LIÈGE université
Sciences Appliquées

Modelling of fluvial dike breaching due to overtopping

A thesis submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy (PhD) in Engineering Science

by

Vincent SCHMITZ

Supervisor: Benjamin DEWALS

Co-supervisor: Sébastien ERPICUM

DOCTORAL COLLEGE IN ARCHITECTURE, ENGINEERING AND GEOLOGY

DECEMBER 2024

This thesis was publicly defended on February 19, 2025, in front of a jury composed of

President:	Prof. Frédéric Nguyen	(ULiège, Belgium)
Members:	Prof. Maarten Arnst	(ULiège, Belgium)
	Dr. Stéphane Bonelli	(INRAE, France)
	Prof. Weiming Wu	(Clarkson University, NY, USA)
Supervisor:	Prof. Benjamin Dewals	(ULiège, Belgium)
Co-supervisor:	Dr. Sébastien Erpicum	(ULiège, Belgium)

Abstract

Fluvial dikes are structures built along watercourses to protect humans and infrastructures from flooding. The risk of dike breaching is increasing due to more frequent extreme weather events, urbanization near these structures, and aging existing dikes. Among the causes of breaching, overtopping is the most common. Accurate prediction of the breaching dynamics and resulting flooding is crucial for effective emergency planning in floodplains.

This thesis focuses on developing a numerical model for homogeneous non-cohesive fluvial dike breaching induced by overtopping. While various models exist for dam breaching, these structures differ fundamentally from fluvial dikes. Embankment dams, which are perpendicular to the main flow, tend to breach symmetrically, whereas fluvial dikes mainly expand in the downstream direction due to the river momentum.

Building on an existing lumped model primarily designed for dam breaching, this work developed a new lumped model dedicated to fluvial dikes. A global sensitivity analysis and an uncertainty analysis applied on the initial dam model identified model parameters whose uncertainty had a critical impact on the variability of the model outputs, namely the breach discharge and expansion. Laboratory tests further clarified the influence of dike geometry and main channel width on breaching dynamics, suggesting that erosion at the downstream breach extremity is dictated by higher velocities than assumed in the initial dam model. This led to the introduction of an “effective breach width”, representing the fraction of the breach width conveying most of the water. Incorporating this concept considerably improved the model’s ability to predict the breach discharge and expansion.

However, accuracy on the breach discharge prediction remained limited, as it relied on simplified formulas borrowed from frontal weirs, which do not satisfactorily predict dike breach discharge. Various empirical or semi-empirical side weir formulations performed well for simplified side breach configurations but proved insufficient for dike breaching cases. Machine learning was then employed to better capture the breach hydrodynamics. A new analytical model was introduced and coupled to a decision-tree-based machine learning technique. This approach demonstrated strong performance in predicting the breach discharge from dike breaching laboratory tests.

The thesis also compared the adapted lumped model with its dynamic coupling to a 2D hydrodynamic model. Based on experimental observations and the global sensitivity analysis we conducted, the model parametrization was adjusted to further enhance breach erosion predictions. The adapted dike model outperformed the original dam model in predicting discharge and expansion for fluvial dike breach. It also provided more conservative flood extent predictions, as demonstrated for a hypothetical breach scenario along Belgium’s Albert Canal, making it particularly valuable for supporting evacuation planning in floodplains.

Future work should focus on additional test campaigns to refine the model structure and parametrization, enabling even greater accuracy in predicting fluvial dike breaching dynamics.

Résumé

Les digues fluviales, construites le long des cours d'eau, protègent les populations et les infrastructures contre les crues. Leur risque de rupture est en augmentation en raison du dérèglement climatique, de l'urbanisation croissante et du vieillissement des structures existantes. Prédire avec précision la dynamique de rupture et les inondations associées est essentiel pour une planification d'urgence efficace dans les zones inondables.

Cette thèse développe un modèle numérique dédié aux ruptures de digues fluviales homogènes et non cohésives induites par surverse. Contrairement aux barrages en remblai, qui sont perpendiculaires à l'écoulement principal et se rompent symétriquement, les brèches des digues s'étendent principalement dans la direction aval en raison de la quantité de mouvement importante de l'écoulement dans la rivière.

Dans cette thèse, un nouveau modèle non discrétisé spatialement et dédié aux digues fluviales a été mis au point à partir d'un modèle de rupture de barrages existant. Une analyse de sensibilité appliquée au modèle de rupture de barrage a permis d'identifier les paramètres dont l'incertitude a un impact important sur la variabilité du débit et de l'expansion de la brèche. Des essais en laboratoire ont permis de mieux comprendre l'influence de la géométrie de la digue et de la largeur du canal principal sur la dynamique de rupture. Il a également été mis en évidence que la vitesse de l'écoulement à proximité de l'extrémité aval de la brèche est sensiblement plus importante que la vitesse moyenne à travers la brèche. Cela a conduit à l'introduction d'une « largeur de brèche efficace », représentant la fraction de la brèche par laquelle transite la majeure partie du débit. L'utilisation de ce concept a considérablement amélioré les prédictions du modèle.

Cependant, la formulation du débit de brèche devait être revue car elle reposait sur des formules adaptées aux déversoirs frontaux, alors que la configuration étudiée correspond davantage à un déversoir latéral. Diverses formules empiriques de déversoirs latéraux ont été appliquées à des essais expérimentaux de rupture, mais leur précision s'est révélée insuffisante. Un nouveau modèle analytique a alors été introduit et couplé à une technique de machine learning basée sur les arbres de décision. Cette nouvelle approche a permis des prédictions du débit de brèche nettement plus précises.

Enfin, le nouveau modèle de rupture de digue a été comparé à son couplage avec un modèle hydrodynamique 2D. Avec une paramétrisation ajustée sur base d'observations expérimentales et de tests numériques, il s'est révélé plus fiable que le modèle de barrage, avec des prédictions plus conservatives, comme démontré dans un scénario de rupture hypothétique le long du canal Albert en Belgique.

De futurs travaux de recherche devraient inclure davantage d'essais expérimentaux, sur base desquels il serait possible d'affiner la structure et la paramétrisation du modèle et ainsi renforcer la précision des prévisions de rupture de digues fluviales.

Remerciements

Cette thèse est le fruit de nombreux échanges et rencontres. Je tiens donc à adresser mes remerciements à toutes celles et tous ceux qui y ont contribué, de près ou de loin, via leur aide scientifique ou tout simplement leur soutien.

L'environnement de travail dans lequel j'ai évolué m'a permis de pleinement m'épanouir en alliant à la perfection professionnalisme et moments de détente plus informels.

Mes pensées vont tout d'abord à Benjamin Dewals qui m'a accompagné tout au long de cette aventure. Son suivi m'a permis de garder le cap tout en me permettant d'explorer de nouvelles pistes. Je remercie également tout particulièrement Michel Pirotton pour sa bienveillance inconditionnelle et ses précieux conseils. Nos nombreux échanges m'ont permis de considérablement améliorer mes connaissances en hydraulique, et au-delà. Sur le plan numérique, Pierre Archambeau et Stéphane Champailler ont toujours été à l'écoute et désireux de m'aider. Leurs multiples bouées de sauvetage m'ont été vitales dans l'océan de codes que j'ai dû traverser par moments. Je remercie également Sébastien Erpicum, Maarten Arnst, Kamal el Kadi Abderrezzak et Frédéric Nguyen pour les échanges constructifs que nous avons pu avoir, notamment lors de mes comités de thèse.

Ce travail est aussi le fruit d'une collaboration plus étroite avec d'anciens collègues devenus académiques dans d'autres institutions. Merci à Vasilis Kitsikoudis and Xuefang Li pour le travail que nous avons pu accomplir ensemble jusqu'ici. Ce fût un réel plaisir de travailler ensemble, et j'espère que nous en aurons encore l'occasion à de nombreuses reprises dans le futur ! Je tiens également à remercier Grégoire Wylock et Julie Bouglé pour le travail qu'ils ont réalisé sous ma supervision lors de leur TFE et stage. Leur travail a clairement contribué à enrichir la thèse ici présente.

Ma thèse faisant suite à celle d'Ismail Rifai, je lui dois beaucoup pour avoir particulièrement bien documenté ses recherches et pour avoir été disponible lorsqu'une question subsistait. Le démarrage de ma thèse a également été grandement facilité par Louis Goffin, mon premier collègue de bureau et prédécesseur en tant qu'assistant. Son accueil chaleureux au +1/544 et sa transmission de savoir m'ont permis de me sentir à l'aise dès mon arrivée. Deux ans plus tard, ce fût à mon tour d'accueillir dans ce bureau un nouveau collègue, Clément Delhez, avec qui je n'aurais pu souhaiter meilleure entente. Outre notre complicité, nous avons construit ensemble un système de travail collaboratif et vertueux qui a permis à chacun de nous de se remettre en question et de regagner en motivation lorsque c'était nécessaire.

En termes de motivation, j'adresse mes remerciements aux fidèles de la grande cafet' de la première heure (Franz, Nayoun et Raph) et à ceux qui ont suivi (Gaby, Morgan, Angel et Maël). Nos temps de midi furent autant d'occasions de déconnecter avant d'entamer une après-midi de travail requinqué.

Dédicace spéciale à Maxime (DJ MX), Loris (le binôme), Solène (Soso), Damien (Piou-Piou) et Zoïs (papy) pour les bons moments et les fous-rires que nous avons partagés.

La bienveillance et l'entraide entre doctorants ont largement contribué à mon épanouissement dans l'équipe. Pour cela, je voudrais remercier Christophe, Joris, Pratik, Daniela, Utashi, Kawthar Maria Paula et Alexandre, mais aussi Anass, Julien, Carlos, Hector et Guillian. Merci tout particulièrement à Laurence Defrère de nous avoir gâté via toutes ses petites attentions, ainsi qu'à Max et Greg sur qui j'ai toujours pu compter pour me donner un coup de main lors de mes passages au labo.

Enfin, je remercie mes parents et ma belle-famille pour leur soutien indéfectible. Last but not least, merci à toi Audrey pour m'avoir soutenu, compris et supporté quotidiennement tout au long de ces cinq années de thèse (et au-delà).

Contents

Introduction

1. Context

For thousands of years, humans have built homes and infrastructures along rivers, which provide them with drinking water, food, fertile lands, transportation, and power generation. Besides these numerous benefits, rivers also represent a serious threat when flooding occurs, impacting surrounding communities and farmlands as water levels rise above the riverbanks.

To increase the banks' height and protect the hinterland from flooding, raised embankment structures have been built along rivers. These hydraulic structures, called fluvial dikes or levees, often stretch over long distances to prevent the flow from bypassing them. Ancient civilizations already adopted this defense system. The first significant dike system was constructed more than 3,000 years ago along the left bank of the Nile River (Hassan et al., 2017). Since then, their use has spread worldwide, with advancements in construction practices (Nakamura et al., 2024).

Fluvial dikes generally increase people's sense of safety, inducing a so-called "levee effect", i.e., people tend to settle down and expand their activities closer to the rivers (Ding et al., 2023). Although these areas become less prone to flooding, the consequences associated with a fluvial dike breaching increase significantly. A potential breach can result in more severe consequences than regular flooding. Instead of water gradually spilling into the floodplain, a large water volume is suddenly released, generating a highly energetic flow near the breach. This flow not only floods the area but also destroys structures and alters the topography close to the failure site (Olson et al., 2015). Furthermore, when the river level drops, flooding is prolonged as water retreats more slowly from the inundated lands due to the blockage caused by the remaining dikes. Additionally, poor dike maintenance and the surge of extreme weather events exacerbate the probability of dike failures.

Such catastrophic events happen regularly. For instance, in 2005, Hurricane Katrina hit the Southeastern part of the United States and caused the saturation and partial destruction of the levee system in New Orleans. Consequently, 80% of the city was flooded, with local water depths reaching up to 3m. Economic damages exceeded $135 billion, and about 1,000 people died (Plyer, 2016). After the event, the water took weeks to recede.

These events often trigger national-level reactions, such as the creation of dedicated organizations (e.g., National Levee Safety Program in the U.S.A.) that work to reduce flood risks for people, businesses, critical infrastructure, and the environment by optimizing evacuation plans, urbanization planning, etc. Within this context, a comprehensive understanding of the breaching dynamics of fluvial dikes is crucial to achieve a sound assessment of flood risks and to design appropriate countermeasures.

Fluvial dikes may fail due to several factors including overtopping, seepage, or soil instability. Among these, external erosion by overtopping is the most common cause of breaching in earthen dikes, i.e., the predominant type of dike (ASCE/EWRI Task Committee, 2011). So far, most works have focused on the breaching of dams, i.e., structures perpendicular to the main flow. While embankment dams tend to develop symmetrical breaches, breaches in fluvial dikes typically extend downstream due to the momentum of the river flow, which creates a highly

non-uniform flow profile through the dike breach (Kakinuma and Shimizu, 2014). Therefore, most analytical developments and empirical formulas designed for dams cannot be directly applied to fluvial dikes.

To address this gap, the present thesis is dedicated to understanding and modelling the breaching of fluvial dikes induced by overtopping. Here, the focus was set on homogeneous dikes made of non-cohesive material. This simple dike type provides a solid foundation for developing and validating models, making it an ideal starting point for investigating the complex dynamics of dike breaching.

2. Main contributions

This thesis primarily aims at improving our ability to model the breaching of homogeneous, non-cohesive fluvial dikes due to overtopping. The project is structured around the flowchart presented in Figure 1, which outlines its main features.

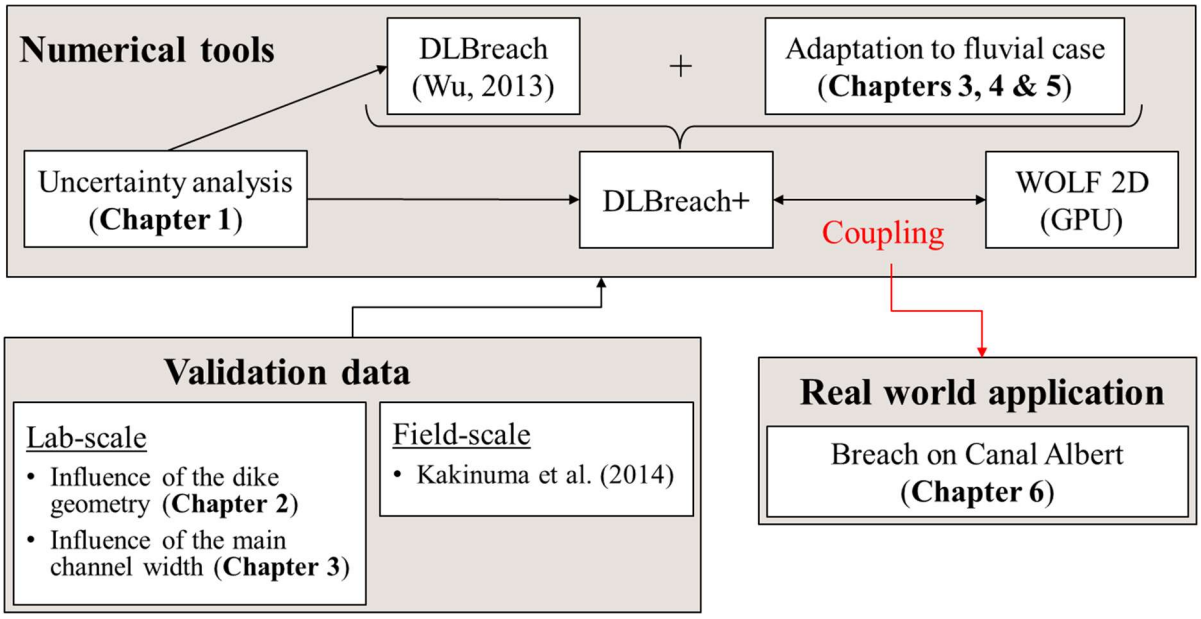


Figure 1. Flowchart of the thesis structure.

