RC2: 'Comment on egusphere-2025-3962', Anonymous Referee #2, 11 Dec 2025

In their paper, the authors present an innovative urban flood monitoring approach. Intersecting segmented flood masks derived from imagery recorded by low-cost trail cameras and lidar data, they estimate flood water surface elevations for two flood events. Maximum flood depths and extents were then compared with results from a 2D hydrodynamic model.

I have read the paper with interest and think it can be published after major revision. My detailed comments are included below.

Major concerns

1. HEC-RAS model

While I agree in general that comparing flood extent and/or depth from the authors' new method with results from a 2D hydrodynamic model might be an interesting analysis, the paper in its current for lacks important details regarding how the HEC-RAS model was implemented:

[Q1] The details of the HEC-RAS model implementation are currently included in the Zenodo data supplement. Because the HEC-RAS model itself is secondary to the development of our analysis pipeline, we do not want to overwhelm the readers with too many details in the main text given the current length of the manuscript. We will migrate the relevant elements into the main text methods section and supplementary information.

How was the model grid set up?

We have added the following to the supplement:

> ***(Text S1) Model domain:*** *The computational mesh was generated from a TIN-interpolated, and gap filled, 0.5 m resolution DTM generated from 2019 USGS 3DEP aerial lidar data. The base mesh was generated with 10 m node spacing. Breaklines were added for channel centerlines, culvert inflows and storm drains. Mesh refinement was applied within 5 m of these breaklines, reducing node spacing to 1 m for the area of interest.*

What infiltration method was used in the rain-on-grid approach, and how was it parameterized? What land use classifications and corresponding roughness coefficient were used?

We have added the following to the supplement:

> ***(Text S1) Mannings roughness (n) and runoff:*** *30 m resolution National Land Cover Database (NLCD) were used to define spatially variable roughness coefficients, using HEC-RAS manual reference values for each classification. This was refined using vector polygons of road surfaces and building footprints from the Illinois Department of Transportation. Within building footprints n was assigned a high value of 10 which prevents the routing of runoff from those cells.*

Land Cover Classification	Manning's N Value
NoData	0.035
Roads	0.01
Buildings	10
Banks	0.04
MainChannel	0.03
Open Water	0.035
Developed, Open Space	0.035
Developed, Low Intensity	0.08
Developed, Medium Intensity	0.12
Developed, High Intensity	0.15
Barren Land Rock/Sand/Clay	0.025
Deciduous Forest	0.15
Evergreen Forest	0.15
Mixed Forest	0.1
Shrub/Scrub	0.9
Grassland/Herbaceous	0.04
Pasture/Hay	0.045
Cultivated Crops	0.05
Woody Wetlands	0.07
Emergent Herbaceous Wetlands	0.07

Table S2: Reference landcover based Manning 's roughness coefficients taken from USACE 2024.

Rainfall Runoff: *The curve number (CN) method was used to calculate initial infiltration losses and runoff generation. The same NLCD landcover, and IDOT building and road layers were used to define spatially variable values for CN, abstraction ratio and minimum infiltration rate. Reference values were taken from the HEC-RAS hydraulics manual, with abstraction ratios suggested by Hawkins and Jiang 2023. No additional infiltration losses are calculated after the initial rainfall-runoff conversion.*

Name (Land Cover: Soil Hydric Group)	**Curve Number**	**Initial Abstraction Ratio**	**Abstraction Ratio**	**Minimum Infiltration Rate**
Developed, Medium Intensity : D	86	0.05	0.082	1.270
Developed, Open Space : B/D	74	0.05	0.176	3.485

Emergent Herbaceous Wetlands : A	76	0.05	0.158	7.600
Grassland/Herbaceous : D	80	0.05	0.125	1.270
Buildings :	100	0.05	0.000	0.000
Roads : C	98	0.05	0.010	0.000

Was storm drain infrastructure modeled, or only surface flow?

