41ST IAHR WORLD CONGRESS SINGAPORE
22 – 27 JUNE 2025
INNOVATIVE WATER ENGINEERING FOR SUSTAINABLE DEVELOPMENT

Book of Extended Abstracts
41st IAHR World Congress
22-27 June 2025
Singapore

USING COMPREHENSIVE FIELD DATA FOR VALIDATING LARGE-SCALE HYDRODYNAMIC HINDCASTING OF AN EXTREME FLOOD EVENT

Pratik CHAKRABORTY[1], Christophe DESSERS[1], Pierre ARCHAMBEAU[1], Michel PIROTTON[1], Sébastien ERPICUM[1], Benjamin DEWALS[1]

[1] Hydraulics in Environmental and Civil Engineering (HECE), University of Liège (ULiège), Belgium
email: p.chakraborty@uliege.be

ABSTRACT

Hydrodynamic modelling is a crucial tool for simulating large-scale floods, yet validation remains challenging due to discrepancies between modelled and observed datasets, especially in urbanized floodplains. The July 2021 Vesdre Valley flood in Belgium provides a unique opportunity for rigorous validation, supported by an extensive post-flood survey with over 16,000 high-water marks recorded by the river authority (SPW). This study evaluates the performance of a high-resolution 2D hydrodynamic model (*WOLF*) by comparing simulated and observed flood extents and water depths. To address discrepancies arising from structural artifacts in built-up areas, three harmonization approaches—Buffered-Buildings, Filling-of-Holes, and Inverse Adaptation—are tested. Validation using the Jaccard similarity index demonstrates strong spatial agreement, with a domain-average value of 0.86. Water depth comparisons reveal consistent agreement across most sectors, though localized deviations highlight room for potential improvement in feature resolution. Nonetheless, findings suggest that even with extensive field data available, comprehensive validation remains a challenge, thereby necessitating careful consideration and application of appropriate post-processing for effective model validation.

Keywords: Hydrodynamic modelling; Vesdre Valley; flood modelling validation; inundation mapping; Jaccard similarity index.

1. Introduction

Hydrodynamic modelling is often employed for simulating large-scale floods, yet validation of these models with measured data is complex due to the often sparse and unevenly distributed nature of flood observation data. Survey data density and spatial extent frequently involve trade-offs, limiting comprehensive validation opportunities. The July 2021 floods in Belgium's Vesdre Valley (Dewals et al., 2021) presented a significant case study, as this catastrophic event was followed by an unusually detailed post-flood survey conducted by the Service public de Wallonie (SPW). This survey provided an extensive dataset of observed flood levels across the region, offering an exceptional basis for validating a hydrodynamic model's performance.

The present study aims to validate a hydrodynamic hindcast of the July 2021 floods in the Vesdre valley by comparing its simulated flood extents and water depths to the SPW survey data. Given that the hydrodynamic model and SPW data differ in the way building footprints are dealt with, this research explores three methods to bridge these differences in the datasets.

Studies employing flood modelling chains are well-documented in the literature, with validation efforts often centred around data from measuring stations, such as hydrographs or water-level records at key locations. While these datasets provide valuable insights into riverine dynamics, they often fail to capture the complexity of inundation across floodplains. Validation exercises that incorporate extensive floodplain observations are comparatively infrequent, primarily due to the limited availability of such detailed datasets.

What sets the Vesdre Valley flood study apart is the sheer scale and density of the observational dataset which includes over 16,000 high-water marks – meticulously recorded across the affected floodplain. To the best of the authors' knowledge, this level of detail is unprecedented for validating flood extents and depths. This comprehensive dataset promises a granular evaluation of model performance, providing critical insights into both localized and large-scale hydrodynamic behaviour.