Dike or dam breach models can be classified into three categories: statistical (or parametric) models, distributed physically based models, and simplified physics-based models. Statistical models are computationally efficient but lack generality as they entirely rely on data from specific cases without considering underlying physics (De Lorenzo and Macchione, 2014; Chen et al., 2019; Lee, 2019). Distributed models are detailed and might provide more accurate results but to the expense of a significant computational load (Cantero-Chinchilla et al., 2019; Onda et al., 2019; Pheulpin et al., 2020; Shimizu et al., 2020). Simplified physics-based models offer an interesting trade-off. Without spatially distributing the flow description nor the embankment morphology, they enable simulating hydraulic and dike breach variables (e.g., time-evolution of breach discharge and dimensions) by describing selected physical processes (Wu, 2013; Zhong et al., 2017; Peter et al., 2018; Tsai et al., 2019; Li et al., 2020). For this reason, the adopted strategy was to reimplement a spatially non-discretized physics-based numerical model

initially developed for dam breaching, DLBreach (Wu, 2013, 2016), and to adapt it to a fluvial configuration.

A deeper understanding of the initial numerical model behavior was first required. As detailed in Chapter 1, a global sensitivity and uncertainty analysis was undertaken to gain insights into the model's response to variations in the input parameters and to reveal the global impact of input uncertainties on model outputs. The findings demonstrated that model results are highly dependent on the specific dam configuration considered, suggesting that results from sensitivity and uncertainty analyses conducted on a single configuration cannot be generalized.

Once the key parameters in the initial model were identified, the main differences in the breaching process of dams and dikes needed to be clarified to identify required model adaptations. Experimental tests were conducted to better understand the physical processes involved in fluvial dike breaching. Chapters 2 and 3 are respectively dedicated to studying the influence of dike geometry and main channel width on the breaching process, two parameters whose influence had not been systematically investigated until now.

Two fundamentally different processes were identified:

1. The non-symmetrical velocity distribution across the dike breach is critical for accurately capturing the non-symmetrical erosion process. In Chapter 3, a new parameter is introduced to account for the fact that flow through the breach is not uniformly distributed along the breach width but instead concentrated in the most downstream portion. This phenomenon induces greater water velocity near the downstream breach extremity (Kakinuma and Shimizu, 2014; Charrier, 2015), resulting in more intense erosion at this location. Conversely, limited erosion occurs at the upstream extremity, leading to largely non-symmetrical breach expansion in fluvial dikes. Incorporating this new parameter into the initial dam breaching model significantly enhances its prediction accuracy for breach expansion and breach discharge.

2. Due to the significant differences in the breach flow characteristics, empirical or semi-analytical formulas designed for dam breach discharge prediction are often unsuitable for fluvial dikes. In Chapter 4, the performance of various empirical dike breach discharge formulas is evaluated using both simplified and realistic dike breaching laboratory experiments. The analysis reveals that formulas performance highly depends on the tested configuration, and no single formula consistently outperforms the others. To address this, Chapter 5 introduces decision-tree-based machine learning models and a new analytical model to achieve better results. These findings show that integrating analytical or empirical models with machine learning significantly improves the accuracy of dike breach discharge estimation.

Returning to the thesis's initial goal, Chapter 6 compares two dike breaching modelling approaches against data from field experiments (Kakinuma et al., 2013). The first approach describes the flow and morphodynamic evolution of the dike breach using the simplified, physically based model adapted for fluvial configurations. The second approach uses a dynamic coupling of a 2D hydrodynamic model (WOLF 2D - GPU) with the breach evolution module

of the non-discretized dike breaching model, whose goal is now limited to describing the dike breach morphodynamics. Finally, the applicability of the latter approach is demonstrated based on a real-world case study of a potential breach along the Albert Canal in Wallonia (Belgium).

Chapter 7 concludes the thesis and provides perspectives for further enhancing fluvial dike breaching models.

References

ASCE/EWRI Task Committee, 2011. Earthen Embankment Breaching. Journal of Hydraulic Engineering 137, 1549–1564.

Cantero-Chinchilla, F.N., Castro-Orgaz, O., Dey, S., 2019. Prediction of Overtopping Dike Failure: Sediment Transport and Dynamic Granular Bed Deformation Model. Journal of Hydraulic Engineering 145.

Charrier, G., 2015. Etude expérimentale des ruptures de digues fluviales par surverse (Doctoral dissertation, Aix-Marseille).

Chen, Z.Y., Ping, Z.Y., Wang, N.X., Yu, S., Chen, S.J., 2019. An approach to quick and easy evaluation of the dam breach flood. Science China Technological Sciences 62, 1773–1782.

De Lorenzo, G., Macchione, F., 2014. Formulas for the peak discharge from breached earthfill dams. Journal of Hydraulic Engineering 140, 56–67.

Ding, M., Lin, P., Gao, S., Wang, J., Zeng, Z., Zheng, K., Zhou, X., Yamazaki, D., Gao, Y., Liu, Y., 2023. Reversal of the levee effect towards sustainable floodplain management. Nature Sustainability 6, 1578–1586.

Hassan, F.A., Hamdan, M.A., Flower, R.J., Shallaly, N.A., Ebrahem, E., 2017. Holocene alluvial history and archaeological significance of the Nile floodplain in the Saqqara-Memphis region, Egypt. Quaternary Science Reviews 176, 51–70.

Kakinuma, T., Shimizu, Y., 2014. Large-scale experiment and numerical modeling of a riverine levee breach. Journal of Hydraulic Engineering 140.

Kakinuma, T., Tobita, D., Yokoyama, H., Takeda, A., 2013. Levee breach observation at Chiyoda experimental flume, in: 12thInternationalSymposium River Sedimentation (ISRS), IRTCES, Kyoto, Japan.

Lee, K., 2019. Simulation of Dam-Breach Outflow Hydrographs Using Water Level Variations. Water Resources Management 33, 3781–3797.

Li, Y., Chen, A., Wen, L., Bu, P., Li, K., 2020. Numerical simulation of non-cohesive homogeneous dam breaching due to overtopping considering the seepage effect. European Journal of Environmental and Civil Engineering.

Nakamura, S., Nakai, F., Ito, Y., Okada, G., Oki, T., 2024. Levee system transformation in coevolution between humans and water systems along the Kiso River, Japan. Hydrology and Earth System Sciences 28, 2329–2342.

Olson, K., Matthews, J., Morton, L.W., Sloan, J., 2015. Impact of levee breaches, flooding, and land scouring on soil productivity. Journal of Soil and Water Conservation 70, 5A–11A.

Onda, S., Hosoda, T., Jaćimović, N.M., Kimura, I., 2019. Numerical modelling of simultaneous overtopping and seepage flows with application to dike breaching. Journal of Hydraulic Research 57, 13–25.

Peter, S.J., Siviglia, A., Nagel, J., Marelli, S., Boes, R.M., Vetsch, D., Sudret, B., 2018. Development of Probabilistic Dam Breach Model Using Bayesian Inference. Water Resources Research 54, 4376–4400.

Pheulpin, L., Bacchi, V., Bertrand, N., 2020. Comparison Between Two Hydraulic Models (1D and 2D) of the Garonne River: Application to Uncertainty Propagations and Sensitivity Analyses of Levee Breach Parameters, Springer Water.

Plyer, A., 2016. Facts for Features: Katrina Impact | The Data Center https://www.datacenterresearch.org/data-resources/katrina/facts-for-impact/ [Accessed 06-12-2024].

Shimizu, Y., Nelson, J., Arnez Ferrel, K., Asahi, K., Giri, S., Inoue, T., Iwasaki, T., Jang, C., Kang, T., Kimura, I., Kyuka, T., Mishra, J., Nabi, M., Patsinghasanee, S., Yamaguchi, S., 2020. Advances in computational morphodynamics using the International River Interface Cooperative (iRIC) software. Earth Surface Processes and Landforms 45, 11–37.

Tsai, C.W., Yeh, J., Huang, C., 2019. Development of probabilistic inundation mapping for dam failure induced floods. Stochastic Environmental Research and Risk Assessment 33, 91–110.

Wu, W., 2013. Simplified physically based model of earthen embankment breaching. Journal of Hydraulic Engineering 139, 837–851.

Wu, W., 2016. Introduction to DL Breach–A simplified physically based dam/levee breach model. Clarkson University, NY.

Zhong, Q., Chen, S., Deng, Z., 2017. Numerical model for homogeneous cohesive dam breaching due to overtopping failure. Journal of Mountain Science 14, 571–580.

Chapter 1

Global Sensitivity Analysis of a Dam Breaching Model: To Which Extent Is Parameter Sensitivity Case-Dependent?

This chapter corresponds to the journal paper "Global Sensitivity Analysis of a Dam Breaching Model: To Which Extent Is Parameter Sensitivity Case-Dependent?" by V. Schmitz, M. Arnst, K. El Kadi Abderrezzak, M. Pirotton, S. Erpicum, P. Archambeau, and B. Dewals, published in 2022 in Water Resources Research. The PhD candidate developed the methodology, implemented the numerical models, analyzed the results, wrote the manuscript and generated the figures.

Global Sensitivity Analysis of a Dam Breaching Model: to Which Extent is Parameter Sensitivity Case-Dependent?

V. Schmitz[1], M. Arnst[2], K. El Kadi Abderrezzak[3,4], M. Pirotton[1], S. Erpicum[1], P. Archambeau[1] & B. Dewals[1]

[1] *Research Group of Hydraulics in Environmental and Civil Engineering (HECE), University of Liège, Liège, Belgium.*

[2] *Computational and Stochastic Modeling, University of Liège, Liège, Belgium*

[3] *National Laboratory for Hydraulics and Environment (LNHE), EDF R&D, Chatou, France.*

[4] *Saint Venant Laboratory for Hydraulics, Chatou, France.*

Corresponding author: Vincent Schmitz (V.Schmitz@uliege.be)

Key Points:

- Based on a simplified physically based dam breaching model, a global sensitivity analysis was conducted using Sobol indices of total order.
- The analysis was applied on twenty-seven configurations at both laboratory and field scales.
- Most influential parameters and output uncertainty magnitudes were assessed and turned out to strongly depend on the dam configuration.

ABSTRACT

Failure of dams and dikes often leads to devastating consequences in protected areas. Numerical models are instrumental tools to assess flood risk and guide emergency management, but numerous uncertainties affect model outcomes. Identifying uncertain model input parameters that induce high uncertainties in model outputs is essential. This paper focuses on two model outputs: the maximum breach discharge and the time to reach this peak. Using our implementation of a simplified physically based dam breaching model developed by Wu (2016), a global sensitivity analysis was conducted based on Sobol indices of total order. Unlike in most previous studies, many different configurations (twenty-seven), both at laboratory and field scales, were considered. For each of them, input variables were ranked according to the significance of the contribution of their uncertainty to the output variability, and the dependency between reference configurations and sensitivity analysis results was highlighted. Depending on the considered case study, input parameter uncertainties with the largest impact on output variability were different, and so was the magnitude of output uncertainties. We demonstrate that sensitivity analysis results obtained for a specific configuration cannot be transferred as it to other configurations. Finally, sensitivity and uncertainty quantification results were combined in a decision tree to determine which input parameter uncertainty is the most critical in each configuration and what standard deviation in the output variables is expected. The global sensitivity analysis procedure presented here may apply to a wide variety of models in environmental sciences.

1. Introduction

With the surge of extreme meteorological events and intensification of urbanization downstream of hydraulic structures, the need for predicting failure of dams and dikes has become of paramount importance for establishing emergency response procedures (Zhong et al., 2021). To this end, numerical models are instrumental tools for simulating the embankment breaching process. Existing models can be classified into three categories. First, statistical (or parametric) models are based solely on regression analysis of data from past events or laboratory campaigns. They describe some breaching parameters (e.g., final breach width, failure duration or maximum breach discharge) as a function of dam or reservoir properties. These simple, computational-efficient models may lack generality because they entirely rely on data from specific cases without considering underlying physics (De Lorenzo and Macchione, 2014; Chen et al., 2019; Lee, 2019). Conversely, distributed physically based models can describe the phenomenon in greater detail, as they solve the flow and sediment governing equations using a computational mesh of the domain. Their results may be accurate but only if reliable data is available and if physical processes are numerically well represented, e.g., erosion of non-homogeneous dam material, slope failure or 3D-flow patterns (Cantero-Chinchilla et al., 2019; Onda et al., 2019; Pheulpin et al., 2020; Shimizu et al., 2020). Additionally, the time required to run distributed physically based models can be substantial (ASCE, 2011). Simplified physics-based models offer a good trade-off (Wu, 2013). Without spatially distributing the flow description nor the embankment morphology, they enable simulating hydraulic and dam breach variables (e.g., time-evolution of breach discharge and dimensions) by describing selected physical processes (Wu, 2013; Zhong et al., 2017;

Peteretal., 2018; Tsai et al., 2019; Li et al., 2020). Simplified physics-based models are computationally efficient and enable uncertainty analyses that require many runs, e.g., the Monte Carlo method. In this paper, breaching of homogeneous dams made of non-cohesive material is computed using our implementation of the physics-based lumped fully coupled hydro-morphodynamic model developed by Wu (2016).

In all numerical models, input variables can be subject to uncertainties, which may emerge from multiple sources (Kiureghian and Ditlevsen, 2009; Peter et al., 2018). Depending on their origin, they are generally classified as parametric or modelling uncertainties. The first group gathers uncertainties on physical parameters. They may originate either from inaccuracies in measurements or from the natural variability of parameters (e.g., spatial distribution of geometrical parameters or soil composition of embankments). Modelling uncertainties can be encountered when uncertainties arise from the model structure itself (Gupta et al., 2012), or when using parameters tied to parametrization of complex physical phenomena (e.g., regression parameters). Assigning purely deterministic values to uncertain parameters can be suboptimal and alternative characterizations may be preferred, such as variation intervals or probabilistic distributions.

Uncertainty analyses were developed to gain insights into model response to variations in input parameters (Borgonovo and Plischke, 2016). They are generally divided into three main steps (Helton et al., 2006; Arnst and Ponthot, 2014): (1) uncertainty characterization, (2) uncertainty propagation, and (3) exploitation of the results in accordance with the study goals. In Step 1, probability distributions are attributed to input variables. In Step 2, a numerical model is used to propagate those uncertainties to the model outputs. In Step 3, a common approach consists in leading a sensitivity analysis to better understand the relation that exists between output and input uncertainties. Sensitivity analyses are defined as local or global (Reed et al., 2022). In the first case, uncertainties in the model outputs are evaluated by slight modifications in the input values around a reference configuration. Although being computationally efficient, this approach is not suitable for nonlinear models as it explores a limited fraction of the input space and often assumes that input variables do not interact in the model (Saltelli and Annoni, 2010; Rakovec et al., 2014; Saltelli et al., 2019). To tackle those limitations, global sensitivity analyses are led to reveal the global impact of input uncertainties on model outputs by exploring the entire input space of interest.