We have added the following to the supplement:

> ***(Text S1) Stormwater system:*** *The combined sewer-stormwater system was modeled as a 1D pipe network in HEC-RAS. The location of stormwater inflows were taken from the Illinois Department of Natural Resources (IDNR) and Heartlands Institute survey of the Prairie Du Pont Watershed. Precise information on the topology and hydraulics of the sewer-stormwater system is unavailable, and connection between inflows was inferred based on published IDNR and USACE reports and maps (USACE 2024; IDNR 2023). Pipe diameter of 0.7 m, based on IDNR survey, and n of 0.015 m were used, based on USACE reported values. The pipe network was connected to the know drainage ditch outfall.*

Are there storage areas within the model domain?

> Because the study area is separated from nearby lakes and reservoirs by the major drainage ditches and roads no additional storage areas were included.

Without additional detail, a review of this portion of the analysis is nearly impossible. Even if detail is added, I still question the value of comparing the authors results with those from an uncalibrated HEC-RAS model; I also suspect the lack precipitation data form within the study area adds substantial uncertainty to model results (it sounds like data from only one rain gauge was used for each event, and gauges were located at a distance of 6 and 8 km from the study area, respectively).

[Q3] We Thank the reviewer for their comments, but even without quantitative calibration we believe there is significant value in our model comparison. A major motivation of our study is monitoring approaches for areas without sufficient data for pluvial model calibration. This approach is not unique to our study. There are multiple prior studies where uncalibrated, and otherwise simplified 2D models are used to evaluate new DTM based methods (e.g. Samela et al. 2020; Preisser et al. 2022). To that end, the comparison is intended to identify major similarities and differences in characteristic behavior between camera-derived flood extents and a rain-on-grid model, not as an absolute ground-truth. While factors like precipitation uncertainty will modify maximum flood timing and extent, they are unlikely to alter the behaviors implicit to a rain-on-grid model which distinguish it from our image-based estimates. We will revise the methods and discussion to clarify our conceptual approach and qualify the limits of direct camera to model validation:

> ***Lines 372-375: "****Because the model itself is only qualitatively calibrated, its output is not treated as a direct validation for absolute water levels estimated from images. Instead, it characterizes similarity or divergence in flood behavior predicted by each method, based. This is quantified both in terms of the relative agreement in predicted flood extent, and spatial flood connectivity, between the two methods."*

> ***Lines 707-715:*** *"Despite these challenges, our results demonstrate how empirically-derived WSEs can complement and strengthen traditional hydraulic modeling workflows. Our method provides continuous, high-resolution estimates of water level and extent that are directly tied to real flood behavior, capturing sub-decimeter changes in WSE and floodwater connectivity that would otherwise be missed by point-based flood monitoring and modeling approaches. While further validation of camera-derived extents would be necessary for confident direct calibration, this level of precision is valuable for the initial validation of uncalibrated models, an important tool for preliminary flood-risk analysis in settings with no gauges or rapidly changing infrastructure performance."*

I think one of the potentially important applications of the proposed method is mentioned in the discussion (lines 588-589): data for calibration of hydrodynamic models is limited, particularly for pluvial flooding. Here, estimates of water surface elevations and flood extents from cameras could fill an important data gap. If the authors could demonstrate that they can calibrate their HEC-RAS model using camera-based observations, that would strengthen the paper considerably.

While we agree with the reviewer that there is significant future opportunity in using camera-based observations to calibrate flood models (which we discuss in Lines 810-815), we feel that this effort is beyond the scope of the study presented here. The focus of the current study is application of the computer vision methods to estimate spatial flood extent, and introducing an additional model calibration element is likely to detract from that focus. The reviewer's suggestion would require us to simultaneously evaluate the performance of our camera-based methods while also applying those methods to calibrate the HEC-RAS model, introducing ambiguity into the interpretation of both elements of the study.

