM2D apart however is not only its strong agreement with these rigorously verified maps but also its unparalleled computational efficiency. Unlike the ~~city's~~ states modeling framework, which—due to computational constraints—cannot generate a statewide flood hazard map in a single simulation, RIM2D is capable of simulating high-resolution flood extents over the entire urban domain in a matter of minutes, as seen in Figure 3. This opens the door to ensemble-based flood hazard assessments using large numbers of synthetic rainfall scenarios generated via stochastic weather generators. Such probabilistic mapping provides a more robust basis for flood risk planning than traditional deterministic approaches (Bodoque et al., 2023; Khosh Bin Ghomash et al., 2023). Another key advantage of using RIM2D is the ability to implement a statewide flood model. This allows the model to capture inundation that extends beyond the boundaries of small catchments, which are typically the focus of official flood hazard maps in Berlin. Floodwaters can move across these hydrological boundaries in both directions—either entering a catchment from adjacent areas or spreading out from within the catchment into neighboring regions. This effect is particularly significant in flat landscapes, where minimal elevation differences allow even shallow floodwaters to overflow natural or man-made catchment divides. Such cross-boundary flow highlights the value of running simulations at the state scale, as it provides a more comprehensive understanding of flood dynamics that smaller-scale models might miss.

In this context, RIM2D presents a valuable tool not only for real-time flood forecasting but also strategic urban resilience planning, enabling authorities to efficiently assess a wide range of future flood scenarios~~under deep uncertainty~~.

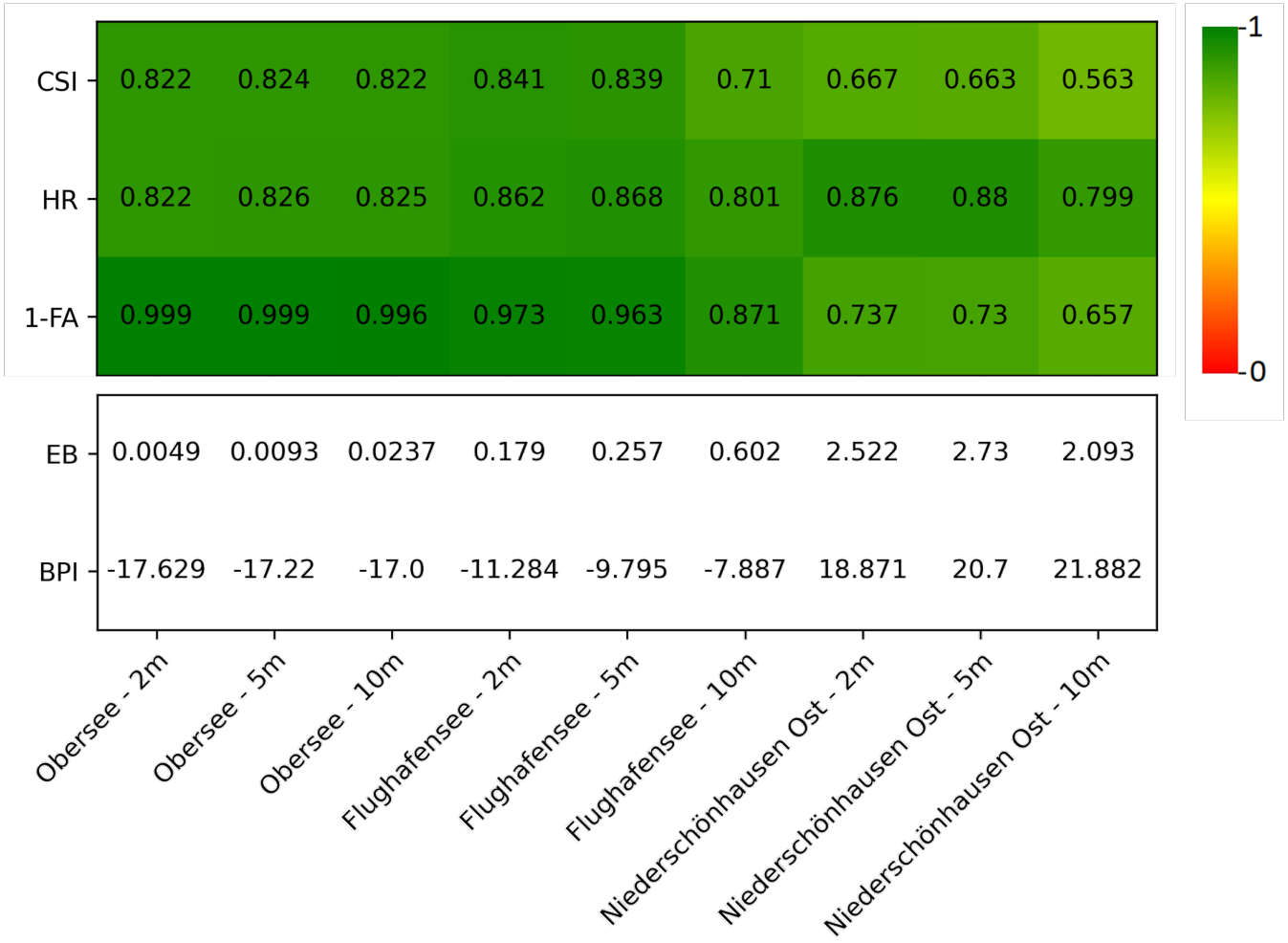

Figure 6. Comparison of RIM2D-simulated flooded areas with official city flood maps at spatial resolutions of 2, 5, and 10 meters for a 100-year flood event in the Obersee, Flughafensee, and Niederschönhausen Ost catchments. Evaluation metrics include the Critical Success Index, Hit Rate, False Alarm Ratio, Error Bias, and Bias Percentage Indicator.