Four main sensitivity analysis settings were listed by Saltelli et al. (2004), namely factor (or input parameter) prioritization, factor fixing, variance cutting and factor mapping. When the pursued objective is to identify the uncertain input parameters that induce the most critical variability in model outputs, one refers to factor prioritization. Conversely, factor fixing aims at pointing out input parameters whose variability has a negligible impact on model outputs, allowing to discard their influence in subsequent analyses. Variance cutting is used to assess the minimum number of input parameters that should be fixed to limit the uncertainty on model outputs below a prescribed threshold. Finally, factor mapping identifies which input values lead to output values contained in a specific range of the output space. This approach allows highlighting combinations of input values that lead to non-physical results (Spear and Hornberger, 1980; Pianosi et al., 2016).

This paper highlights the dependency between global sensitivity analysis results and test configuration through factor prioritization and factor fixing approaches. Within this context, variance-based methods are convenient as they rank inputs or subsets of inputs as a function of the influence of their uncertainty on the variability of resulting outputs (Saltelli et al., 2008). Among them, the use of Sobol indices of total order is widespread (Sobol, 1993; Saltelli et al., 2010). Those indicators illustrate the portion of the global output variance caused by the uncertainty on a specific input subset, including the interactions of this subset with any other input subset.

While previous studies were often limited to the analysis of a single case study (Tables S1 and S2 in Supplement), we apply here a global sensitivity analysis based on Sobol indices of total order to twenty-seven embankment configurations, in both laboratory and field scales. The results of this study highlight that the input parameter uncertainties contributing most to the output variability differ considerably from one configuration to the other. The magnitude of output uncertainties also varies with the considered configuration. Here, we demonstrate that sensitivity analysis results obtained for a specific dam breaching case may not be transferred to any other case.

The remaining of the paper is organized as follows: the computational model and the simulated dam configurations are introduced in Section 2. The input sampling methods and the global sensitivity analysis technique are described in Section 3. In Section 4, results of the sensitivity analysis are presented and discussed for twenty-seven tests in laboratory and field-scale configurations. Conclusions are drawn in Section 5.

2. Physical data and methods

This section starts with the description of our implementation of a model initially developed by Wu (2016), i.e., the numerical model on which the sensitivity analysis was applied (Section 2.1). It is composed of three interacting modules: (i) a hydrodynamic module, (ii) a sediment transport module and (iii) a morphodynamic module, in which the breach geometry is described as the combination of a flat top reach corresponding to the top of the breach and a downstream reach located on the dike floodplain face (Figure 1). A brief description of the considered experimental configurations is provided in Section 2.2.

2.1. Computational model

Based on initial and boundary conditions, the model hydrodynamic module evaluates the water level in the main channel and on the different parts of the embankment, the breach discharge, and the resulting shear stress using two main assumptions: critical flow on the dam crest and uniform flow on the downstream face. Those values are then introduced in the sediment transport module to compute both suspended-load and bed load concentrations in the flow through the breach assuming non-equilibrium sediment transport. Using the eroded material volume, the new dam geometry is generated by the morphodynamic module. The procedure is repeated by feeding the hydrodynamic module with the updated dam geometry. At the start of the computation, the geometry of an initial notch in the dam crest needs to be defined by the user, as the model is not able to predict the location of breach initiation.

The computational model was coded using *Matlab* software and explicit resolution schemes were used. The time step was 0.5 s for all tests. The present section summarizes the three modules of the numerical model, and detailed flow charts are provided in Supplement (Figures S1, S2 and S3).

2.1.1. Hydrodynamic module

The hydrodynamic module (Figure S1, in Supplement) computes first the water level in the main channel using mass balance:

$$\frac{dz_s}{dt} = \left[Q_{in} - Q_b - Q_d - Q_{out}(z_s)\right]\frac{1}{A_{res}}, \tag{1}$$

with z_s the water level in the main channel, A_{res} the water area in the main channel, Q_{in} the inflow discharge, Q_b the discharge through the breach, Q_d the drain discharge and Q_{out} the outflow discharge at the downstream end of the main channel. If the water level is lower than the breach bottom elevation, z_b, the breach discharge remains zero. Otherwise, the breach discharge is computed as

$$Q_b = c_1 b_{top}\left(z_s - z_b\right)^{3/2} + c_2 m\left(z_s - z_b\right)^{5/2}, \tag{2}$$

with c_1 and c_2 two weir coefficients, b_{top} the breach bottom width on the flat top reach (Figure 1) and m the breach side slope. Since non-cohesive dams are considered here, m can also be

expressed as $m = (\tan \varphi_r)^{-1}$, where φ_r is the dam material repose angle considered similar for both wet and dry material.

The water levels on the flat top and downstream reaches (Figure 1) were computed assuming critical flow on the dam crest and uniform flow over the downstream face.

$$\begin{aligned} &\text{Flat top reach:} && h_{top} = \max\left(\frac{2}{3}(z_s - z_b); 0\right), \\ &\text{Downstream reach:} && Q_b = \frac{1}{n} A R^{2/3} S_d^{-1/2}, \end{aligned} \tag{3}$$

with A the flow area over the considered reach, R the hydraulic radius, S_d the slope of the breach downstream side (H:V) and n the Manning roughness coefficient computed based on Strickler's formula:

$$n = \max\left(d_{50}^{1/6} / A_n; n_{min}\right), \tag{4}$$

where d_{50} is the median size of the embankment material, A_n an empirical coefficient and n_{min} the minimum value allowed for n, as defined by Wu (2016).

On both reaches, the bed shear stress can then be calculated as

$$\tau_b = \frac{\rho g n^2 Q_b^2}{A^2 R^{1/3}}, \tag{5}$$

where ρ is the water density and g the gravitational acceleration.

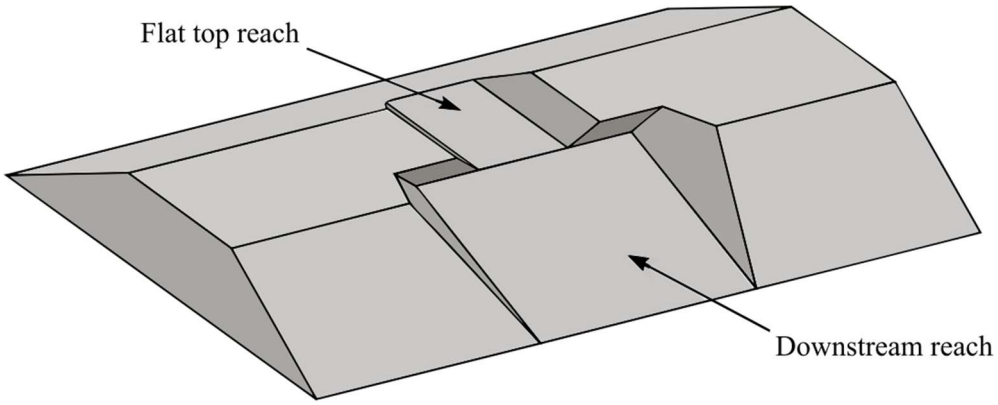


Figure 1. Scheme of the embankment geometry used in Wu's (2016) model.

2.1.2. Sediment transport module

The bed shear stress is introduced in the sediment transport module (Figure S2, in Supplement) to evaluate the eroded material volume on each reach. The effective shear stress is computed as follows:

$$\tau_e = \tau_b' + \lambda_0 \tau_c \frac{\sin\phi}{\sin\varphi_r}, \tag{6}$$

with ϕ the angle between a horizontal plane and the considered reach, i.e., ϕ = 0 on the flat top reach and $\phi = \tan^{-1}\left(1/S_d\right)$ on the downstream slope. $\tau_c = \theta_{cr}\left(\rho_s - \rho\right) g d_{50}$ is the critical shear stress involving the critical Shields parameter, θ_{cr}, and the sediment density, ρ_s. τ_b' is the grain shear stress based on an empirical coefficient $A_{n'}$ and writes

$$\tau_b^{'} = \left(\frac{n'}{n}\right)^{3/2} \tau_b \qquad \text{with} \quad n' = \max\left(\frac{d_{50}^{1/6}}{A_{n'}}; n_{min}\right). \tag{7}$$

λ_0 is a correction factor introduced by Wu (2016):

$$\lambda_0 = 1 + \lambda_{0,a}\left(\frac{\tau_b'}{\tau_c}\right)^{\lambda_{0,b}} e^{2\sin\phi/\sin\varphi_r}, \tag{8}$$

with $\lambda_{0,a}$ and $\lambda_{0,b}$ two empirical coefficients.

The equilibrium sediment concentration on both reaches is made of two contributions:

$$\begin{aligned} &\text{Suspended load:} \quad C_* = \frac{1}{C_a^* \rho_s}\left(\frac{U^3}{gR\omega_s}\right)^{C_b^*} * \left[1 + \left(\frac{1}{C_c^*}\frac{U^3}{gR\omega_s}\right)^{C_d^*}\right]^{-1}, \\ &\text{Bed load:} \qquad q_b^* = q_{b,a}^*\left(\max\left[\frac{\tau_e}{\tau_c} - 1; 0\right]\right)^{q_{b,b}^*} \sqrt{\left(\frac{\rho_s}{\rho} - 1\right) g d_{50}^3}, \end{aligned} \tag{9}$$

with $\left\{C_a^*, C_b^*, C_c^*, C_d^*\right\}$ and $\left\{q_{b,a}^*, q_{b,b}^*\right\}$ empirical coefficients obtained through experimental data fitting (Zhang, 1961; Wu et al., 2000), U the flow velocity on the considered reach and ω_s the sediment settling velocity computed using the formula of Wu and Wang (2006), which is based on sediment density, ρ_s, sediment median size, d_{50}, and Corey shape factor, S_p.

The actual sediment concentration at the outlet of each reach, $C_{t,out}$, is linked to the actual concentration at the inlet, $C_{t,in}$, through the following formula:

$$C_{t,\text{out}} = C_{t^*} + (C_{t,\text{in}} - C_{t^*})\exp(-\frac{\Delta}{L_s}), \tag{10}$$

with $C_{t*} = C_* + B_w q_b^* / Q_b$ the total equilibrium sediment concentration and B_w the width of the water surface. Δ is the length of the considered reach and $L_s = \lambda B_w$ is the mixing length characterizing the adjustment of sediment from a non-equilibrium state to the equilibrium state. λ is a numerical parameter to be adjusted in each simulation.

The eroded volume is computed on each reach using the following equation, in which p denotes the dike material porosity:

$$\frac{dV_b}{dt} = \frac{Q_b(C_{t,in} - C_{t,out})}{1-p}. \tag{11}$$

2.1.3. Morphodynamic module

In line with Wu (2016), uniform erosion on all breach faces is assumed. The morphodynamic module (Figure S3, in Supplement) computes the breach depth variation on each reach as

$$\frac{dz_b}{dt} = \frac{1}{A_{e,tot}} \frac{dV_b}{dt} \quad \text{with} \quad A_{e,\text{tot}} = A_{\text{bottom}} + 2A_{\text{side}}, \tag{12}$$

where A_{bottom} and A_{side} are the breach bottom and breach side areas of the considered reach, respectively.

The breach bottom width variation on the flat top reach is computed as

$$\frac{db_{top}}{dt} = 2\frac{dz_b}{dt}\left(\frac{1}{\sin\varphi_r} - \frac{1}{\tan\varphi_r}\right), \tag{13}$$

whilst the breach top width variation on this reach writes

$$\frac{dB_{top}}{dt} = \frac{2}{\sin\varphi_r}\frac{dz_b}{dt}. \tag{14}$$

On the downstream reach, Wu (2016) assumed that the breach width variation is influenced by the ratio between the breach width on the flat top reach and on the downstream reach. He introduced a correction factor c_b expressed as

$$c_b = \min\left[1, \max\left(0, c_{b,coef}\frac{b_{\text{top}}}{b_{\text{D/S}}} - \left(c_{b,coef} - 1\right)\right)\right], \tag{15}$$

where $c_{b,coef}$ is an empirical parameter and b_{top} and $b_{\text{D/S}}$ are the breach bottom widths on the flat top and the downstream reach, respectively.

Breach width variations on the downstream reach write

$$\frac{db}{dt} = 2\frac{dz_b}{dt}\left(\frac{\max(c_b, \cos\varphi_r)}{\sin\varphi_r} - \frac{1}{\tan\varphi_r}\right) \quad \text{and} \quad \frac{dB}{dt} = \frac{2c_b}{\sin\varphi_r}\frac{dz_b}{dt}. \tag{16}$$

To close the loop, updated geometrical variables are finally introduced in the hydrodynamic module and a new iteration is performed.

2.2. Test cases

The sensitivity analysis was applied to various experimental tests to assess how the configuration influences the relation between uncertainties on input and output variables. Frank (2016) investigated experimentally the spatial dam breach process due to overtopping for dams made of homogeneous non-cohesive material. A total of forty-five three-dimensional embankment breach tests in frontal configuration and laboratory conditions were conducted. The influence of several parameters was analysed, including dam scaling, inflow discharge, sediment grain size, main channel width, dam cross-section, and reservoir volume. In the latter case, the pumping system was regulated to simulate a reservoir of a larger volume than the real physical reservoir. For each test, hydrographs (breach, inflow, and drainage discharges) and the upstream water level were recorded. Twenty-seven laboratory experiments performed by Frank (2016) were first considered. Table 2 summarizes the non-dimensional characteristics of these tests on which an uncertainty analysis was applied while dimensional data are gathered in Table S3 in Supplement. Since Frank (2016) considered half embankments, parameter values have been adapted in this study to represent entire dams. To be consistent with other configurations, both dam side slopes were set equal to 1:2 (V:H) in Test 28, whilst equal to 1:1.5 in Frank's (2016) original experiments.

To assess the impact of geometric scale on the results of sensitivity analysis, an upscaling was applied to the laboratory tests to obtain field-scale configurations: dike heights were multiplied by a factor 10, while non-dimensional parameters were kept constant. The objective here is to show how dependencies between uncertainties in model inputs and model outputs evolve with the geometric scale. It is not intended to investigate how physical processes differ between the laboratory and field scales, but merely to assess potential differences in model behaviour when the geometric scale is varied.

Table 1. Summary of parameters involved in the numerical model.