We will revise the discussion to more directly call out the potential use of these methods for future flood model calibration as follows:

> **Lines 847-863***: "Camera-based observations provide a promising avenue to address these calibration gaps. Depending on the data available and the precision required, camera-derived information could support multiple levels of model calibration. At a minimum, observations of flood presence, extent, and connectivity can serve as semi-quantitative validation of model structure and behavior. More detailed or well-distributed camera installations could function as stream-gauge surrogates, enabling direct calibration of key model parameters such as surface roughness, stormwater capacity, or flood wave timing. These approaches could ultimately facilitate both post-event model evaluation and real-time model adjustment, bridging gaps in empirical data for urban flood forecasting. When possible to implement, camera-derived WSEs offer a rare empirical reference for validating modeled spatiotemporal patterns of inundation. For example, these high-resolution, time-resolved observations enabled direct comparison with outputs from an uncalibrated HEC-RAS Rain-on-Grid simulation of the July 4 flood event, revealing a close match in peak flood depth, timing,*

and extent. This proof-of-concept highlights the strong potential of integrating image-derived data into calibration workflows for 2D hydrodynamic models, particularly in high-flow scenarios where floodwaters are hydraulically connected and drainage networks are overwhelmed. Beyond event reconstruction, such observations can support real-time model updating, performance evaluation of stormwater infrastructure, and planning for flood mitigation in poorly instrumented or rapidly evolving urban settings, providing a practical, data-driven way to reduce uncertainty in urban flood simulations."

2. Extension of flood extends beyond the camera field of view

The authors apply a flood fill procedure to estimate flood extents outside the camera field of view. I question this approach, which can't account for overland flow dynamics, infiltration, etc. I think the authors might be better off using the flood fill approach only within the field of view.

[Q5] We thank the reviewer for their comment. However, there is an important conceptual distinction between flood-fill models and process-based flood models. By prescribing a surface water level the flood-fill extents are agnostic to the factors such as infiltration that produced the water level. Further, flood fill models do not contain an explicit time-dependence, each flood extent is independently generated from a single image observation. Within the context of pluvial flooding within an urban area, where the fraction of impervious surfaces is quite high and the study area contains many small, internally draining depression, a flood-fill model is a suitable approximation for extrapolation over short (~100m) distances.

Specific to this contribution, extrapolation serves an important role in cross-site comparison. In a low-relief environment we can expect discrete in flood connectivity, and categorical disagreements between sites reveal missing dynamics, namely storm-water infrastructure. The main point is that using flood fill models to propagate flood extents from multiple cameras improves our ability to identify these dynamics.

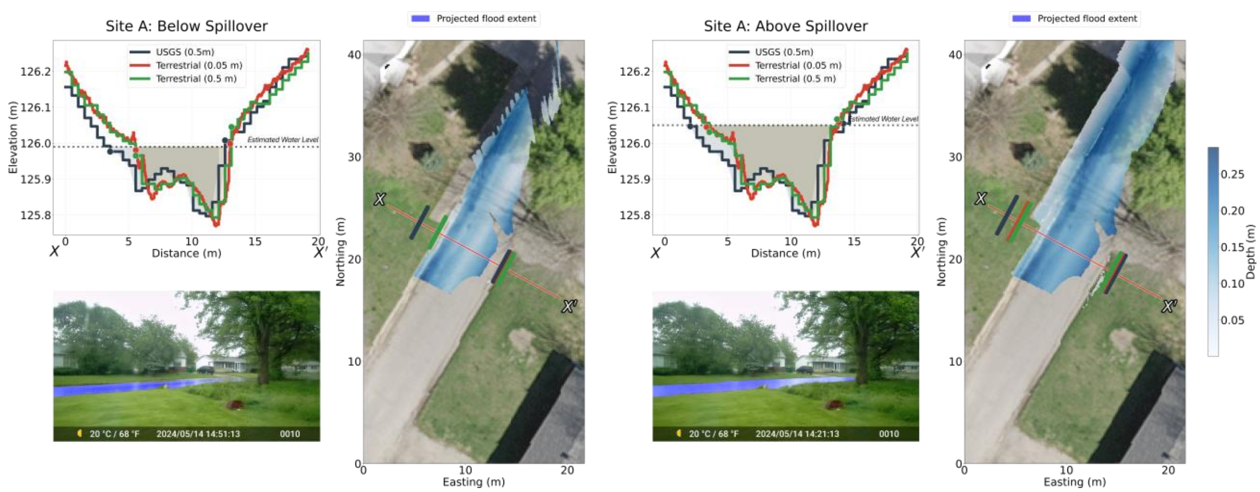
Given the widespread use of static water-level models for rapid flood assessment, it is valuable to discuss their behavior and limitations in the context of emerging flood data sources such as cameras (Gallien et al., 2014; Hong et al., 2024; Li et al., 2022; Preisser et al., 2022). We have expanded the discussion to highlight the limitations of the flood-fill approach and suggest avenues for integrating data from multiple cameras in future work.