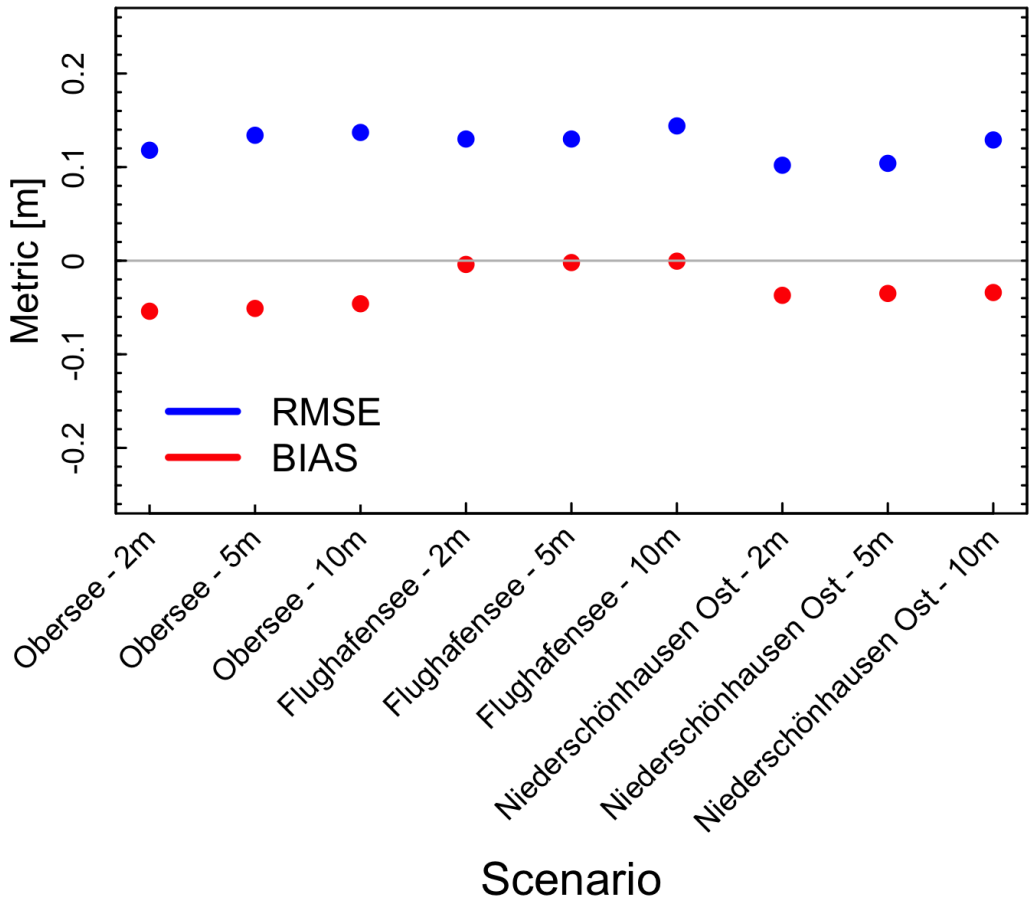

Figure 7. Comparison of RIM2D-simulated water depths with official city flood maps at spatial resolutions of 2, 5, and 10 meters for a 100-year flood event in the Obersee, Flughafensee, and Niederschönhausen Ost catchments. Evaluation metrics include Root Mean Square Error (RMSE) and Bias.

3.3 Flood Impacts

We now turn our attention to the simulation of the June 2017 pluvial flood event in Berlin. The results are presented in Figure 8, which includes maps of (a) maximum water depth, (b) maximum flow velocity, (c) the time at which each grid cell reaches its maximum water depth, and (d) the number of affected individuals within the domain. These outputs are based on the simulation with a spatial resolution of 5 meters. The figure demonstrates RIM2d's capability to spatially represent impact-based outcomes of the 2017 flood. From Figure 8a, it can be seen that most areas of the city experienced water depths between 5 and 20 cm, with localized hot-spots reaching depths of approximately 50 cm. These values are consistent with photographic and video documentation of the event. Regarding flow velocities (Figure 8b), generally low velocities are observed throughout the domain—a typical characteristic of pluvial flooding in relatively flat urban areas like Berlin. Figure 8c