	Symbol	Description
Input parameters	S_u, S_d	Upstream and downstream slopes of the dam, respectively
	L_k	Dam crest length
	h_d	Dam height
	φ_r	Sediment repose angle
	ρ_s	Sediment density
	d_{50}	Sediment median size
	p	Dam material porosity
	Q_{in}	Inflow discharge
Model parameters	c_1, c_2	Weir efficiency coefficients involved in the breach discharge evaluation (Eq. (2))
	$A_n,\ A_{n'},\ n_{min}$	Parameters in Strickler's formula: $n = \max\left(\frac{d_{50}^{1/6}}{A_n}; n_{min}\right); n' = \max\left(\frac{d_{50}^{1/6}}{A_n'}; n_{min}\right)$, with n and n' total and effective Manning's coefficients, respectively
	θ_{cr}	Critical Shields parameter
	$\lambda_{0,a}, \lambda_{0,b}$	Coefficient and exponent involved in the effective shear stress computation (Damgaard et al., 1997)
	S_p	Corey shape factor
	$C_{*a}, C_{*b}, C_{*c}, C_{*d}$	Regression coefficients involved in the suspended load concentration (Eq. (9))
	$q^*_{b,a}, q^*_{b,b}$	Regression coefficients involved in the bed load transport capacity formula (Eq. (9))
	λ	Empirical coefficient involved in mixing length computation
	$c_{b,coef}$	Artificial breach widening limitation coefficient (Eq. (15))

Table 2. Frank's (2016) experimental parameters. L_r = reservoir length; l_r = reservoir width; A_r = reservoir area; h_d = dam height; L_k = dam crest length; $z_{n,ini}$ = initial notch depth; B_{ini} = initial notch width; d_{50} = median grain size; Q_{in} = inflow discharge; h_{cr} = critical flow depth.

	Test ID	L_r/h_d	l_r/h_d	A_r/h_d^2	h_d (mm)	L_k/h_d	$z_{n,ini}/h_d$	B_{ini}/h_d	d_{50}/h_d	d_{50} (mm)	Q_{in} (l/s)	h_{cr}/h_d
Scaling	11				150				$5.7\ 10^{-3}$	0.86	2.62	
	10	4.07	10/3	13.57	300	1/3	1/3	4/3	$5.8\ 10^{-3}$	1.75	14.8	$9.4\ 10^{-2}$
	8				600				$6.3\ 10^{-3}$	3.78	83.6	
	12				150				$5.7\ 10^{-3}$	0.86	1.31	
	13	4.07	10/3	13.57	300	1/3	1/3	4/3	$5.8\ 10^{-3}$	1.75	7.4	$5.9\ 10^{-2}$
	14				600				$6.3\ 10^{-3}$	3.78	41.8	
	15				150					0.43	2.62	
	16	4.07	10/3	13.57	300	1/3	1/3	4/3	$2.9\ 10^{-3}$	0.86	14.8	$9.4\ 10^{-2}$
	17				600					1.75	83.6	
Discharge	18										4.4	$2.6\ 10^{-2}$
	19										8.8	$4.2\ 10^{-2}$
	20	11.47	20/3	76.47	300	1/3	1/3	4/3	$5.8\ 10^{-3}$	1.75	18.4	$6.8\ 10^{-2}$
	21										35.2	$10.5\ 10^{-2}$
Sediment	22								$2.9\ 10^{-3}$	0.86		
	20	11.47	20/3	76.47	300	1/3	1/3	4/3	$5.8\ 10^{-3}$	1.75	18.4	$6.8\ 10^{-2}$
	23								$1.3\ 10^{-2}$	3.78		
Initial breach width	20							4/3				
	24							5/3				
	25	11.47	20/3	76.47	300	1/3	1/3	2	$5.8\ 10^{-3}$	1.75	18.4	$6.8\ 10^{-2}$
	26							8/3				
	27							4				
Crest length	28					0						
	20	11.47	20/3	76.47	300	1/3	1/3	4/3	$5.8\ 10^{-3}$	1.75	18.4	$6.8\ 10^{-2}$
	29					4/3						
Reservoir water surface area	40			76.4								
	36			299								
	41			367								
	42	11.47	20/3	521	300	1/3	1/15	4/15	$5.8\ 10^{-3}$	1.75	0	0
	43			1410								
	44			2298								

3. Uncertainty quantification data and methods

This section first provides a general description of uncertainty analysis procedures (Section 3.1). Uncertain input variables are then characterized in the numerical model, and uncertainty quantification methods are presented for independent and dependent input variables (Section 3.2). Using the standard Monte Carlo method, uncertainty is propagated through the computational model and a global sensitivity analysis based on Sobol indices is presented (Section 3.3).

3.1. Uncertainty analysis procedures

Step 1: Uncertainty characterization of input variables

Model predictions are directly affected by input uncertainties, highlighting the need for systematically quantifying them to reliably predict physical processes, such as embankment failures (Pappenberger and Beven, 2006; Froehlich, 2008; ASCE, 2011). Different approaches exist to quantify uncertainties on physical parameters, depending on data availability and modelling objectives, i.e., description of a specific case or general representation of typical embankment failures. In the latter case, when the data set is large, maximum likelihood estimation is often used to define parameter distributions (Rossi, 2018). In contrast, when scarce data is available, Bayesian inference is preferred (Saltelli et al., 2004). When information on parameters of interest is too limited or not available at all (e.g., model parameters), plausible parameter distributions may be defined based on modeler expertise. Dependency between input variables is seldom considered, while this may influence input probability distributions significantly (Jacques et al., 2006; Da Veiga et al., 2009, 2021; Li et al., 2010; Baroni and Tarantola, 2014; Pheulpin et al., 2022). In the present case, nineteen independent input variables and two sets of dependent input variables are involved. Specific embankment configurations are considered, for which almost no information on parameter distributions is available. Therefore, probability density functions (PDF) associated to each input were mostly generated based on expert judgement, with a special care for dependent input variables (see Sections 3.2.1 and 3.2.2).

Step 2: Uncertainty propagation

The uncertainty in input variables must be propagated through the numerical model to assess its impact on the outputs. This procedure is crucial to properly interpret model outcomes, since it determines the variation range of model predictions (Pappenberger and Beven, 2006). It is most often implemented using a Monte Carlo sampling method (Metropolis and Ulam, 1949). Many simulations are run using different sets of input variable values generated according to their probability distributions. As the number of runs increases, approximations of the model output statistical moments converge, e.g., mean and variance (Janssen, 2013). The classical Monte Carlo sampling method is computationally expensive, making it suitable for fast models only. To reduce the number of runs required to reach a given accuracy on output statistical moments, advanced Monte Carlo sampling methods were developed (Hou et al., 2019). One of the most frequently used is the Latin Hypercube sampling method (McKay et al., 1979). Instead of selecting input values purely based on their probability distributions, this method favours sets of input values located in underrepresented areas of the input space

(Helton and Davis, 2003). When considering smooth but computationally expensive models containing a limited number of input variables, a surrogate model (or metamodel) is often fitted to the original numerical model to mimic the relationship between input and output variables while improving computational efficiency (Jin et al., 2001; Queipo et al., 2005; Sudret, 2008). An alternate method is the point estimate method initially introduced by Rosenblueth (1975). Based on a limited number of model evaluations, it aims at computing output statistical moments using statistical moments of the input variables. In this paper, uncertainty propagation was performed using our implementation of the physically-based fully coupled hydro-morphodynamic model developed by Wu (2016). Thanks to the computational efficiency of this model, a regular Monte Carlo sampling method was adopted here.

Step 3: Results exploitation

Uncertainty propagation results are generally exploited in two different ways. On the one hand, uncertainty quantification focuses on the output variability characterization, including output main statistical moments, e.g., mean and variance, or complete probability distribution (Walker et al., 2003). On the other hand, sensitivity analyses highlight the relationships between uncertainties on specific input and output variables (Borgonovo and Plischke, 2016). Among them, global sensitivity analyses are generally preferred as derivatives and other local methods provide an incomplete picture of model response over the range of variability in the model inputs (Saltelli et al., 2019). In this context, screening techniques are widespread, with Morris method being the most common (Morris, 1991). This method evaluates two indicators. The first one reflects the mean output sensitivity to the variation of a specific input variable, while the second one describes the nonlinear behaviour of this dependency within the input space (Iooss and Lemaître, 2015). Regression-based analyses provide an algebraic expression of the relationships between outputs and input variables (Helton et al., 2006). One of the simplest and most popular approaches is the least square linearization technique, which consists in a multi-linear regression between model output and input variables. A coefficient, characterizing the relative contribution of each input uncertainty to the output variability, is assigned to each input. The larger this coefficient, the larger the impact of the considered input on the output variability. An example of more advanced regression-based analysis is the permutation feature importance, which relies on machine learning regression and evaluates how permutations of input parameters alter model predictions (Breiman, 2001). Moment-independent methods are usually preferred when the focus is set on the entire output distribution, especially when the output distribution is highly skewed or when it is multi-modal. The delta method computes sensitivity indicators that represent the normalized expected shift in the output distribution induced by uncertainties on a specific input variable (Borgonovo, 2006, 2007; Pianosi and Wagener, 2015). The use of a variance-based sensitivity analysis is shown to be particularly general in its applicability and in its capacity to reflect nonlinear processes and the effects of interactions among variables (Hall et al., 2009). The Sobol index of first order represents the output variability caused by a specific uncertain input subset. In contrast, the Sobol index of total order also includes the interactions of this subset with any other input subsets. Both indices aim at assessing the relative significance of input variables uncertainties (Iooss and Lemaître, 2015). It allows spotting the most critical input uncertainties as well as the non-influential ones, i.e., factor prioritization and factor fixing.

Tables S1 and S2 in Supplement provide an overview of existing studies applying uncertainty analysis to dam or dike breach models. In each of these studies, only a single case study was considered. Part of them solely focused on uncertainty quantification, e.g., Abdedou et al. (2020), Froehlich (2008), Froehlich and Goodell (2012), Peter et al. (2018), Vorogushyn et al. (2011) and Westoby et al. (2015). Others performed global sensitivity analysis. Among them, Alhasan et al. (2016) and Bellos et al. (2020) used a Morris screening technique, whilst Ahmadisharaf et al. (2016), Goeury et al. (2022) and Sattar (2014) performed regression-based analyses. Goeury et al. (2022) also used the moment-independent delta method because of the presence of multi-modal output distributions. Variance-based methods, such as Sobol indices, were used by Kalinina et al. (2020), Pheulpin et al. (2020) and Tsai et al. (2019), but also considering each only a single case study. In contrast, in the present work, Sobol indices of total order were computed for each input subset in twenty-seven embankment configurations both at laboratory and field scales.

3.2. Input distributions and sampling methods

Uncertain input variables are gathered in Table 1. They are divided into two types, namely model parameters that need to be set based on experimental observations and input parameters, e.g., material characteristics and dam geometry parameters.

Values of the different input variables are selected randomly according to their respective probability distributions to feed the numerical model. It is therefore necessary to define a probability distribution function (PDF), or multivariate probability distribution assigned to each input variable or group of input variables, respectively. The approach differs for variables whose values are assumed independent from each other and variables with dependent values. Both procedures are introduced hereafter.

3.2.1. Independent input variables

A Beta distribution has been assigned as a PDF to all independent variables due to its high versatility, wide-spread use in sensitivity analyses in environmental sciences and beyond (e.g., Benke et al. (2008), Hall et al. (2005), Kalinina et al. (2020) and Mokhtari and Frey (2005)), and its finite support. This last feature is of great interest as it allows narrowing the analysis to behavioural input samples, e.g., avoiding negative values for uncertain physical inputs (Wagener and Pianosi, 2019). It is expressed as

$$PDF(x;\alpha,\beta)=\frac{x^{\alpha-1}\left(1-x\right)^{\beta-1}}{\mathrm{B}(\alpha,\beta)} \qquad \text{with} \quad x=\frac{\tilde{x}-\tilde{x}_{min}}{\tilde{x}_{\max}-\tilde{x}_{min}}\in[0;1], \tag{17}$$

where x is the normalized form of the considered variable $\tilde{x}$ that varies within the interval $[\tilde{x}_{min};\tilde{x}_{\max}]$, α and β two real and positive parameters to be fixed and B the beta function.

In this work, we determined the values of parameters α and β from two features of the input variable x. First, we enforced the PDF to exhibit a maximum at a reference value, x_{ref}, so that

$$\left.\frac{\mathrm{d}PDF(x;\alpha,\beta)}{\mathrm{d}x}\right|_{x=x_{ref}} = 0 \quad \text{with} \quad x_{ref} \in]0;1[. \tag{18}$$

The second feature we considered is the variance of the variable x. However, due to the approximate knowledge of this characteristic, a qualitative approach based on expert knowledge is used here. One of the two beta distribution parameters is fixed manually while the other is determined using Eq.(18). PDF characteristics of each uncertain input variable of the present study are presented in Table S4 in Supplement.

Note that the choice of a particular type of PDF and its parameters may have an influence on the global sensitivity analysis results. This aspect is discussed in Section 0.

3.2.2. Dependent input variables

A multivariate probability distribution is needed to represent input dependency properly and generate relevant input samples. The typical case of regression coefficients $\left\{C_a^*, C_b^*, C_c^*, C_d^*\right\}$ and $\left\{q_{b,a}^*, q_{b,b}^*\right\}$ obtained through experimental data fitting (Wu, 2016) is handled here. The presented methods (Methods 1 and 2) aim at generating sets of regression coefficients that lead to meaningful regression curves with regards to experimental observations (Figure S4, in Supplement). No explicit expression of the multivariate probability distribution is provided through these methods.

Method 1: sample generation based on bootstrapping

Let an empirical data vector containing n experimental points be written as

$$\boldsymbol{d} = \{d_1, d_2, \ldots, d_n\} \text{ with } d_i = (x_{exp,i}, y_{exp,i}). \tag{19}$$

Based on these n experimental points, the regression function, f, can be expressed as

$$y_{reg} = f(\boldsymbol{c}, x), \tag{20}$$

with $\boldsymbol{c}$ the vector containing the regression coefficients.

k samples of length n are generated with replacement from the experimental points of vector $\boldsymbol{d}$. This procedure, called *bootstrapping* (Davison and Hinkley, 1997), leads to k vectors of length n: $\{\boldsymbol{d}^1, \boldsymbol{d}^2, \ldots, \boldsymbol{d}^k\}$. For each vector, the value of the parameters contained in $\boldsymbol{c}$ is determined to best fit the experimental data using least square minimization method in a trust-region algorithm, leading to k sets of regression coefficients: $\{\boldsymbol{c}^1, \boldsymbol{c}^2, \ldots, \boldsymbol{c}^k\}$. As illustrated in Figure S4 in Supplement, each set $\boldsymbol{c}^i$ can be considered as a meaningful random sample of the considered dependent inputs with regards to experimental observations. This physically meaningful method was used to generate results presented in Section 4. As a comparison, a different sampling method (Method 2) is presented hereafter.

Method 2: sample generation based on parameters decoupling

Based on a limited number of regressions *k*, this method aims at decoupling dependent regression parameters and generating independent marginal PDF. Conversely to Method 1, sets of regression coefficients can then be generated without performing any additional regression, reducing the computation time. In this context, a principal component analysis (PCA) is used to minimize the correlation between variables. However, zero correlation implies decoupling in the sole case where parameters follow a Gaussian distribution. This condition is met through a change of variables. Special care should be taken to avoid alteration of parameters dependency through the procedure. A flow chart of the method is provided in Figure 2.