2. Methodology

2.1. Numerical model

The integrated modeling system *WOLF*, developed by the *HECE* research group at *ULiège*, was employed to perform the hindcast of the Vesdre Valley flood event. The 2D depth-averaged hydrodynamic model utilized outputs from the *WOLFHydro* hydrological model (Dessers et al., 2024) as spatially distributed runoff inputs (source terms). A computational mesh with a resolution of 5 m × 5 m was generated for the entire valley. High-resolution topographic and bathymetric data (LiDAR v2021-22 - 0.5 m × 0.5 m) were resampled to this coarser resolution using a *mean* operator.

Boundary conditions at the downstream limit were specified using a Froude number approach, ensuring stability and consistency with flow dynamics. Variable Manning's coefficients, derived from land use data, were applied to account for spatial variations in surface roughness and flow resistance.

2.2. Data used for validation

This study leverages data from both observed flood extents and model outputs. The observed flood data, collected by the Service public de Wallonie (SPW), includes a rasterized inundation map generated from 16,000 high-water marks. These points were interpolated between (or extrapolated when necessary) using an Inverse Distance Weighted (IDW) approach to produce a continuous inundation map.

Table 1. Summary of data used in the process of validation

Data	Version	Description
Observed water depths (SPW)	2021	Raster of observed water depths derived from field observations of water levels and the DEM (LiDAR) of 2013-14.
PICC	2024	Building footprints (polygonal) data.
Evaluation sectors	-	Author-defined spatial units to base model evaluation on.

2.3. Validation Methods

Comparing flood extents between observed and modelled data is challenging due to fundamental differences in how each dataset is derived and represented. Observed flood maps, like those from the SPW, are interpolated from point-based field data and often depict inundation within built-up areas, treating structures as permeable to flooding (Schubert and Sanders, 2012). In contrast, the hydrodynamic model represents buildings and other raised structures as impermeable and elevated, resulting in "holes" in the simulated flood extent. This mismatch introduces structural artifacts that can distort validation outcomes.

To address these discrepancies and ensure a fairer comparison, we tested three distinct methods for harmonizing the observed and modelled flood datasets. Each method adjusts one or both datasets to align their treatment of built-up areas.

Buffered-Buildings approach: In the Buffered-Buildings (BB) approach, we introduce a buffer around each building footprint in the model's DEM to simulate inundation within these areas when surrounded by water. Since buildings are treated as elevated impermeable structures, they often appear as dry patches in the modelled flood extent, while the observed map, generated from interpolated data, shows these areas as inundated. To resolve this, a buffer zone around each building is checked for inundation status. If any cell within the buffer shows flood depth above a minimum threshold, the building footprint is marked as inundated. This buffer approach is sensitive to the buffer size, so we test various buffer widths (0, 5, and 10 meters) to ensure accurate alignment with the observed map.

Filling-of-Holes approach: The Filling-of-Holes (FoH) approach simplifies the issue of building discrepancies by directly addressing "holes" in the model's flood extent. These holes—non-inundated cells within otherwise inundated areas—primarily represent building footprints and are often shown as flooded in the observed map due to differences in processing. In this approach, we identify and "fill" any dry patches that are fully enclosed by flooded cells in the model output. By reassigning these patches as inundated, we approximate how the observed map depicts inundation in built areas. This method is efficient but may occasionally overestimate flooded areas by filling genuinely dry patches due to local high terrain rather than buildings. It is also prone to underestimation in case a block of buildings is clearly inundated but doesn't technically constitute a *hole* because of not being *completely* surrounded by water.

Inverse Adaptation approach: The Inverse Adaptation (IA) approach takes an alternative strategy by adjusting the observed flood map to match the model's representation of buildings and high topography. Here, all inundated cells within building footprints on the observed flood map are set to dry, mirroring the model's approach that treats these structures as impermeable. This adaptation removes the need to simulate flooding in building footprints in the model output and maintains elevated dry areas accurately. While this approach may lead to an overall underestimation of flooded area by removing building footprints from the observed map, it improves the alignment between observed and modelled data, allowing a focused comparison of flood extents and water depths.