Lines 787-808: *"Flood fill methods are well-suited for short duration pluvial events in low relief, urban areas. Because the study sites within a self-contained depression, it is unlikely that there are substantial gradients in water surface elevation. This is supported by the 2D model results which, exclusive of edge effects, predicts a difference in water elevation between Sites A and B of only 0.5 cm after initial merging of the flood patches. Static elevation-based methods are widely used for rapid flood mapping, including in emergency management contexts (Gallien et al., 2024; Wang et al. 2024; Zheng et al., 2018). The cross-camera comparison used in this study is an effective tool for identifying potential failure modes within these models."*

3. Validation of the new method

The study would also be strengthened if estimated water surface elevations could be validated using other data sources. I understand that depth measurements may not be available, but could the authors estimate depth at strategic locations based on visible markers and compare those to estimates from their approach for the corresponding location? Also, some expanded discussion of uncertainty as a function of distance from the camera location would be beneficial.

Independent depth measurements are not available for the study site, and indeed the lack of such data is the primary motivation behind this project. However, there are identifiable markers of discrete jumps in water level, such as road overflow points (Vandele et al. 2019). Based on the lidar DTMs, and aerial imagery we compared projected flood extents, with road elevation profiles above and below spillover. This is included in a revised supplementary Figure:



"SI 7: Road topographic profiles based on the 0.5m USGS DTM (black), 0.005m terrestrial DTM (red), and 0.5m terrestrial DTM (green). b) Flood extent interpolated from intersection of the projected terrestrial lidar points with the flood mask. C) Corresponding image with flood mask overlay.

To qualitatively validate the water level extraction method we examined observation immediately above and below overtopping of the road boundary. These were compared against cross-road elevation profiles extracted from both USGS and terrestrial DTMs. Prior to spillover, the projected flood extent ends at the road boundary, with water level slightly below the curb elevation found in the terrestrial DTM profile. After spillover, the project extent expands to fill the small paved area above the road surface, before stopping at the lawn boundary. This is consistent with the image observation, and topographic profile of a second spillover onto the lawn itself. Further from camera, toward the NW, decreased pixel resolution leads to the projected extent bleeding beyond the road, potentially upwardly biasing estimated water levels. The extracted water level is approximately 3 cm higher than the elevation contour best aligned with the projected flood boundary below spillover, and approximately 1cm higher after spillover.

The magnitude of both these biases decreases with larger flood extents due to more gradual elevation gradients, and the lack of curb shadows."

We have added discussion of distance dependent pixel resolution and its influence on camera pose, and water level estimation:

Line 332-355:* "Because pixel resolution decreases with distance from the camera, multiple 3D points may project onto a single image pixel; therefore, all inundated points are retained and a one-to-one pixel–point correspondence is not enforced. Together, these inundated points represent the portion of the ground surface that is underwater at the time the image was captured. This set of inundated points is interpolated into a 0.05-meter resolution raster representing the visible flood extent in the image. This interpolation step reduces bias associated with distance-dependent differences in point density and avoids over-representation of regions where many 3D points project to a single pixel. Water surface elevation (WSE) is estimated from the rasterized flood boundary rather than from individual pixels or raw point projections. Canny edge detection is applied to the rasterized inundation extent to identify the flood boundary, and the 90th and 95th percentiles of the resulting edge elevation distribution (WSE_{90} and WSE_{95}) to account for potential topographic noise or obstruction of the water edge in the time lapse images. Assuming a flat water surface, elevations along the flood boundary should exhibit a sharp peak at the upper end of the elevation distribution. The consistency and sharpness of this peak are another parameter useful to evaluate the camera pose estimation, as errors in estimated camera orientation or translation produce unrealistically large elevation differences between near- and far-field water edges."*

Other comments

Figure 1b – this image is difficult to interpret, perhaps change the color scheme?