First, *k* sets of *p* regression coefficients $\{\boldsymbol{c}^1, \boldsymbol{c}^2, \ldots, \boldsymbol{c}^k\}$ are generated according to the procedure related to Method 1. Then, parameter values are adapted to fit a beta distribution support ([0;1]), so that

$$c_{j,beta}^{i} = \frac{c_j^i - \min\left(c_j^1, c_j^2, ..., c_j^k\right)}{\max\left(c_j^1, c_j^2, ..., c_j^k\right) - \min\left(c_j^1, c_j^2, ..., c_j^k\right)}, \tag{21}$$

where c_j^i is a value of the j^{th} regression parameter observed in $\{\boldsymbol{c}^1, \boldsymbol{c}^2, \ldots, \boldsymbol{c}^k\}$.

A beta distribution is then fitted on the discretised marginal probability distribution of each regression parameter using maximum likelihood estimates of the beta distribution parameters α and β (Fisher, 1912). The fitted beta distributions are then used to map the regression parameters in beta CDF coordinates, so that

$$c_{j,betaCDF}^{i} = CDF_{beta}\left(c_{j,beta}^{i}, \left(\alpha;\beta\right)^j\right), \tag{22}$$

with $\left(\alpha;\beta\right)^j$ the fitted parameters of the beta distribution related to the j^{th} regression parameter.

The resulting regression parameters are then transposed in gaussian coordinates:

$$c_{j,gauss}^{i} = CDF_{gauss}^{-1}\left(c_{j,betaCDF}^{i}, \mu, \sigma^2\right), \tag{23}$$

with μ and σ^2 the mean and the variance of the gaussian distribution. For the sake of simplification, it was arbitrarily decided to use $\mu = 0$ and $\sigma^2 = 1$ in this study. Note that this choice has no implication on the results.

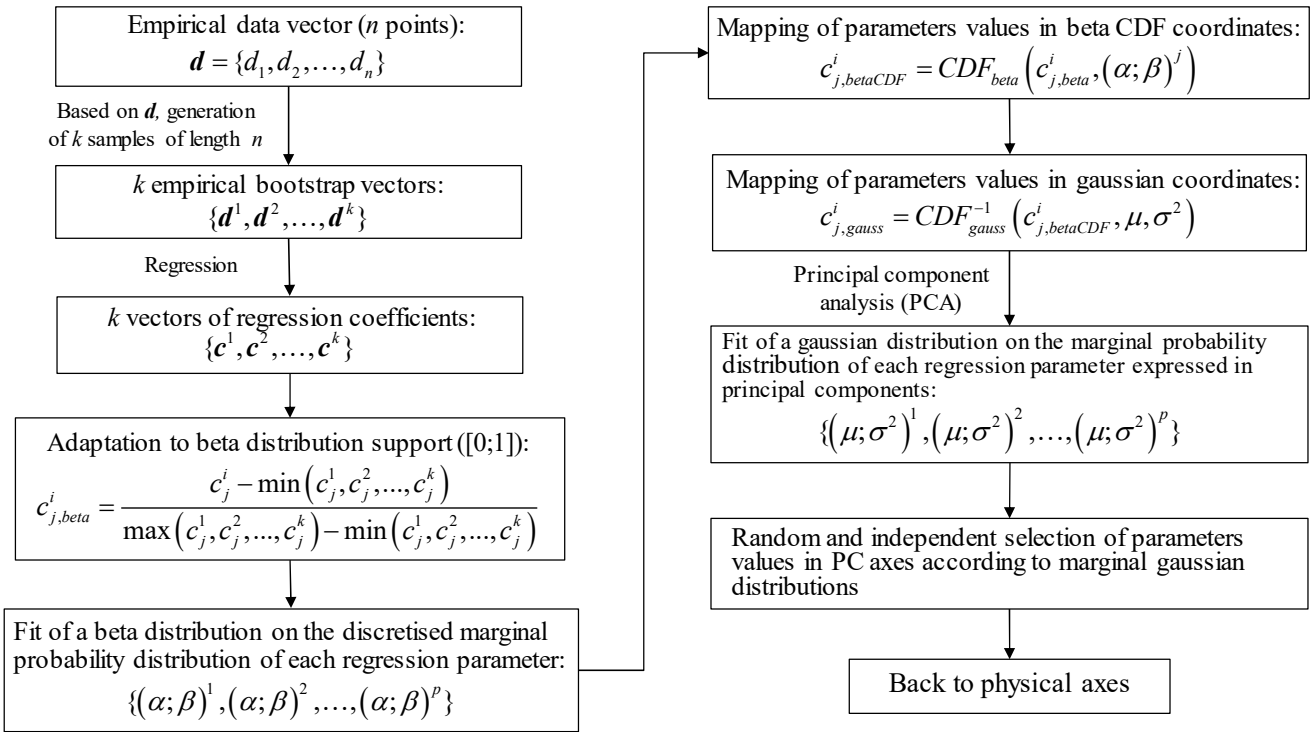


Figure 2. Flow chart of Method 2.

A principal component analysis (PCA) is then applied on the resulting data to minimize covariance. At this point, data approximately follow a Gaussian distribution. Then, gaussian distributions are fitted on the data expressed in the principal axes using maximum likelihood estimates of the gaussian distribution parameters μ and σ^2. Based on those PDF, parameter values are selected randomly and independently before being transposed in the initial physical axes to constitute the desired samples.

Although this method does not provide an explicit form of the multivariate probability distribution, it generates independent PDF for each regression coefficients. The computation time required for the sampling procedure is thus reduced. Nonetheless, those PDF are approximations of the multivariate probability distribution, and the accuracy of the results should thus be assessed systematically.

In this work, weir efficiency coefficients (c_1, c_2) are assumed dependent and varying proportionally. Their probability distributions are indirectly characterized through a linear mapping, so that

$$\begin{cases} c_1 = c_{eff} \cdot c_1^{ref} \\ c_2 = c_{eff} \cdot c_2^{ref} \end{cases} \quad \text{with } c_{eff} \text{ a random variable with } c_{eff}^{ref} = 1. \tag{24}$$

3.3. Global sensitivity analysis

The procedure of input uncertainty characterization was presented for independent and dependent input variables in Section 3.2. Those uncertainties were propagated to model outputs through our implementation of the numerical model of Wu (2016). Statistical descriptors are used to quantify the resulting uncertainties and to characterize the relation between uncertainties on input and output variables. In this work, Sobol indices of total order were used to highlight the most influential input uncertainties. It can be expressed as the ratio between significance descriptors and the global output variance.

Let the random variable $\boldsymbol{X} = \left(X_1,\ldots,X_{m_1}\right)$ be partitioned into n statistically independent random subsets of variables, so that $\boldsymbol{X}^1 = \left(X_1,\ldots,X_{m_1}\right), \boldsymbol{X}^2 = \left(X_{m_1+1},\ldots,X_{m_2}\right)$, … , $\boldsymbol{X}^n = \left(X_{m_1+\ldots+m_{n-1}+1},\ldots,X_{m_1+\ldots+m_n}\right)$. The significance descriptor related to subset j may be approximated as (Sobol, 2001; Arnst and Ponthot, 2014):

$$s_{\boldsymbol{X}^j} = E\left[V\left(g(\boldsymbol{X}^j,\boldsymbol{x}^{\sim j})\right)\right] \approx s^{\nu}_{\boldsymbol{X}^j} = \frac{1}{2\nu}\sum_{l=1}^{\nu}\left[g(\boldsymbol{x}_l^j,\boldsymbol{x}_l^{\sim j}) - g(\tilde{\boldsymbol{x}}_l^j,\boldsymbol{x}_l^{\sim j})\right]^2, \qquad (25)$$

where $\boldsymbol{X}^{\sim j}$ denotes the random variable that contains all the components of $\boldsymbol{X}$ that are not components of $\boldsymbol{X}^j$, $P_{\boldsymbol{X}^{\sim j}}$ and $P_{\boldsymbol{X}^j}$ the probability distributions of $\boldsymbol{X}^{\sim j}$ and $\boldsymbol{X}^j$, respectively, $\{\boldsymbol{x}_l^j, 1\le l\le\nu\}$ and $\{\tilde{\boldsymbol{x}}_l^j, 1\le l\le\nu\}$ two independent ensembles of independent and identically distributed (i.i.d.) samples from $P_{\boldsymbol{X}^j}$ and $\{\boldsymbol{x}_l^{\sim j}, 1\le l\le\nu\}$ an independent ensemble of i.i.d. samples from $P_{\boldsymbol{X}^{\sim j}}$, and g the model output function. As expressed Eq. (25), the significance descriptor associated to $\boldsymbol{X}^j$ represents the mean output variance obtained when the values of all input variables are fixed, except those related to input parameters contained in $\boldsymbol{X}^j$. Sobol indices of total order are finally obtained by dividing this indicator by the total output variance, i.e., the variance obtained when all input variables are unknown and vary according to their respective PDF.

In Eq. (25), the larger ν the better the approximation of the significance descriptor. Noting that $g(\boldsymbol{x}_l^j,\boldsymbol{x}_l^{\sim j})$ and $g(\tilde{\boldsymbol{x}}_l^j,\boldsymbol{x}_l^{\sim j})$ correspond to output values of two different runs, 2ν runs are required to evaluate Eq. (25).

4. Results and discussion

This section starts with the comparison of the two sampling methods of dependent inputs introduced in Section 3.2.2 and in Text S1 in Supplement. Using Method 1, a global sensitivity analysis is then applied to each configuration presented in Section 2.2. The focus is set on two model outputs, namely the peak breach discharge, $Q_{b,peak}$, and the time to peak, t_{peak}, i.e., time between the overtopping initiation and the peak breach discharge occurrence. All information used to characterize the independent input variables uncertainty is gathered in Table S4 in Supplement.

4.1. Comparison of dependent input samples generation methods

Before examining global sensitivity analysis results, we compare Method 1 and Method 2 used to generate samples of dependent input variables (Section 3.2.2). Figure S4 in Supplement shows regression curves obtained using random samples of $\left\{C_a^*, C_b^*, C_c^*, C_d^*\right\}$ and $\left\{q_{b,a}^*, q_{b,b}^*\right\}$ generated using Method 1. All regression curves seem in agreement with experimental point clouds, proving that Method 1 provides meaningful random samples. Conversely, Method 2 provides results that approximate the results obtained by Method 1, but computational cost is reduced. The objective is to determine which method presents the best trade-off between computational efficiency and accuracy. In this analysis, results are computed using Test 10 (reference configuration).

Sobol indices of total order were computed for the maximum breach discharge, $Q_{b,peak}$, and the related elapsed time since overtopping initiation, t_{peak}. Using Method 1, relative variations in Sobol indices are smaller than 1% when $\nu > 4000$ (Figure S5, in Supplement) and the results are thus considered as converged. Figure S6 in Supplement compares those values with results obtained by Method 2 using $k = 500$, $k = 1000$ and $k = 2000$. In all cases, results converged for $\nu = 4000$. Increasing the value of k does not improve accuracy on Sobol indices related to $Q_{b,peak}$. Regarding t_{peak}, a significant improvement is observed when k rises from 500 to 1000, with no additional accuracy gain when k reaches 2000. Therefore, $k = 1000$ was chosen. The maximum error on Sobol indices is about 5% compared to Method 1 for both model outputs.

The computing time required to generate regression samples using Method 1 ($k = \nu = 4000$) and Method 2 ($k = 1000$; $\nu = 4000$) is presented in Table 3. Method 2 is about four times faster than Method 1. Nonetheless, the time required to perform a numerical simulation of Test 10 is about 0.21 s. Since 4000 runs should be launched for each of the 21 parameters or parameters subsets, the total simulation time reaches about 18,500 s. The computation time saved on the regression coefficient samples generation is thus marginal in the present case. For this reason and to optimize results accuracy, Method 1 has been used to lead the global sensitivity in the following section.

Table 3. Comparison of computation time required to generate regression coefficient samples with Method 1 ($k = \nu = 4000$) and Method 2 ($k = 1000$; $\nu = 4000$).

	Method 1	Method 2			
	Regression procedure = Sample generation ($\nu = k = 4000$)	Regression procedure ($k = 1000$)	Sample generation ($\nu = 4000$)	Total time	Computation time reduction
$\{C_a^*, C_b^*, C_c^*, C_d^*\}$	411.9 [s]	103 [s]	0.6 [s]	103.6 [s]	**-75%** (-308.3 [s])
$\left\{q_{b,a}^*, q_{b,b}^*\right\}$	5.1 [s]	1.3 [s]	0.2 [s]	1.5 [s]	**-71%** (-3.6 [s])

4.2. Application of the global sensitivity analysis

Sobol indices of total order related to the peak breach discharge and the time to peak are presented in Figure 3a and 3b for laboratory-scale configurations and in Figure 4a and 4b for field-scale configurations. In those figures, the *y*-axis corresponds to the values of the Sobol indices while numbers along the *x*-axis refer to the tested configurations listed in Table 2. Circles aligned on a same vertical line refer to Sobol indices computed in a similar test configuration. One colour is associate to each input variable. Only parameters with a Sobol index larger than 5% for at least one configuration are presented. Vertical dotted lines separate groups of configurations in which the value of a single parameter is varied.

When considering laboratory-scale configurations, uncertainties in seven input parameters appear to be critical for both outputs of interest, namely c_{eff}, A_n, $A_{n'}$, n_{min}, θ_{cr}, λ and h_d. Additionally, uncertainties in the inflow discharge (Q_{in}) play a major role in the determination of $Q_{b,peak}$, while t_{peak} is also slightly impacted by uncertainties in q_b^* parameters, i.e., $q_{b,a}^*$ and $q_{b,b}^*$.

Results appear to be particularly different between configurations with no inflow discharge but an initial water level slightly higher than the initial breach depth (Tests 36 and 40 to 44) and tests with an inflow discharge (all other tests). In the latter case, uncertainties in input parameters affecting erosion, i.e., effective shear stress and sediment concentration are almost systematically the most critical for both outputs, as already pointed in many previous works (Sattar, 2014; Westoby et al., 2015; Alhasan et al., 2016; Peter et al., 2018; Kalinina et al., 2020). Those include A_n (parameter in Strickler formula), θ_{cr} (critical Shields parameter) and λ (empirical coefficient involved in mixing length computation). They dictate how fast the breach expands and are therefore key parameters in the breaching process modelling. Surprisingly, the influence of uncertainties in q_b^* and C_* parameters is mostly negligible for both outputs, i.e., influence of parameters related to bed load transport capacity and suspended load concentration, respectively. This can however be understood by considering their low uncertainty resulting from the bootstrapping method presented in Section 3.2.2 for input parameters with dependent PDF. A small uncertainty associated to parameter $A_{n'}$ also explains why it exhibits Sobol indices of negligible values.

When increasing the inflow discharge (Tests 18 to 21), Sobol indices related to bed load concentration (A_n and θ_{cr}) and the ones related to λ exhibit opposite trends for both outputs. As illustrated in Figure S7 in Supplement, suspended load corresponds to a larger fraction of the overall sediment concentration when the inflow discharge rises, to the expense of bed load. This induces a smaller relative impact of uncertainties in A_n and θ_{cr}, which leads to a higher relative influence of uncertainties in λ. Similar trends are observed when the initial breach width enlarges (Tests 20 and 24 to 27). The mixing length, L_s, involved in Eq. (10) is the product of λ and the water free surface width on the considered reach, B_{ini}. By increasing the initial breach width, B_{ini} increases as well and uncertainties in λ have more impact on L_s, which strongly

influences erosion dynamics. In this case, Sobol indices related to λ rise substantially to the expense of other parameters.