presents the time of which cell experiences max water depth [s] during the simulation. This indicator can be vital for providing insights into how quickly different areas experience peak flood severity, which is essential for evaluating potential damage to critical infrastructure (Bachmann et al., 2021), estimating exposure durations, and planning targeted mitigation strategies. Lastly, Figure 8d shows the affected people during the flood. Individuals are considered "affected" if they reside in areas impacted by flooding—essentially, those who would "get wet feet." This definition is consistent with established literature in flood risk assessment, where affected population is defined as those living in cells with positive water depth in a given flood scenario e.g. (Winsemius et al., 2013; Ward and de Bruin, 2018). The potential impact is estimated using population density data [persons/m^2] assigned to each grid cell within the simulation domain. Vulnerability is defined by the level of impact, which is determined by the maximum water depth [m] in each cell. To reduce uncertainty associated with very shallow flooding, a threshold of 10 cm is applied: if the maximum water depth in a cell is below 0.1 m, the population in that cell is not considered affected; if the depth exceeds 0.1 m, all individuals in the cell are assumed to be affected. For this analysis, we use the 2020 WorldPop population density dataset for Germany (WorldPop, 2020), with a spatial resolution of 100 meters. The RIM2D simulations max water depths are aggregated to the WorldPop grid resolutions and overlayed with the data. A WorldPop cell is classified as affected if the maximum simulated water depth within it exceeds 0.1 m ("wet feet" threshold). The entire population assigned to that cell by WorldPop is then counted as affected. Values below 1 person may arise from the WorldPop data because many cells, especially sparsely populated areas such as parks, industrial zones, or outskirts, contain only a fraction of one person according to the WorldPop disaggregated data.

While looking at the figure, a higher concentration of affected individuals (indicated in dark green) can be seen in the city center compared to the outskirts, primarily due to the higher population density in central Berlin. It is also worth noting that RIM2D simulation outputs enable the assessment of additional impact dimensions. For example, RIM2D allows users to generate human instability maps and identify areas where vehicles may begin to float, following the method proposed by Jonkman and Penning-Rowsell (2008), based on the simulation outputs. These analyses are particularly useful in flash flood or fluvial scenarios with high flow velocities, where risks to human safety and vehicle stability are more pronounced. However, in the case of Berlin, due to the characteristics of the pluvial flood and the resulting low water velocities indicators such as human instability and vehicle floatation were not computed, as the associated risks were considered negligible under the observed flow conditions.