[Q7] We will adjust this figure to accentuate the small-scale variation of the floodplain, while preserving the context of the upland bluffs.

Section 2.3 – I recommend revising this section. It is difficult to follow the detailed accounts of start and end times. It might be better to display this as a figure or omit some of the detail not necessary for understanding the larger picture.

[Q8] We will add additional annotation of key elements of flood timing and duration to figures 3 & 4 and will remove unnecessary details from the text. However, we feel that some narrative description of the events as observed by the cameras is necessary to build reader intuition and understanding and better prime them for presentation of the flood extent estimates.

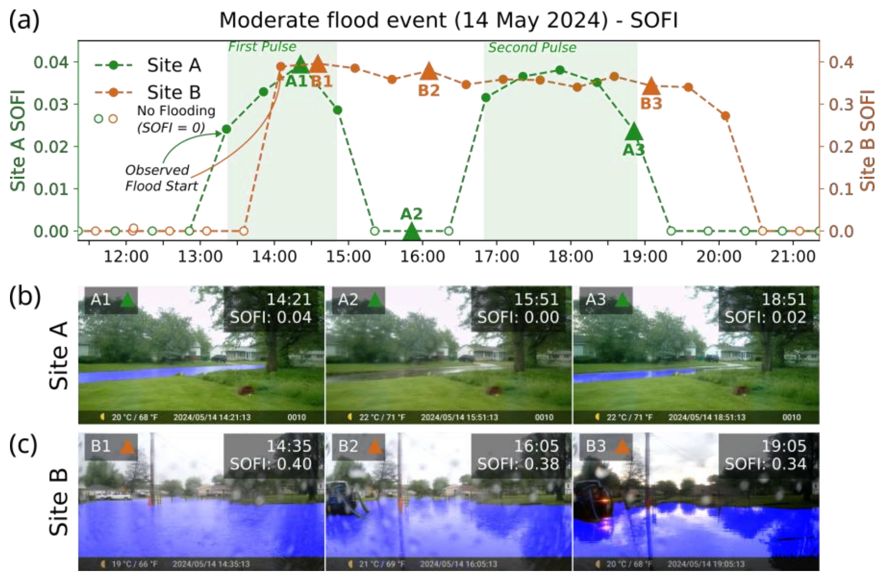


Figure 3: (a) SOFI time series for 14 May moderate severity case study event. Representative flooded images from (b) Site A and (c) Site B. Segmented flood masks are shown in blue.

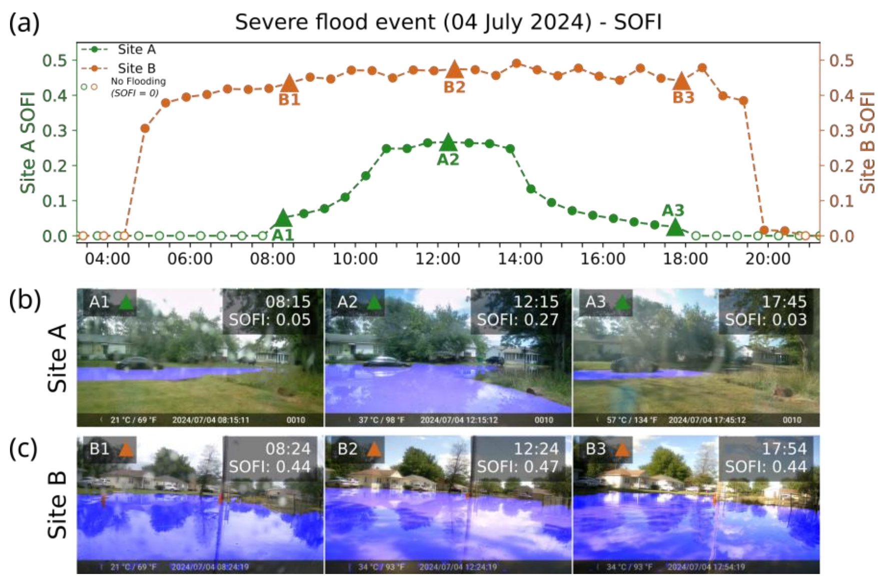


Figure 4: (a) SOFI time series for 04 July severe case study event. Representative flooded images from (b) Site A and (c) Site B. Segmented flood masks are shown in blue.