In Tests 27 (large initial breach width) and 29 (large dike crest width), the relative impact of uncertainties in Q_{in} becomes remarkable in the determination of $Q_{b,peak}$. In Test 27, the initial breach width is particularly large. In this case, additional breach expansion becomes less decisive for the maximum breach discharge value. This trend is illustrated in Figure S8 in Supplement, which shows a decreasing influence of uncertainties in input parameters when the initial breach width increases, except for the inflow discharge coefficient. In Test 29 (large dam crest width), the breaching initiation phase (i.e., phase of limited overtopping flow and slow erosion) is particularly long due to the extremely large dike crest that must be eroded. The larger the dike crest, the smoother the resulting breach hydrograph, as highlighted by Schmitz et al. (2021) for fluvial dikes. Therefore, the maximum breach discharge becomes less sensitive to variations in input parameters, as highlighted in Figure S8 in Supplement. Nonetheless, the relation between uncertainties in Q_b and Q_{in} is not altered, and the value of the inflow discharge coefficient becomes more important in the determination of $Q_{b,peak}$. Conversely, t_{peak} is extremely dependent on the duration of the breaching initiation phase, for which influence of erosion related parameters is predominant. For this reason, uncertainties in those parameters keep a dominant role for the determination of t_{peak}.

Although being negligible in most tested configurations, the impact of uncertainties in parameter n_{min} becomes more significant in Test 15 (very small median grain size) because the value of n_{min} is only considered in the computation of Manning's coefficients (Eq. (4) and (7)) when the median grain size, d_{50}, is particularly small. In this case, n_{min} highly impacts the erosion process by replacing the role of parameter A_n. In the opposite case, its value is not seen at all by the numerical model and its Sobol index becomes zero.

When no inflow discharge is injected (Tests 36 and 40 to 44), observations are quite different. When the reservoir area increases, the water level in the reservoir decreases more slowly, leading to a greater water level above the breach bottom, h_{top}. Through Eq. (2), the influence of c_{eff} on Q_b and $Q_{b,peak}$ grows substantially, to the expense of the relative influence of erosion-related parameters. The dam breaching process presents in this case a rather long initiation phase because it is only driven by a small difference between water level in the reservoir and initial breach depth. For this reason, the dam height, h_d, is critical to determine the duration of this initial phase and thus t_{peak}, so do erosion parameters.

Trends observed for field-scale configurations (Figure 4a, 4b, and S9 in Supplement) are extremely similar to the ones presented for laboratory-scale tests. Though, several differences deserve to be highlighted. Due to the increased value of the median grain size in field-scale configurations, parameter n_{min} plays no role anymore. Conversely, Sobol indices related to S_u and L_k slightly rise for $Q_{b,peak}$. This difference is, however, not significant since they hardly rise above the 5% limit value while being just below this threshold in laboratory-scale configurations. The same explanation holds for Sobol indices of q_b^* parameters related to t_{peak}. The small impact of scaling on sensitivity analysis results is confirmed by configurations gathered in the three scale families (Tests 8 and 10 to 17). Besides the punctual significant

impact of n_{min} in Test 15 (very small median grain size), results are poorly influenced by limited scale modifications in laboratory and field-scale cases. When analysing a real-scale dam breaching with the presented numerical model, this observation is of great interest since it suggests that a sensitivity analysis might be applied to a down-scaled embankment to reduce computational cost while leading to similar results.

While sensitivity analysis allows ranking parameters according to their relative impact on specific model outputs, uncertainty quantification assesses the magnitude of the output uncertainties. Both aspects should be combined to properly underline the most critical parameter uncertainties, i.e., parameter uncertainties with a critical impact on highly uncertain outputs. Based on the sensitivity analysis led in this study and uncertainty quantification results illustrated in Figures S10, S11 and S12 in Supplement, this combined approach was implemented in Figure 5 to generate a decision tree. This diagram aims at identifying which parameter uncertainties should be minimized in priority, according to the output uncertainty in the considered configuration. A_n turns to be the most critical parameter globally. Though, when the median grain size, d_{50}, drops below around 1 mm, the impact of n_{min} becomes predominant and tends to cancel the influence of A_n.

Overall, these results suggest that uncertainties in the estimation of the value of Manning's coefficient, n, considerably influence the model outputs. This aspect is confirmed in Figure S14 in Supplement, in which relative variations in the calculated bed load transport capacity are given as a function of predefined relative variations in the value of parameter n. It shows that a predefined relative variation in parameter n leads to a larger relative variation in the bed load transport capacity. This is due to the model structure and equations, which involve parameter n at a power above unity, such as in Eq. (5) and indirectly in Eq. (9). Hence a given degree of uncertainty in n leads to greater degree of uncertainty in the model outcomes. This issue is strongly intertwined with the assumption of uniform flow on the downstream face of the dam (Eq. (3)). Although widespread (e.g., Cai et al. (2022), Macchione (2008) and Shen et al. (2020)), this assumption remains an intrinsic limitation of simplified models, while a spatially distributed (e.g., in 2D) approach would be needed to properly capture the complex flow processes involved in dam breach hydraulics.

To assess the influence on the results of the assumed type of PDF, we have undertaken new computations in which the Beta distributions associated to the most influential inputs (c_{eff}, A_n, Q_{in}, $\theta_{\sigma'}$, λ and h_d, as identified in Figure 5) were replaced by truncated normal distributions. Each truncated normal distribution was defined using the same information as used to generate the corresponding Beta distribution (Section 3.2.1), i.e., the mode, the variance, and the variation interval. Figures S12 and S13 in Supplement show that this modification has no significant impact on the sensitivity analysis and uncertainty quantification results, i.e., the decision tree provided in Figure 5 is not altered.

An alternate way to test the sensitivity of the results to the input uncertainty characterization is to keep the same probability distribution function, but to vary its parametrization, e.g., its mean and variance. Bello et al. (2022) led this kind of analysis on a limited number of input variables to which a normal probability distribution function was assigned. They independently varied

the mean and variance of each input PDF by ± 50% and assessed the impact on the sensitivity analysis results. It appeared that the relative impact of the input uncertainties on the model output remained qualitatively the same, e.g., the most critical input uncertainty did not vary. However, the magnitude of the output uncertainties was considerably influenced. Unsurprisingly, it is expected that the larger the variations in the mean, variance, or variation interval extremities of the input probability distribution functions, the greater the impact on the analysis results. Though, an in-depth exploration of combinations of PDF parametrizations gets cumbersome when considering many inputs variables.

The results presented here highlight the need to conduct a global sensitivity analysis and an uncertainty quantification for each specific configuration of interest, especially when the main physical processes represented by the model vary from one configuration to the other (e.g., in the present case, when Q_{in} is equal to or different from zero). This conclusion applies also to numerous other applications in water resources management and environmental sciences, in which the relative weight of various physical processes represented in the model changes with the modelled configuration.

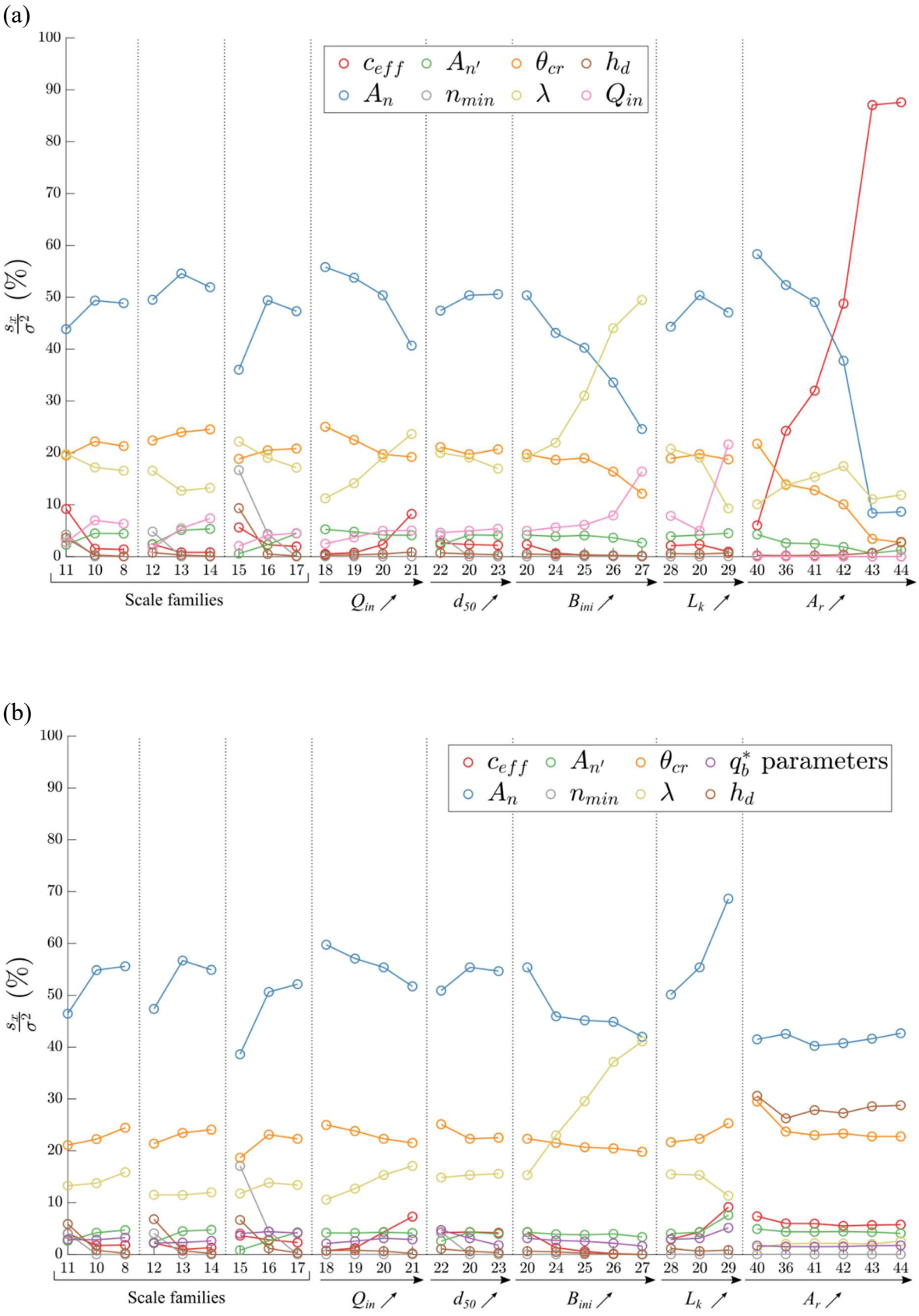


Figure 3. Sobol indices of total order for (a) peak breach discharge, and (b) time to peak in laboratory-scale configurations (numbers = Frank's (2016) test ID).

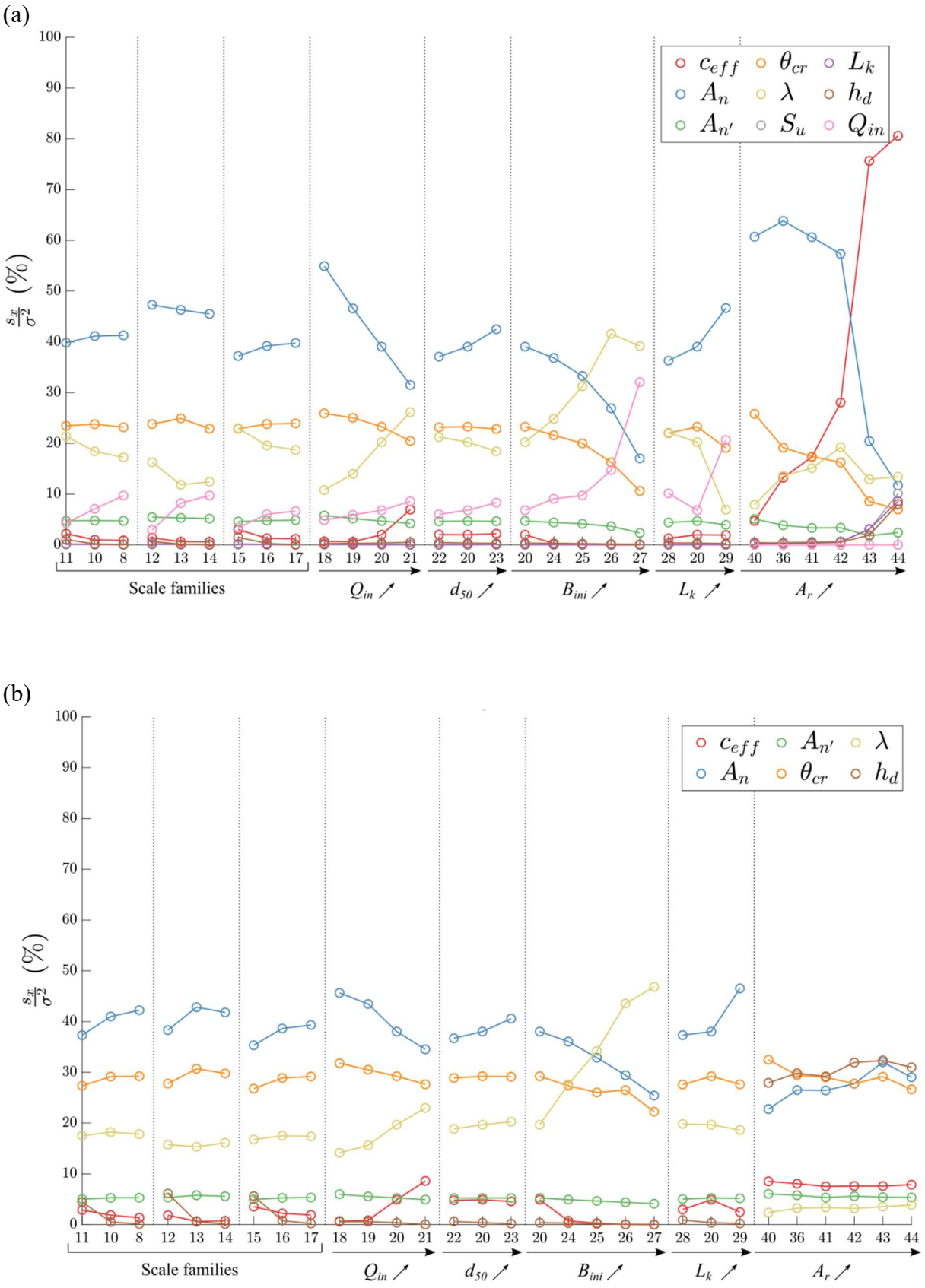


Figure 4. Sobol indices of total order for (a) peak breach discharge, and (b) time to peak in field-scale configurations (numbers = Frank's (2016) test ID).