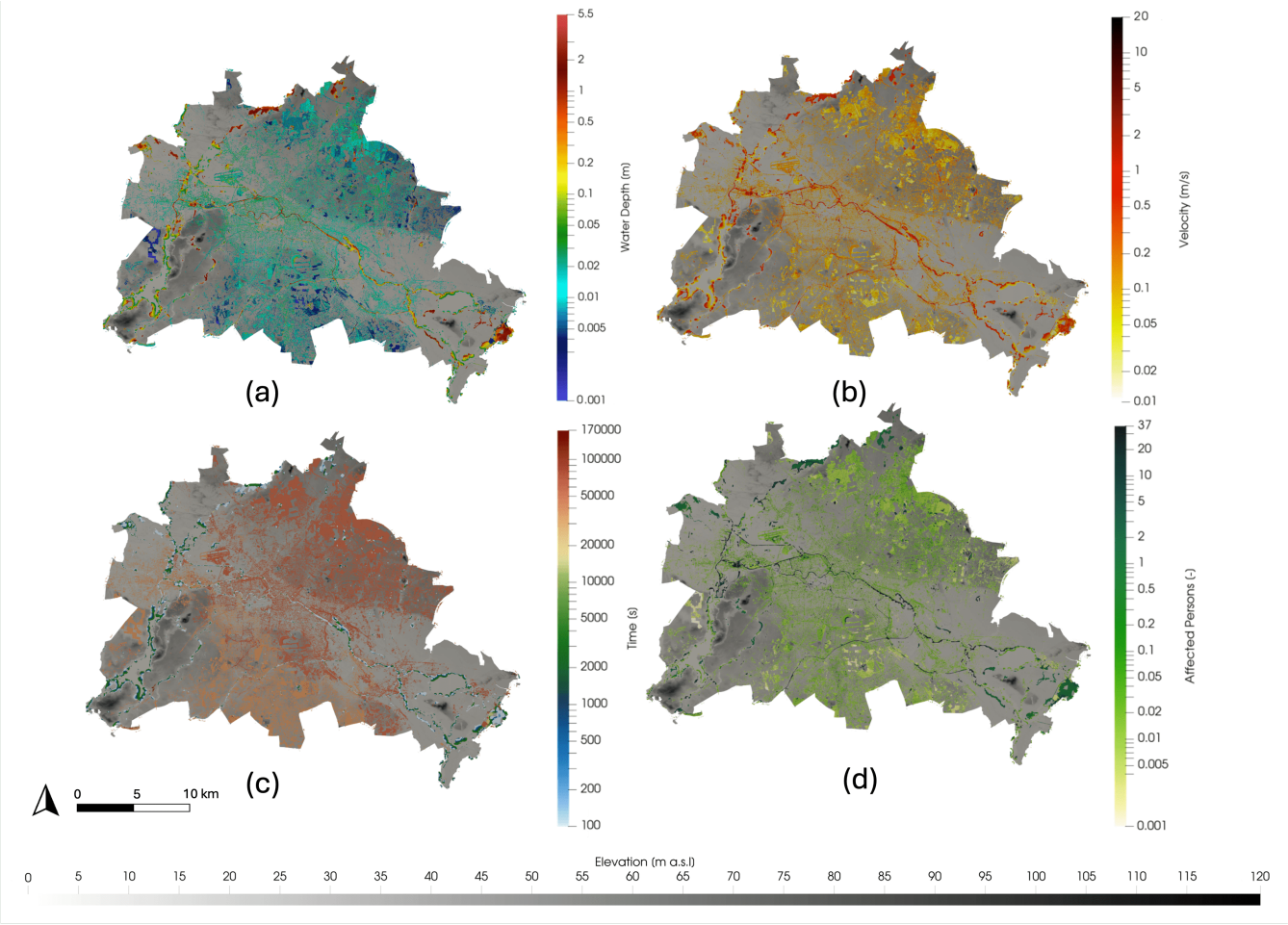

Figure 8. (a) Simulated max water depths [m], (b) max velocities [m/s], (c) time of max water depth [s] and (d) affected population [number of ~~Persons~~persons] for the June 2017 flood event at dx = 5 m.

3.4 Uncertainties

As with all model simulations, inherent uncertainties are expected in the results. The uncalibrated hydrodynamic model presented in this study inevitably contributes to the uncertainties already present in meteorological forecasts. Nevertheless, hydrodynamic modeling typically introduces the least uncertainty compared to other sources (de MOEL and Aerts, 2011). This uncertainty could be further reduced by avoiding ad hoc model configurations and uncalibrated simulations—as done in this feasibility study—and instead pursuing a more robust implementation. Such an approach would involve calibration and/or validation using historical flood events or supplementary data sources, such as volunteered geographic information (VGI) (Drews et al., 2023). With the expertise and collaboration of the ~~city's~~ states relevant authorities, implementing a more refined model would be entirely feasible.

Results are subject to typical uncertainties in parameters and model inputs such as Manning roughness (friction), infiltration rates, sewer drainage capacity, DEM accuracy, and rainfall input, which usually alter flooded area by ±10–20 percent and water depths by ±15–25 percent for realistic parameter ranges, as shown in sensitivity studies with RIM2D and other comparable high-resolution urban models (Khosh Bin Ghomash et al., 2025b; Apel et al., 2024; Shrestha et al., 2022; Alipour et al., 2022). As an example for the case of Berlin, to probe sensitivity to sewer capacity, we ran two additional simulations with sewer capacities reduced by 50 and 100 percent (assuming partial or full blockage or saturation from prior events). This resulted in an estimated increases in flooded areas by approximately 8 and 13 percent and average water levels in the domain by approximately 17 and 23 percent respectively compared to the base case, highlighting the importance of accurate sewer parameterization in urban pluvial modeling. Given the demonstrated runtimes of RIM2D, operational ensemble forecasting exploring these uncertainties is achievable (e.g., hundreds or thousands of Monte Carlo simulations can be completed in hours on multi-GPU setups, as demonstrated in (Khosh Bin Ghomash et al., 2025b)). A comprehensive sensitivity study or calibrated risk assessment for Berlin is, however, beyond the scope of this paper, whose aim is to demonstrate that state-scale, high-resolution, real-time pluvial flood forecasting with a physically based model is now technically feasible with multi-GPU processing. For operational deployment of RIM2D, it is advisable to incorporate uncertainty maps derived from ensembles of meteorological forecasts, something which is made possible by RIM2Ds efficient computation runtimes. This strategy enables clearer visualization of the uncertainties inherent in the input data. Given RIM2D's computational efficiency, these forecast ensembles can be rapidly translated into inundation maps that explicitly represent flood forecast uncertainty (Alfonso et al., 2016; Bodoque et al., 2023). Such maps not only illustrate the range of possible outcomes but also enhance decision-making under uncertainty by supporting risk-informed planning and emergency response (Bodoque et al., 2023). Furthermore, they offer a valuable tool for communicating risk to non-technical stakeholders, helping to improve public understanding and stakeholder engagement.

4 Conclusions

This study highlights the capabilities of the RIM2D hydrodynamic model, emphasizing its multi-GPU computation features that support its application in flood impact forecasting, early warning systems, and urban planning for large metropolitan areas. Using the June 2017 pluvial flood in Berlin, Germany, and standardized rainfall scenarios, we demonstrate that RIM2D delivers reliable simulations of inundation extent and flow dynamics with highly efficient runtimes—making timely flood forecasts and warnings feasible. Crucially, achieving such performance at high resolutions across a large domain like Berlin is not possible without multi-GPU computation. In addition, RIM2D outputs can be used to generate spatially distributed impact indicators, offering detailed, site-specific insights into potential flood impacts. These outputs go beyond traditional flood depth maps, supporting more targeted disaster management and enabling more informative public alerts, which could substantially enhance current flood response and warning systems.