Line 322: What do the authors mean by "rainfall was uniformly applied to the domain"? Precipitation data from one gauge was applied to the entire model domain? Uniform intensity? Please clarify.

[Q9] A single rain gauge record was used for each flood event, and that hyetograph was applied to the entire model domain grid, with no spatial variability in precipitation. While a simplification, the dominance of local pluvial runoff in the study area, and the short duration of the case-study flood events, likely limits the influence of watershed scale precipitation gradients, and would not alter the basic behaviors of the rain-on-grid model relevant for comparison to the camera-derived estimates. We have clarified this in the methods:

> ***Lines 350-353:*** *"This model is implemented using the Hydrologic Engineering Center's River Analysis System (HEC-RAS), configured with a "rain-on-grid" unsteady boundary condition to simulate overland water flow across an 89.6 km² model domain covering the study site (USACE, 2022). The base terrain is the 0.5 m USGS DTM. Rainfall records defined the unsteady inputs the model domain, assuming spatially uniform precipitation."*

Line 365: Please explain how SOFI values should be interpreted.

[Q10] SOFI is the fraction of the total fraction of an image classified as flooded. It is included as a semi-quantitative metric of flood magnitude to contextualize the estimated water level, and flood extent trends. However, the absolute values of SOFI are a function of both the physical flood extent, and the perspective of the camera. For example, a camera installed directly Infront of a flood source will see SOFI initially rise very quickly, before leveling off as flooding fills the FOV (see SI Figure 2). We have added additional text explaining this interpretation:

> ***Line 255-259****"This ratio is referred to as the Static Observer Flooding Index (SOFI), following the approach of Vitry et al. (2019), providing a simple proxy for flood intensity as seen from a fixed observation point. SOFI has been shown to correlate strongly with changes in water level for a given location (Moy de Vitry et al. 2019). The shape and magnitude of SOFI response depend strongly on the geometry of a camera relative flooding, and as such values cannot be directly compared between study sites."*

Reply Citations:

Hong, Y., Kessler, J., Titze, D., Yang, Q., Shen, X., & Anderson, E. J. (2024). Towards efficient coastal flood modeling: A comparative assessment of bathtub, extended hydrodynamic, and total water level approaches. Ocean Dynamics, 74(5), 391-405.

Gallien, T. W., Sanders, B. F., & Flick, R. E. (2014). Urban coastal flood prediction: Integrating wave overtopping, flood defenses and drainage. Coastal Engineering, 91, 18-28.

Li, Z., Mount, J., & Demir, I. (2022). Accounting for uncertainty in real-time flood inundation mapping using HAND model: Iowa case study. Natural Hazards, 112(1), 977-1004.

Silverman, A. I., Brain, T., Branco, B., sai venkat Challagonda, P., Choi, P., Fischman, R., ... & Toledo-Crow, R. (2022). Making waves: Uses of real-time, hyperlocal flood sensor data for emergency management, resiliency planning, and flood impact mitigation. *Water Research*, *220*, 118648.

Gold, A., Anarde, K., Grimley, L., Neve, R., Srebnik, E. R., Thelen, T., ... & Hino, M. (2023). Data from the drain: A sensor framework that captures multiple drivers of chronic coastal floods. Water Resources Research, 59(4), e2022WR032392.

Negri, R., Ceferino, L., & Cremen, G. (2025). Prioritizing urban areas for the deployment of hyperlocal flood sensors using stakeholder elicitation and risk analysis. Natural Hazards Review, 26(3), 04025020.