Thereafter, a robust validation metric, the Jaccard similarity index, has been adopted to assess agreement of extents.

$$J(M, O) = \frac{M \cap O}{M \cup O} \tag{1}$$

where *J* – Jaccard similarity index, *M* – Boolean raster of modelled flood extents (i.e., 1 and 0 indicating inundation and non-inundation respectively) and *O* – Boolean raster of observed flood extents.

To assess model performance in terms of water depths, we compare the modelled and observed rasters by calculating the difference between them. This raster-based comparison provides spatially continuous insights into the model's accuracy. Errors are quantified using standard metrics (e.g., mean absolute error) across predefined river reaches, (evaluation sectors in Table 1). These sectors facilitate localized performance assessments and help identify specific areas where the model either over- or underestimates flood depths.

3. Results and discussion on validation outcomes

3.1. Comparison of flood extents

Figure 1 depicts the performance of the model with respect to spatial extents using the Jaccard similarity index.

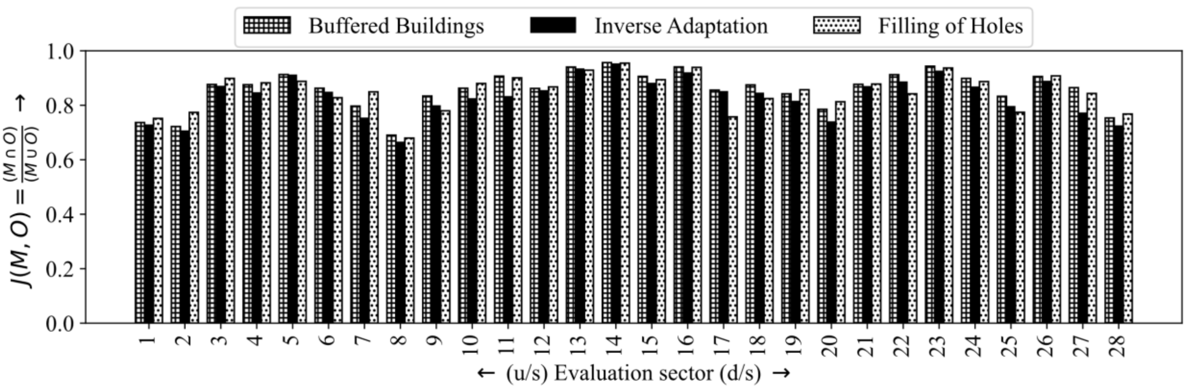

Fig. 1. Jaccard similarity indices (*J*) for different evaluation sectors using different approaches to harmonize the observed (*O*) and modelled (*M*) flood-extents rasters (u/s and d/s refer to upstream and downstream respectively).

The model achieves a high score in almost every sector, peaking at 0.96 and a lowest value of 0.69. By and large, the average value is around 0.86, which is remarkable when compared to literature (de Arruda Gomes et al., 2021). We also find that the three techniques employed yield marginally different values but no one method consistently yields better or worse outcomes than others in all sectors. The reasons for the same shall be detailed in the subsequent extensive presentation of the work.

3.2. Comparison of water depths

Figure 2 depicts the differences in water surface elevations (WSE) between modelled and observed datasets across 28 evaluation sectors spanning the Vesdre Valley. Overall, the results exhibit a generally consistent performance, with most evaluation sectors showing median WSE differences between -0.4 m to +0.4 m. This alignment suggests that the model reasonably captures the flood dynamics across the valley. However, sector-specific variations are evident.