The use of multiple graphics processing units (GPUs) optimized for parallel computing allows RIM2D to achieve simulation runtimes suitable for operational flood forecast systems. In this study, we show that for the 2017 Berlin event, RIM2D

can generate flood forecasts 347× faster than real-time at dx = 10 m resolution, 84× at dx = 5 m, and 8.7× at dx = 2 m, using up to eight high-performance GPUs. Specifically, simulations at dx = 10 m and 5 m resolution complete in 8 and 34 minutes, respectively—well within the timeframe of an short-term operational forecast cycle with radar nowcasts. Comparable speeds are also achievable through large CPU HPC clusters (>1000 cores), but multi-GPU setups currently offer 5–10× better performance-per-watt and lower operational costs for this class of urban-domain simulations (Caviedes-Voullième et al., 2023). These speeds are only achievable through multi-GPU setups; using a single GPU, the dx = 5 m simulation takes nearly two hours, and simulations at dx = 2 m or finer are not feasible. This performance advantage is especially critical for flood scenarios, where the timing and accuracy of warnings significantly affect damage mitigation and public safety. In the context of real-time prediction, several promising machine learning, deep learning, and artificial neural network approaches also exist as complementary alternatives, often achieving sub-second runtimes through emulators trained on hydrodynamic simulations, though they may require extensive training data and lack full physical interpretability (Zeng et al., 2025).

Traditional, flood forecast and early warning systems rely heavily on rainfall forecasts and often provide coarse, district-level alerts with limited spatial detail. However, events like the Berlin flood highlight the limitations of such systems, especially when accurate water depth and inundation maps are essential for targeted response. RIM2D addresses this gap by enabling high-resolution simulations of flood extent and water levels. Given its simple model configuration and the broad availability of input data, RIM2D offers strong potential for broader adoption in real-time flood forecasting and emergency planning. The hardware required to run RIM2D simulations is relatively affordable, especially compared to traditional large-scale computing clusters. This cost-efficiency makes it feasible for flood forecast centers to adopt the model without significant investment in complex IT infrastructure. Furthermore, most major cloud-based HPC providers—such as Google, Amazon, and Deutsche Telekom—offer GPU services that support multi-GPU computing, allowing technical infrastructure to be outsourced at relatively low cost. Given the model's validated performance, straightforward implementation, and fast runtimes, we advocate for enhancing current forecasting practices by integrating high-performance hydraulic models like RIM2D and the derived flood impact forecasts. Doing so could significantly improve the precision and usefulness of flood warnings. The scientific groundwork, methods, and tools for operationalizing such impact forecasts are already mature and ready for deployment.

In conclusion, this study demonstrates that the technology is now mature enough for integration into early warning systems, even in large urban areas like Berlin. It provides sufficient lead time for emergency response and delivers far more detailed and actionable information than traditional flood warning systems. High-resolution, time-sensitive data on flood depth, velocity, and other impact indicators offer clearer insights into potential flood severity and support more accurate impact assessments. Additionally, this type of information is significantly more accessible and understandable for the public, emergency managers, and decision-makers than alerts based solely on parameters like rainfall forecasts (Thieken et al., 2023).

Code availability. RIM2D is open-source for scientific use under the EUPL1.2 license. Access is granted upon request. The simulations were performed with version 0.2.

Data availability. The OSM building shapefiles used in this study are freely available at https://download.geofabrik.de/europe/germany.html. The land cover raster, employed to assign surface roughness values within the simulation domain, can be accessed at https://www.mundialis.de/en/germany-2020-land-cover-based-on-sentinel-2-data/. Digital elevation model of Berlin and the city official flood hazard maps can be accessed through the Geoportal Berlin (Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen Berlin, 2022). The Imperviousness Density map from the Copernicus Land Monitoring Service (CLMS) used for infiltration rates calculation can be found at European Environment Agency (2020)