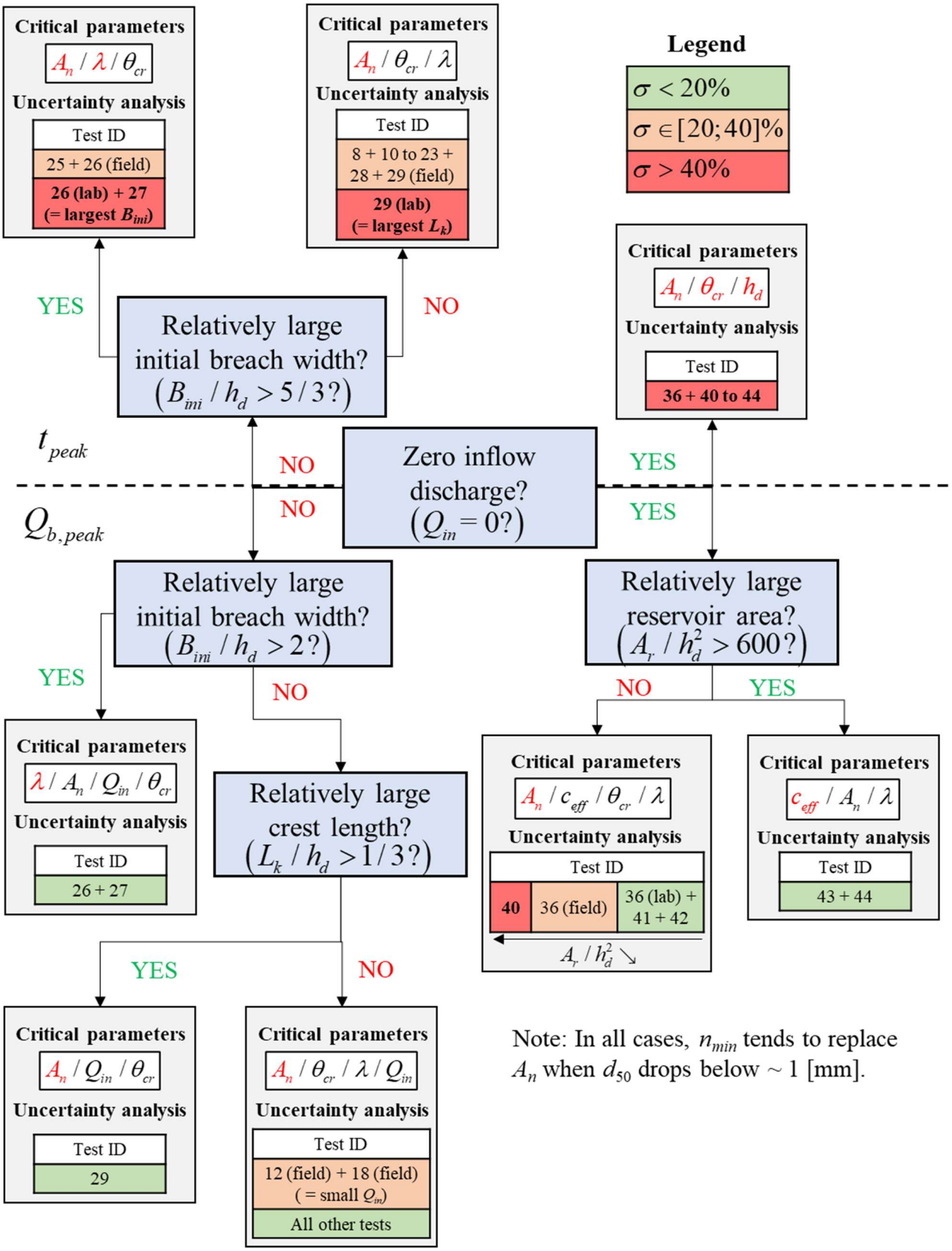


Figure 5. Decision tree indicating most influential parameters (leading parameters are in red) along with output uncertainties magnitude. σ refers to standard deviation values displayed in Figure S12 in Supplement.

5. Conclusion

This paper highlights the need for performing sensitivity analysis to identify most critical input variable uncertainties in embankment breaching numerical models. It also demonstrated the dependency between sensitivity/uncertainty quantification results and test configuration. Uncertainty propagation was performed using our implementation of the model of Wu (2016), in which 21 sets of input variables were deemed uncertain. The performance of two new sampling methods of dependent inputs were assessed, showing that the direct bootstrap sampling method (Method 1) exhibited the best accuracy, although being more computationally demanding.

Probability distributions were associated to independent input variables. Two new methods were proposed to generate samples of regression coefficients (i.e., dependent input variables) complying with field data. Based on available empirical data, Method 1 generated new data clouds by bootstrapping and computed associated regression coefficients for each of them. This method led to meaningful results but had poor computational performance because a regression is performed for each input set. Method 2 used variable changes and PCA to decouple input variables and allow defining independent marginal probability distributions. Although slightly less accurate, this method reduced dependent inputs sampling procedure computation time by more than 70% compared to Method 1. However, in the current study this gain was deemed negligible compared to the time required to run simulations using Monte Carlo procedure. Method 1 was finally used to lead the sensitivity analysis presented in this work.

Sobol indices of total order were computed for each input variable in twenty-seven configurations at both laboratory and field scales. It allowed assessing the relative impact of input uncertainties on the variability of two model outputs (peak breach discharge and time to reach the peak). Results were particularly different between configurations with no inflow discharge but an initial water level slightly higher than the initial breach invert level and tests with an inflow discharge. Conversely, the sensitivity of model predictions with respect to uncertainties in input parameters remains similar irrespective of the considered scale (laboratory vs. field scales), i.e., there was almost no difference between the Sobol indices obtained for laboratory-scale and for field-scale configurations. Parameter A_n turned to be the most critical parameter globally. Though, when the median grain size dropped below 1 mm, the impact of n_{min} became predominant and cancelled A_n influence. It comes out that uncertainties in Manning's coefficient value have a critical impact on model outputs variability. However, an accurate determination of this coefficient is not possible in practice, as it is a model parameter and not a measured quantity. This issue, together with the assumption of uniform flow on the dam downstream face, point at a limitation of simplified models, which may only be truly overcome by adopting a spatially distributed approach (e.g., 2D model) but at the expense of a considerably higher computational burden.

To further generalize the findings of this study, the influence of choosing a particular type of PDF (Beta distribution) to characterize input uncertainties was assessed. The calculations were repeated with an alternate type of distribution functions. The PDF associated to the parameter uncertainties influencing most the output variability were defined as truncated normal distributions instead of the initially assumed Beta distributions. It turned out that this

modification has no substantial impact on the analysis results. In contrast, extensive modifications in the PDF parametrization (e.g., strong modification of the variation interval or the variance) are expected to have more influence on the analysis results (Reed et al., 2022).

The present work allows identifying parameters whose uncertainty is critical for model outputs, depending on the configuration considered. The relevance of using highly uncertain model parameters which strongly influence the results may be questioned. Also, the reliability of the numerical model may be partly assessed by looking at the magnitude of the output uncertainties related to a given configuration. Finally, the procedure described in this paper may be used to determine whether a modification in the model structure brings a notable improvement in results accuracy but also a reduction of the output uncertainty ranges. The global sensitivity analysis procedure and conclusions presented in this study are not limited to the sole case of dam breaching but are also of relevance to a wide variety of modelling applications in environmental sciences. The global sensitivity analysis procedure presented in Section 3 is easily transferrable to other non-linear computational models containing independent and dependent uncertain scalar input variables. Its use is particularly relevant when the objective is to understand which input uncertainties are the most critical (i.e., factor prioritization) and which ones are negligible (i.e., factor fixing). This information allows simplifying subsequent analyses of the numerical model in the considered configuration as some uncertainties may be discarded.

To further generalize our findings and increase their robustness, similar studies should be performed using different numerical models of different types, additional test configurations and different sensitivity indicators, e.g., delta indicator (Borgonovo, 2007). These observations would further support decision-making processes in risk management (Wagener and Pianosi, 2019). When dealing with non-scalar sources of uncertainties, i.e., spatially distributed or time-varying uncertain inputs, other more suitable techniques exist to generate relevant input samples (e.g., Baroni and Tarantola (2014) and Lilburne and Tarantola (2009)). With computational demanding models, the Monte Carlo procedure used in this work should be replaced by an optimized sampling method (e.g., Latin Hypercube Sampling). In this case, Method 2 used to generate samples of dependent inputs might lead to significant computational time saving overall.

Data availability statement

Data were not used, nor created for this research.

References

Abdedou, A., Soulaïmani, A., Tchamen, G.W., 2020. Uncertainty propagation of dam break flow using the stochastic non-intrusive B-splines Bézier elements-based method. Journal of Hydrology 590.

Ahmadisharaf, E., Kalyanapu, A.J., Thames, B.A., Lillywhite, J., 2016. A probabilistic framework for comparison of dam breach parameters and outflow hydrograph generated by different empirical prediction methods. Environmental Modelling and Software 86, 248–263.

Alhasan, Z., Duchan, D., Říha, J., 2016. The probabilistic solution of dike breaching due to overtopping. Proceedings of the 4th European Congress of the International Association of Hydroenvironment engineering and Research, IAHR 2016, 505–512.

Arnst, M., Ponthot, J., 2014. An overview of nonintrusive characterization, propagation, and sensitivity analysis of uncertainties in computational mechanics. International Journal for Uncertainty Quantification 4, 387–421.

ASCE, 2011. Earthen Embankment Breaching. Journal of Hydraulic Engineering 137, 1549–1564.

Baroni, G., Tarantola, S., 2014. A General Probabilistic Framework for uncertainty and global sensitivity analysis of deterministic models: A hydrological case study. Environmental Modelling & Software 51, 26–34.

Bello, D., Alcayaga, H., Caamaño, D., Pizarro, A., 2022. Influence of Dam Breach Parameter Statistical Definition on Resulting Rupture Maximum Discharge. Water (Switzerland) 14.

Bellos, V., Tsakiris, V.K., Kopsiaftis, G., Tsakiris, G., 2020. Propagating dam breach parametric uncertainty in a river reach using the HEC-RAS software. Hydrology 7, 1–14.

Benke, K.K., Lowell, K.E., Hamilton, A.J., 2008. Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model. Mathematical and Computer Modelling 47, 1134–1149.

Borgonovo, E., 2006. Measuring Uncertainty Importance: Investigation and Comparison of Alternative Approaches. Risk Analysis 26, 1349–1361.

Borgonovo, E., 2007. A new uncertainty importance measure. Reliability Engineering & System Safety 92, 771–784.

Borgonovo, E., Plischke, E., 2016. Sensitivity analysis: A review of recent advances. European Journal of Operational Research 248, 869–887.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Cai, Y., Zhang, X., Xue, R., Wang, M., Deng, Q., 2022. Numerical simulation of overtopping breach processes caused by failure of landslide dams. Environmental Fluid Mechanics 22, 839–863.

Cantero-Chinchilla, F.N., Castro-Orgaz, O., Dey, S., 2019. Prediction of Overtopping Dike Failure: Sediment Transport and Dynamic Granular Bed Deformation Model. Journal of Hydraulic Engineering 145.

Chen, Z.Y., Ping, Z.Y., Wang, N.X., Yu, S., Chen, S.J., 2019. An approach to quick and easy evaluation of the dam breach flood. Science China Technological Sciences 62, 1773–1782.

Da Veiga, S., Gamboa, F., Iooss, B., Prieur, C., 2021. Basics and trends in sensitivity analysis : theory and practice in R, Computational science & engineering ; CS23. Society for Industrial and Applied Mathematics.

Da Veiga, S., Wahl, F., Gamboa, F., 2009. Local polynomial estimation for sensitivity analysis on models with correlated inputs. Technometrics 51, 452–463.

Damgaard, J.S., Whitehouse, R.J., Soulsby, R.L., 1997. Bed-load sediment transport on steep longitudinal slopes. Journal of Hydraulic Engineering 123, 1130–1138.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap methods and their application. Cambridge university press.

De Lorenzo, G., Macchione, F., 2014. Formulas for the peak discharge from breached earthfill dams. Journal of Hydraulic Engineering 140, 56–67.

Fisher, R.A., 1912. On an absolute criterion for fitting frequency curves. Messenger of mathematics 41, 155–156.

Frank, P.-J., 2016. Hydraulics of spatial dike breaches. Ph.D. thesis, ETH Zurich.

Froehlich, D.C., 2008. Embankment dam breach parameters and their uncertainties. Journal of Hydraulic Engineering 134, 1708–1721.

Froehlich, D.C., Goodell, C.R., 2012. Breach of duty (not): Evaluating the uncertainty of dam-breach flood predictions. Proceedings of World Environmental and Water Resources Congress 2012: Crossing Boundaries.

Goeury, C., Bacchi, V., Zaoui, F., Bacchi, S., Pavan, S., Abderrezzak, K. El kadi, 2022. Uncertainty Assessment of Flood Hazard Due to Levee Breaching. Water 14, 3815.

Gupta, H.V., Clark, M.P., Vrugt, J.A., Abramowitz, G., Ye, M., 2012. Towards a comprehensive assessment of model structural adequacy. Water Resources Research 48.

Hall, J.W., Boyce, S.A., Wang, Y., Dawson, R.J., Tarantola, S., Saltelli, A., 2009. Sensitivity analysis for hydraulic models. Journal of Hydraulic Engineering 135, 959–969.

Hall, J.W., Tarantola, S., Bates, P.D., Horritt, M.S., 2005. Distributed Sensitivity Analysis of Flood Inundation Model Calibration. Journal of Hydraulic Engineering 131, 117–126.

Helton, J.C., Davis, F.J., 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliability Engineering and System Safety 81, 23–69.

Helton, J.C., Johnson, J.D., Sallaberry, C.J., Storlie, C.B., 2006. Survey of sampling-based methods for uncertainty and sensitivity analysis. Reliability Engineering and System Safety 91, 1175–1209.

Hou, T., Nuyens, D., Roels, S., Janssen, H., 2019. Quasi-Monte Carlo based uncertainty analysis: Sampling efficiency and error estimation in engineering applications. Reliability Engineering and System Safety 191.

Iooss, B., Lemaître, P., 2015. A review on global sensitivity analysis methods, Operations Research/ Computer Science Interfaces Series.

Jacques, J., Lavergne, C., Devictor, N., 2006. Sensitivity analysis in presence of model uncertainty and correlated inputs. Reliability Engineering & System Safety 91, 1126–1134.

Janssen, H., 2013. Monte-Carlo based uncertainty analysis: Sampling efficiency and sampling convergence. Reliability Engineering and System Safety 109, 123–132.

Jin, R., Chen, W., Simpson, T.W., 2001. Comparative studies of metamodelling techniques under multiple modelling criteria. Structural and Multidisciplinary Optimization 23, 1–13.

Kalinina, A., Spada, M., Vetsch, D.F., Marelli, S., Whealton, C., Burgherr, P., Sudret, B., 2020. Metamodeling for uncertainty quantification of a flood wave model for concrete dam breaks. Energies 13.

Kiureghian, A.D., Ditlevsen, O., 2009. Aleatory or epistemic? Does it matter? Structural Safety 31, 105–112.

Lee, K., 2019. Simulation of Dam-Breach Outflow Hydrographs Using Water Level Variations. Water Resources Management 33, 3781–3797.

Li, G., Rabitz, H., Yelvington, P.E., Oluwole, O.O., Bacon, F., Kolb, C.E., Schoendorf, J., 2010. Global sensitivity analysis for systems with independent and/or correlated inputs. Journal of Physical Chemistry A 114, 6022–6032.