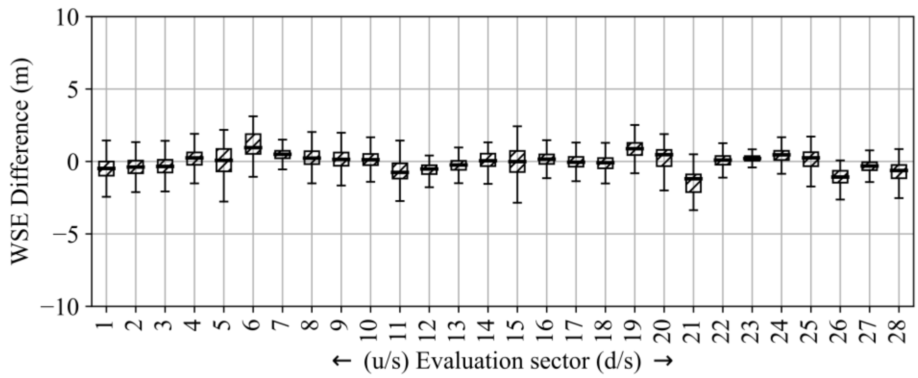

Fig. 2. Box plot of water surface elevation (WSE) differences per evaluation sector (positive values imply overprediction by the model and vice-versa).

4. Discussion on challenges identified

A significant challenge arises when attempting to validate the hydrodynamic model against the observed dataset, which consists of a large number of discrete high-water marks (x, y coordinates) with associated water depths. These high-water marks are typically recorded from building facades and are then used to create a continuous raster of water depths. While it might seem logical to compare the discrete observed water depths directly with the corresponding modelled depths to avoid potential errors introduced by interpolation or extrapolation, this approach is often infeasible.

The reason lies in the spatial placement of the high-water marks: most are located 'within' building footprints, as these represent the structures from which the measurements were taken. In the hydrodynamic model, however, no water depths are assigned to these points (Section 2.2). This discrepancy renders direct comparisons at these discrete locations problematic. As a result, it becomes necessary to work with the continuous raster representation of observed water depths. This, in-turn, introduces another layer of uncertainty since, in this case, the construction of the continuous raster is based on topography data from 2013-14 (which is likely to have evolved in different parts of the domain).

Further, in areas with sparse field data, the accuracy of the continuous water depth raster can be compromised, as the interpolation process may fail to capture local variations in flood extents and water depths accurately. Furthermore, regions where the continuous raster is generated through extrapolation—extending beyond the spatial bounds of the observation points—are particularly prone to high uncertainty. These factors not only impact the reliability of validation metrics, such as the Jaccard similarity index, but also obscure the true performance of the hydrodynamic model, making it difficult to discern whether observed discrepancies stem from model inaccuracies or artifacts of the observation dataset itself.

5. Conclusion

This study highlights the complexities and nuances inherent in validating large-scale hydrodynamic models using detailed observational datasets. While the availability of a comprehensive dataset, such as the high-water marks recorded after the July 2021 Vesdre Valley flood, provides a unique opportunity for rigorous validation, several challenges arise from fundamental differences in how observed and modelled data are generated and represented.

In light of all aforementioned challenges, this study emphasizes the need for careful post-processing and methodological adjustments to enable meaningful validation of hydrodynamic models. The present work highlights the importance of critically evaluating the datasets themselves, identifying structural differences, and applying appropriate corrections. Such efforts not only improve the robustness of model validation but also provide valuable insights for refining flood hazard assessment methodologies in future studies.

References

de Arruda Gomes, M. M., de Melo Verçosa, L. F., & Cirilo, J. A. (2021). Hydrologic models coupled with 2D hydrodynamic model for high-resolution urban flood simulation. Natural Hazards, 108(3), 3121-3157.

Dessers, C., Archambeau, P., Dewals, B., Erpicum, S., Pirotton, M. (2024). WOLFHydro: a modular framework for multi-model hydrological simulations. EGU General Assembly 2024.

Dewals, B., Erpicum, S., Pirotton, M., & Archambeau, P. (2021). July 2021 extreme floods in the Belgian part of the Meuse basin. Hydrolink, 4, 104–107. https://www.iahr.org/library/hydrolink?hid=412

Schubert, J. E., Sanders, B. F. (2012). Building treatments for urban flood inundation models and implications for predictive skill and modeling efficiency. Advances in Water Resources, 41, 49-64.