Appendix A: Supplementary Material: Examples of model comparison with VGI data

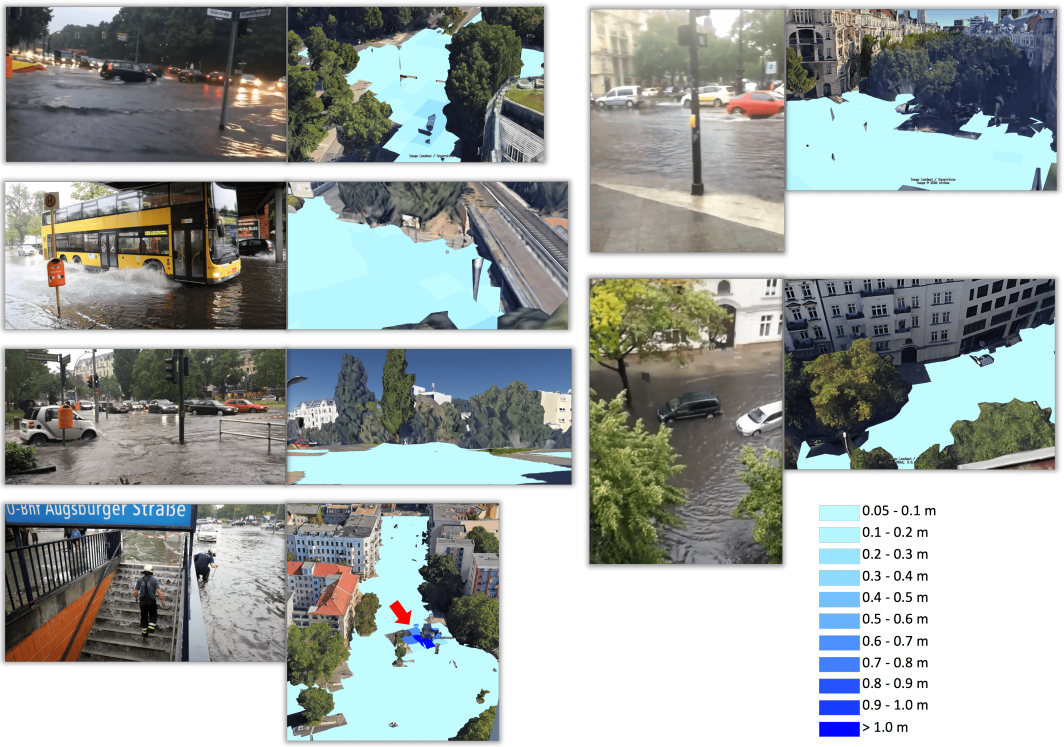

Figure A1. Examples of comparisons between the dx = 5 m RIM2D simulation results to social media posts. The RIM2D outputs are visualized using Google Earth, with background imagery provided by Google Earth (Satellite imagery: © Google Earth 2025). Social media content used for validation is © X Corp. 2025 and reproduced under fair use for academic purposes.

Author contributions. SKBG: Conceptualization, Methodology, Investigation, Simulation, Software, Analysis, Visualization, Model validation, Writing; SD: Visualization, Model validation, Review; HA: Review

Competing interests. The authors declare that they don't have any competing interest within the frame of the content of the manuscript.

Acknowledgements. This ~~research was performed within the frame of the DIRECTED project (https://directedproject.eu/). Funding of the DIRECTED project within the European Union's Horizon Europe – the Framework Programme for Research and Innovation (grant agreement No. 101073978, HORIZON-CL3-2021-DRS-01) is gratefully acknowledged. This~~ work utilized high-performance computing resources made possible by funding from the Ministry of Science, Research and Culture of the State of Brandenburg (MWFK) and are operated by the IT Services and Operations unit of the Helmholtz Centre Potsdam. We thank Johannes Spazier for his helpful consultations on minor aspects of the multi-GPU implementation.

References

118E, D.-A.: Hydraulic dimensioning and verification of drain and sewer systems, 2006.

Alfonso, L., Mukolwe, M., and Di Baldassarre, G.: Probabilistic flood maps to support decision-making: Mapping the value of information, Water Resources Research, 52, 1026–1043, 2016.

Alipour, A., Jafarzadegan, K., and Moradkhani, H.: Global sensitivity analysis in hydrodynamic modeling and flood inundation mapping, Environmental Modelling & Software, 152, 105 398, 2022.

Apel, H., Vorogushyn, S., and Merz, B.: Brief communication: Impact forecasting could substantially improve the emergency management of deadly floods: case study July 2021 floods in Germany, Natural Hazards and Earth System Sciences, 22, 3005–3014, 2022.

Apel, H., Benisch, J., Helm, B., Vorogushyn, S., and Merz, B.: Fast urban inundation simulation with RIM2D for flood risk assessment and forecasting, Frontiers in Water, 6, 1310 182, 2024.

Assumpção, T. H., Popescu, I., Jonoski, A., and Solomatine, D. P.: Citizen observations contributing to flood modelling: Opportunities and challenges, Hydrology and Earth System Sciences, 22, 1473–1489, 2018.

Bachmann, D., Khosh Bin Ghomash, S., and Schotten, R.: Neue entwicklungen in der hochwasserrisikoanalyse: Niederschlagsgeneratoren und kritische infrastrukturen, WasserWirtschaft, 111, 32–38, 2021.

Bates, P. D., Horritt, M. S., and Fewtrell, T. J.: A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling, Journal of Hydrology, 387, 33–45, https://doi.org/10.1016/j.jhydrol.2010.03.027, 2010.

Berghäuser, L., Schoppa, L., Ulrich, J., Dillenardt, L., Jurado, O. E., Passow, C., Samprogna Mohor, G., Seleem, O., Petrow, T., and Thieken, A.: Starkregen in Berlin: Meteorologische Ereignisrekonstruktion und Betroffenenbefragung, 2021.

Bernini, A. and Franchini, M.: A rapid model for delimiting flooded areas, Water resources management, 27, 3825–3846, 2013.

Bodoque, J. M., Esteban-Muñoz, Á., and Ballesteros-Cánovas, J. A.: Overlooking probabilistic mapping renders urban flood risk management inequitable, Communications Earth & Environment, 4, 279, 2023.

Caldas-Alvarez, A., Augenstein, M., Ayzel, G., Barfus, K., Cherian, R., Dillenardt, L., Fauer, F., Feldmann, H., Heistermann, M., Karwat, A., et al.: Meteorological, impact and climate perspectives of the 29 June 2017 heavy precipitation event in the Berlin metropolitan area, Natural Hazards and Earth System Sciences Discussions, 2022, 1–39, 2022.

Caviedes-Voullième, D., Fernández-Pato, J., and Hinz, C.: Performance assessment of 2D Zero-Inertia and Shallow Water models for simulating rainfall-runoff processes, Journal of hydrology, 584, 124 663, 2020.

Caviedes-Voullième, D., Morales-Hernández, M., Norman, M. R., and Özgen-Xian, I.: SERGHEI (SERGHEI-SWE) v1. 0: a performance-portable high-performance parallel-computing shallow-water solver for hydrology and environmental hydraulics, Geoscientific Model Development, 16, 977–1008, 2023.

Cea, L., Sañudo, E., Montalvo, C., Farfán, J., Puertas, J., and Tamagnone, P.: Recent advances and future challenges in urban pluvial flood modelling, Urban Water Journal, pp. 1–25, 2025.

Costabile, P., Costanzo, C., Kalogiros, J., and Bellos, V.: Toward Street-Level Nowcasting of Flash Floods Impacts Based on HPC Hydrodynamic Modeling at the Watershed Scale and High-Resolution Weather Radar Data, Water Resources Research, 59, https://doi.org/10.1029/2023wr034599, 2023.

De Almeida, G. A. and Bates, P.: Applicability of the local inertial approximation of the shallow water equations to flood modeling, Water Resources Research, 49, 4833–4844, 2013.

de Almeida, G. A., Bates, P., Freer, J. E., and Souvignet, M.: Improving the stability of a simple formulation of the shallow water equations for 2-D flood modeling, Water Resources Research, 48, 2012.

de MOEL, H. and Aerts, J.: Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates, Natural Hazards, 58, 407–425, 2011.

Deutscher Wetterdienst (DWD): KOSTRA-DWD-2010R: Grids of Return Periods of Heavy Precipitation (Design Precipitation) over Germany, https://www.dwd.de/DE/leistungen/kostra_dwd_rasterwerte/kostra_dwd_rasterwerte.html, accessed: 2025-05-19, 2017.

Drews, M., Steinhausen, M., Larsen, M. A. D., Dømgaard, M. L., Huszti, L., Rácz, T., Wortmann, M., Hattermann, F. F., and Schröter, K.: The utility of using Volunteered Geographic Information (VGI) for evaluating pluvial flood models, Science of the Total Environment, 894, 164 962, 2023.

European Environment Agency: Imperviousness Density 2018 (raster 10 m), Europe, 3-yearly, https://doi.org/10.2909/3bf542bd-eebd-4d73-b53c-a0243f2ed862, accessed: 2025-05-16, 2020.

Falter, D., Dung, N., Vorogushyn, S., Schröter, K., Hundecha, Y., Kreibich, H., Apel, H., Theisselmann, F., and Merz, B.: Continuous, large-scale simulation model for flood risk assessments: proof-of-concept, Journal of Flood Risk Management, 9, 3–21, https://doi.org/10.1111/jfr3.12105, 2014.

Jonkman, S. and Penning-Rowsell, E.: Human instability in flood flows 1, JAWRA Journal of the American Water Resources Association, 44, 1208–1218, 2008.

Khosh Bin Ghomash, S., Caviedes-Voullieme, D., and Hinz, C.: Effects of erosion-induced changes to topography on runoff dynamics, Journal of Hydrology, 573, 811–828, 2019.

Khosh Bin Ghomash, S., Bachmann, D., Caviedes-Voullième, D., and Hinz, C.: Impact of rainfall movement on flash flood response: A synthetic study of a semi-arid mountainous catchment, Water, 14, 1844, 2022.

Khosh Bin Ghomash, S., Bachmann, D., Caviedes-Voullième, D., and Hinz, C.: Effects of Within-Storm Variability on Allochthonous Flash Flooding: A Synthetic Study, Water, 15, 645, 2023.

Khosh Bin Ghomash, S., Apel, H., and Caviedes-Voullième, D.: Are 2D shallow-water solvers fast enough for early flood warning? A comparative assessment on the 2021 Ahr valley flood event, Natural Hazards and Earth System Sciences, 24, 2857–2874, https://doi.org/10.5194/nhess-24-2857-2024, 2024.

Khosh Bin Ghomash, S., Apel, H., Schröter, K., and Steinhausen, M.: Rapid high-resolution impact-based flood early warning is possible with RIM2D: a showcase for the 2023 pluvial flood in Braunschweig, Natural Hazards and Earth System Sciences, 25, 1737–1749, https://doi.org/10.5194/nhess-25-1737-2025, 2025a.

Khosh Bin Ghomash, S., Yeste, P., Apel, H., and Nguyen, V. D.: Monte Carlo-based sensitivity analysis of the RIM2D hydrodynamic model for the 2021 flood event in western Germany, Natural Hazards and Earth System Sciences, 25, 975–990, 2025b.

Kreibich, H., Hudson, P., and Merz, B.: Knowing what to do substantially improves the effectiveness of flood early warning, Bulletin of the American Meteorological Society, 102, E1450–E1463, 2021.

Landesamt f"ur Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen (LANUV): Jahresbericht 2023, Tech. rep., Landesamt f"ur Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, Recklinghausen, Germany, https://www.lanuk.nrw.de/fileadmin/lanuvpubl/2_jahresberichte/LANUV-Jahresbericht_2023.pdf, accessed: December 19, 2025, 2023.

Lonjou, V., Desjardins, C., Hagolle, O., Petrucci, B., Tremas, T., Dejus, M., Makarau, A., and Auer, S.: Maccs-atcor joint algorithm (maja), in: Remote Sensing of Clouds and the Atmosphere XXI, vol. 10001, pp. 25–37, SPIE, 2016.

Martins, R., Leandro, J., and Djordjević, S.: A well balanced Roe scheme for the local inertial equations with an unstructured mesh, Advances in Water Resources, 83, 351–363, 2015.

Merz, B., Kuhlicke, C., Kunz, M., Pittore, M., Babeyko, A., Bresch, D. N., Domeisen, D. I., Feser, F., Koszalka, I., Kreibich, H., et al.: Impact forecasting to support emergency management of natural hazards, Reviews of Geophysics, 58, e2020RG000 704, 2020.

Najafi, H., Shrestha, P. K., Rakovec, O., Apel, H., Vorogushyn, S., Kumar, R., Thober, S., Merz, B., and Samaniego, L.: High-resolution impact-based early warning system for riverine flooding, Nature communications, 15, 3726, 2024.

Neal, J., Schumann, G., Fewtrell, T., Budimir, M., Bates, P., and Mason, D.: Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK, Journal of Flood Risk Management, 4, 88–95, https://doi.org/10.1111/j.1753-318x.2011.01093.x, 2011.

Pasculli, A., Cinosi, J., Turconi, L., and Sciarra, N.: Learning case study of a shallow-water model to assess an early-warning system for fast alpine muddy-debris-flow, Water, 13, 750, 2021.

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, Hydrology and earth system sciences, 11, 1633–1644, 2007.

Rentschler, J., Avner, P., Marconcini, M., Su, R., Strano, E., Vousdoukas, M., and Hallegatte, S.: Global evidence of rapid urban growth in flood zones since 1985, Nature, 622, 87–92, 2023.

Riembauer, G., Weinmann, A., Xu, S., Eichfuss, S., Eberz, C., and Neteler, M.: Germany-wide Sentinel-2 based land cover classification and change detection for settlement and infrastructure monitoring, in: Proceedings of the 2021 Conference on Big Data from Space, Virtual, pp. 18–20, 2021.

Šakić Trogrlić, R., van den Homberg, M., Budimir, M., McQuistan, C., Sneddon, A., and Golding, B.: Early warning systems and their role in disaster risk reduction, in: Towards the "perfect" weather warning: bridging disciplinary gaps through partnership and communication, pp. 11–46, Springer International Publishing Cham, 2022.

See, L.: A review of citizen science and crowdsourcing in applications of pluvial flooding, Frontiers in Earth Science, 7, 44, 2019.

Seleem, O., Ayzel, G., de Souza, A. C. T., Bronstert, A., and Heistermann, M.: Towards urban flood susceptibility mapping using data-driven models in Berlin, Germany, Geomatics, Natural Hazards and Risk, 13, 1640–1662, 2022.

Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen Berlin: Geoportal Berlin, https://gdi.berlin.de/viewer/main/, accessed: 2025-05-16, 2022.

Shrestha, A., Mascaro, G., and Garcia, M.: Effects of stormwater infrastructure data completeness and model resolution on urban flood modeling, Journal of Hydrology, 607, 127 498, 2022.

Thieken, A. H., Bubeck, P., Heidenreich, A., von Keyserlingk, J., Dillenardt, L., and Otto, A.: Performance of the flood warning system in Germany in July 2021 – insights from affected residents, Natural Hazards and Earth System Sciences, 23, 973–990, https://doi.org/10.5194/nhess-23-973-2023, 2023.

Tügel, F., Nissen, K. M., Steffen, L., Zhang, Y., Ulbrich, U., and Hinkelmann, R.: Extreme precipitation and flooding in Berlin under climate change and effects of selected grey and blue-green measures, EGUsphere, 2025, 1–26, 2025.

Wang, P., Huang, C., and Tilton, J. C.: Mapping three-dimensional urban structure by fusing landsat and global elevation data, arXiv preprint arXiv:1807.04368, 2018.

Ward, P. and de Bruin, S.: The geography of future water challenges, 2018.

Wing, O. E., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., and Erickson, T. A.: Validation of a 30 m resolution flood hazard model of the conterminous United States, Water Resources Research, 53, 7968–7986, 2017.

Winsemius, H., Van Beek, L., Jongman, B., Ward, P., and Bouwman, A.: A framework for global river flood risk assessments, Hydrology and Earth System Sciences, 17, 1871–1892, 2013.

WorldPop: The spatial distribution of population in 2020, Germany, https://dx.doi.org/10.5258/SOTON/WP00670, https://doi.org/10.5258/SOTON/WP00670, school of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Departement de Geographie, Universite de Namur; Center for International Earth Science Information Network (CIESIN), Columbia University. Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076), 2020.

Zeng, B., Huang, G., and Chen, W.: Research progress and prospects of urban flooding simulation: From traditional numerical models to deep learning approaches, Environmental Modelling & Software, 183, 106 213, 2025.

Zhou, Q., Leng, G., Su, J., and Ren, Y.: Comparison of urbanization and climate change impacts on urban flood volumes: Importance of urban planning and drainage adaptation, Science of the Total Environment, 658, 24–33, 2019.