Li, Y., Chen, A., Wen, L., Bu, P., Li, K., 2020. Numerical simulation of non-cohesive homogeneous dam breaching due to overtopping considering the seepage effect. European Journal of Environmental and Civil Engineering.

Lilburne, L., Tarantola, S., 2009. Sensitivity analysis of spatial models. International Journal of Geographical Information Science 23, 151–168.

Macchione, F., 2008. Model for Predicting Floods due to Earthen Dam Breaching. I: Formulation and Evaluation. Journal of Hydraulic Engineering 134, 1688–1696.

McKay, M.D., Beckman, R.J., Conover, W.J., 1979. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics 21, 239–245.

Metropolis, N., Ulam, S., 1949. The Monte Carlo Method. Journal of the American Statistical Association 44, 335–341.

Mokhtari, A., Frey, H.C., 2005. Sensitivity Analysis of a Two-Dimensional Probabilistic Risk Assessment Model Using Analysis of Variance. Risk Analysis 25, 1511–1529.

Morris, M.D., 1991. Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics 33, 161–174.

Onda, S., Hosoda, T., Jaćimović, N.M., Kimura, I., 2019. Numerical modelling of simultaneous overtopping and seepage flows with application to dike breaching. Journal of Hydraulic Research 57, 13–25.

Pappenberger, F., Beven, K.J., 2006. Ignorance is bliss: Or seven reasons not to use uncertainty analysis. Water Resources Research 42.

Peter, S.J., Siviglia, A., Nagel, J., Marelli, S., Boes, R.M., Vetsch, D., Sudret, B., 2018. Development of Probabilistic Dam Breach Model Using Bayesian Inference. Water Resources Research 54, 4376–4400.

Pheulpin, L., Bacchi, V., Bertrand, N., 2020. Comparison Between Two Hydraulic Models (1D and 2D) of the Garonne River: Application to Uncertainty Propagations and Sensitivity Analyses of Levee Breach Parameters, Springer Water.

Pheulpin, L., Bertrand, N., Bacchi, V., 2022. Uncertainty quantification and global sensitivity analysis with dependent inputs parameters: Application to a basic 2D-hydraulic model. LHB 108, 2015265.

Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T., 2016. Sensitivity analysis of environmental models: A systematic review with practical workflow. Environmental Modelling & Software 79, 214–232.

Pianosi, F., Wagener, T., 2015. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. Environmental Modelling & Software 67, 1–11.

Queipo, N.V., Haftka, R.T., Shyy, W., Goel, T., Vaidyanathan, R., Kevin Tucker, P., 2005. Surrogate-based analysis and optimization. Progress in Aerospace Sciences 41, 1–28.

Rakovec, O., Hill, M.C., Clark, M.P., Weerts, A.H., Teuling, A.J., Uijlenhoet, R., 2014. Distributed Evaluation of Local Sensitivity Analysis (DELSA), with application to hydrologic models. Water Resources Research 50, 409–426.

Reed, P.M., Hadjimichael, A., Malek, K., Karimi, T., Vernon, C.R., Srikrishnan, V., Gupta, R.S., Gold, D.F., Lee, B., Keller, K., Thurber, T.B., Rice, J.S., 2022. Addressing Uncertainty in Multisector Dynamics Research. Zenodo.

Rosenblueth, E., 1975. Point estimates for probability moments. Proceedings of the National Academy of Sciences of the United States of America 72, 3812–3814.

Rossi, R.J., 2018. Mathematical statistics: an introduction to likelihood based inference. John Wiley & Sons.

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., Wu, Q., 2019. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. Environmental Modelling & Software 114, 29–39.

Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. Environmental Modelling & Software 25, 1508–1517.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Computer Physics Communications 181, 259–270.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity Analysis: The Primer. Wiley, Chichester (England).

Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., others, 2004. Sensitivity analysis in practice: a guide to assessing scientific models. Wiley Online Library.

Sattar, A.M.A., 2014. Gene expression models for prediction of dam breach parameters. Journal of Hydroinformatics 16, 550–571.

Schmitz, V., Erpicum, S., El Kadi Abderrezzak, K., Rifai, I., Archambeau, P., Pirotton, M., Dewals, B., 2021. Overtopping-Induced Failure of Non–Cohesive Homogeneous Fluvial Dikes: Effect of Dike Geometry on Breach Discharge and Widening. Water Resources Research 57, e2021WR029660.

Shen, G., Sheng, J., Xiang, Y., Zhong, Q., Yang, D., 2020. Numerical Modeling of Overtopping-Induced Breach of Landslide Dams. Natural Hazards Review 21.

Shimizu, Y., Nelson, J., Arnez Ferrel, K., Asahi, K., Giri, S., Inoue, T., Iwasaki, T., Jang, C., Kang, T., Kimura, I., Kyuka, T., Mishra, J., Nabi, M., Patsinghasanee, S., Yamaguchi, S., 2020. Advances in computational morphodynamics using the International River Interface Cooperative (iRIC) software. Earth Surface Processes and Landforms 45, 11–37.

Sobol, I., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and computers in simulation 55, 271–280.

Sobol, I.M., 1993. Sensitivity estimates for nonlinear mathematical models. Mathematical modeling and computational experiments 4, 407–414.

Spear, R.C., Hornberger, G.M., 1980. Eutrophication in peel inlet—II. Identification of critical uncertainties via generalized sensitivity analysis. Water Research 14, 43–49.

Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. Reliability Engineering & System Safety 93, 964–979.

Tsai, C.W., Yeh, J., Huang, C., 2019. Development of probabilistic inundation mapping for dam failure induced floods. Stochastic Environmental Research and Risk Assessment 33, 91–110.

Vorogushyn, S., Apel, H., Merz, B., 2011. The impact of the uncertainty of dike breach development time on flood hazard. Physics and Chemistry of the Earth 36, 319–323.

Wagener, T., Pianosi, F., 2019. What has Global Sensitivity Analysis ever done for us? A systematic review to support scientific advancement and to inform policy-making in earth system modelling. Earth-Science Reviews 194, 1–18.

Walker, W.E., Harremoës, P., Rotmans, J., Sluijs, J.P. van der, Asselt, M.B.A. van, Janssen, P., Krauss, M.P. Krayer von, 2003. Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. Integrated Assessment 4, 5–17.

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M., Hassan, M.A.A.M., Lowe, A., 2015. Numerical modelling of glacial lake outburst floods using physically based dam-breach models. Earth Surface Dynamics 3, 171–199.

Wu, W., 2013. Simplified physically based model of earthen embankment breaching. Journal of Hydraulic Engineering 139, 837–851.

Wu, W., 2016. Introduction to DL Breach–A simplified physically based dam/levee breach model. Clarkson University, NY.

Wu, W., Wang, S.S., 2006. Formulas for sediment porosity and settling velocity. Journal of Hydraulic Engineering 132, 858–862.

Wu, W., Wang, S.S., Jia, Y., 2000. Nonuniform sediment transport in alluvial rivers. Journal of Hydraulic Research 38, 427–434.

Zhang, R.J., 1961. River dynamics. Beijing: Industry Press.

Zhong, Q., Chen, S., Deng, Z., 2017. Numerical model for homogeneous cohesive dam breaching due to overtopping failure. Journal of Mountain Science 14, 571–580.

Zhong, Q., Wang, L., Chen, S., Chen, Z., Shan, Y., Zhang, Q., Ren, Q., Mei, S., Jiang, J., Hu, L., Liu, J., 2021. Breaches of embankment and landslide dams - State of the art review. Earth-Science Reviews 216.

Supplement to "Global Sensitivity Analysis of a Dam Breaching Model: to Which Extent is Parameter Sensitivity Case-Dependent?"

Flow chart of the numerical model

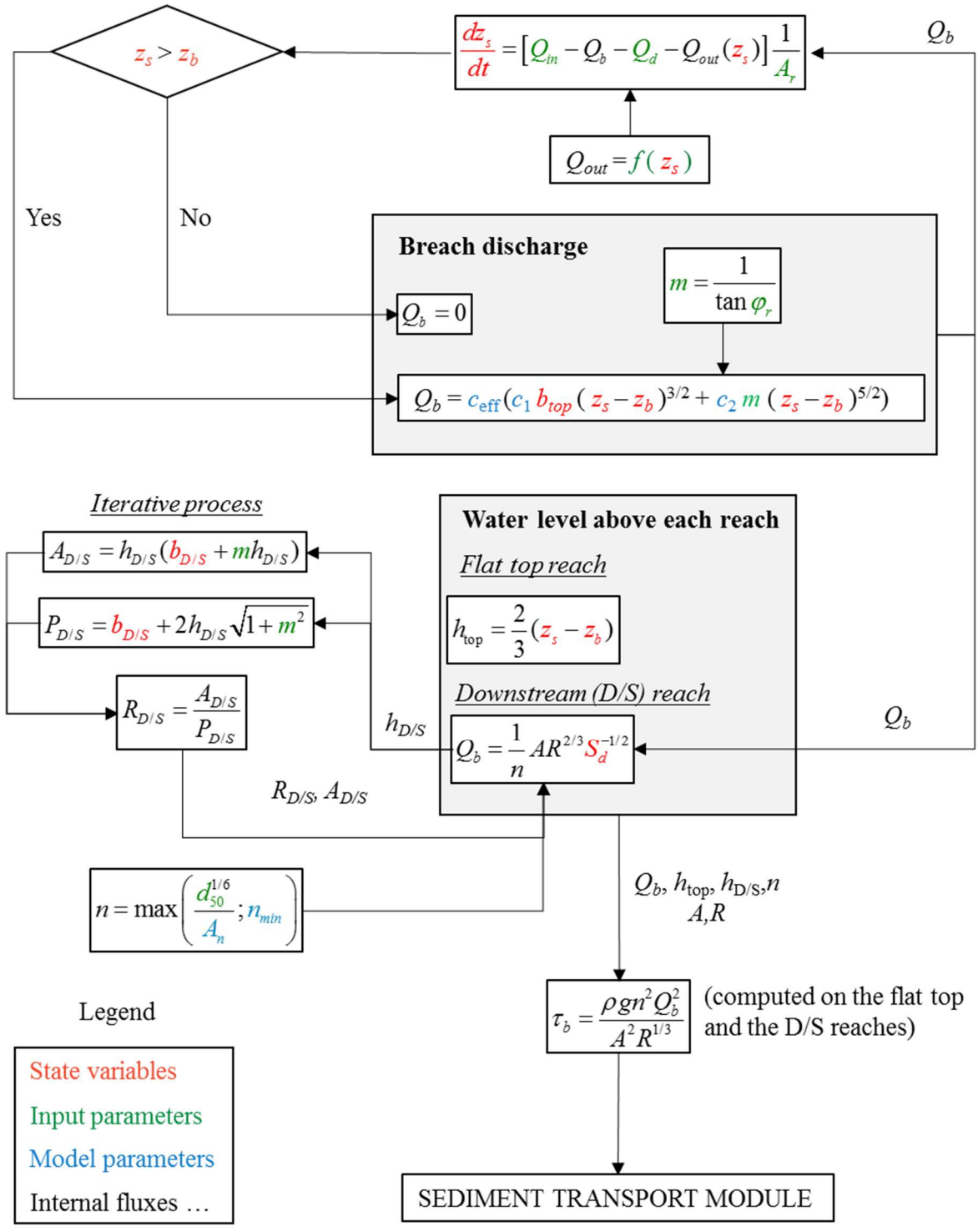


Figure S1. Flow chart of the hydrodynamic module of the implemented numerical model.

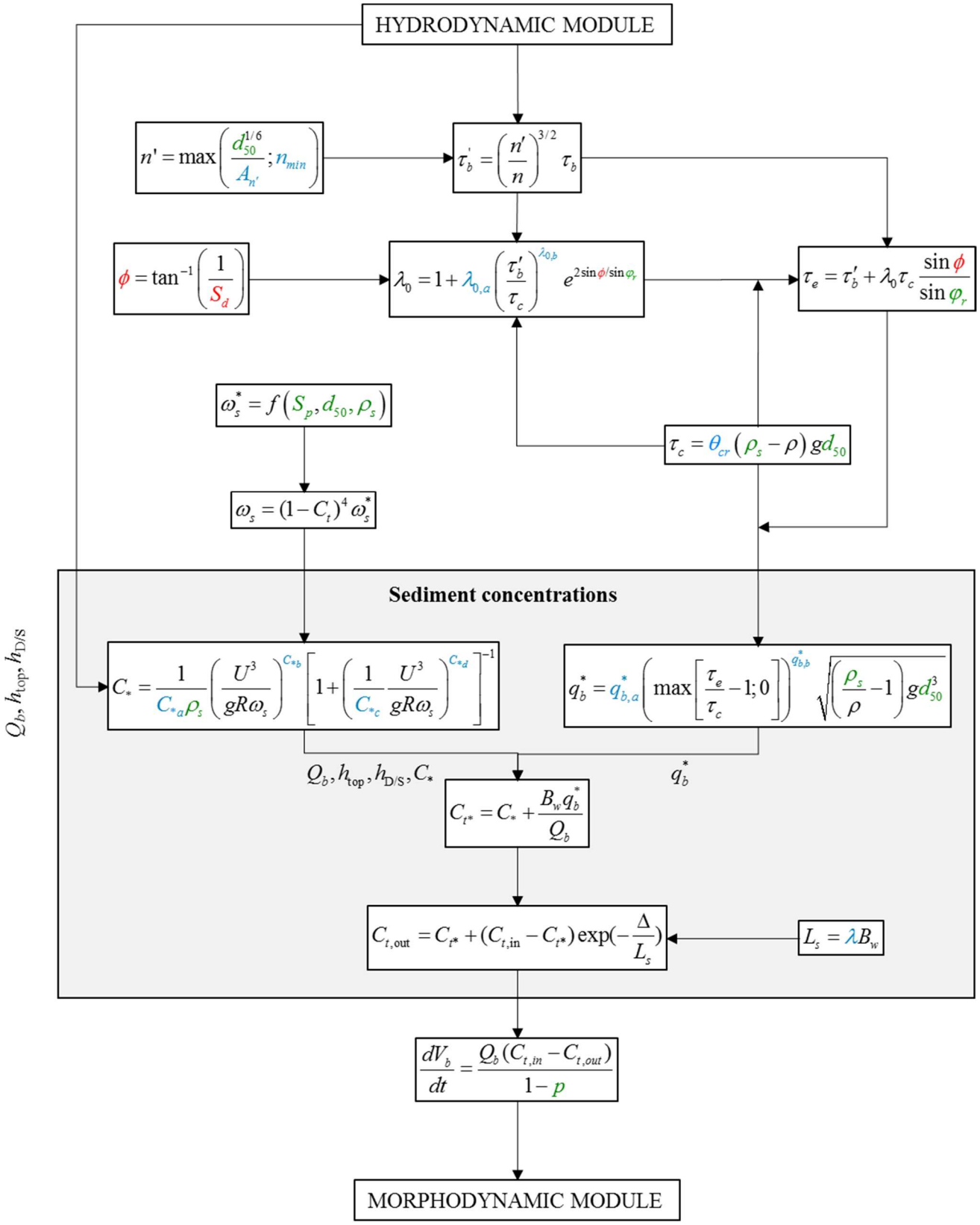


Figure S2. Flow chart of the sediment transport module.

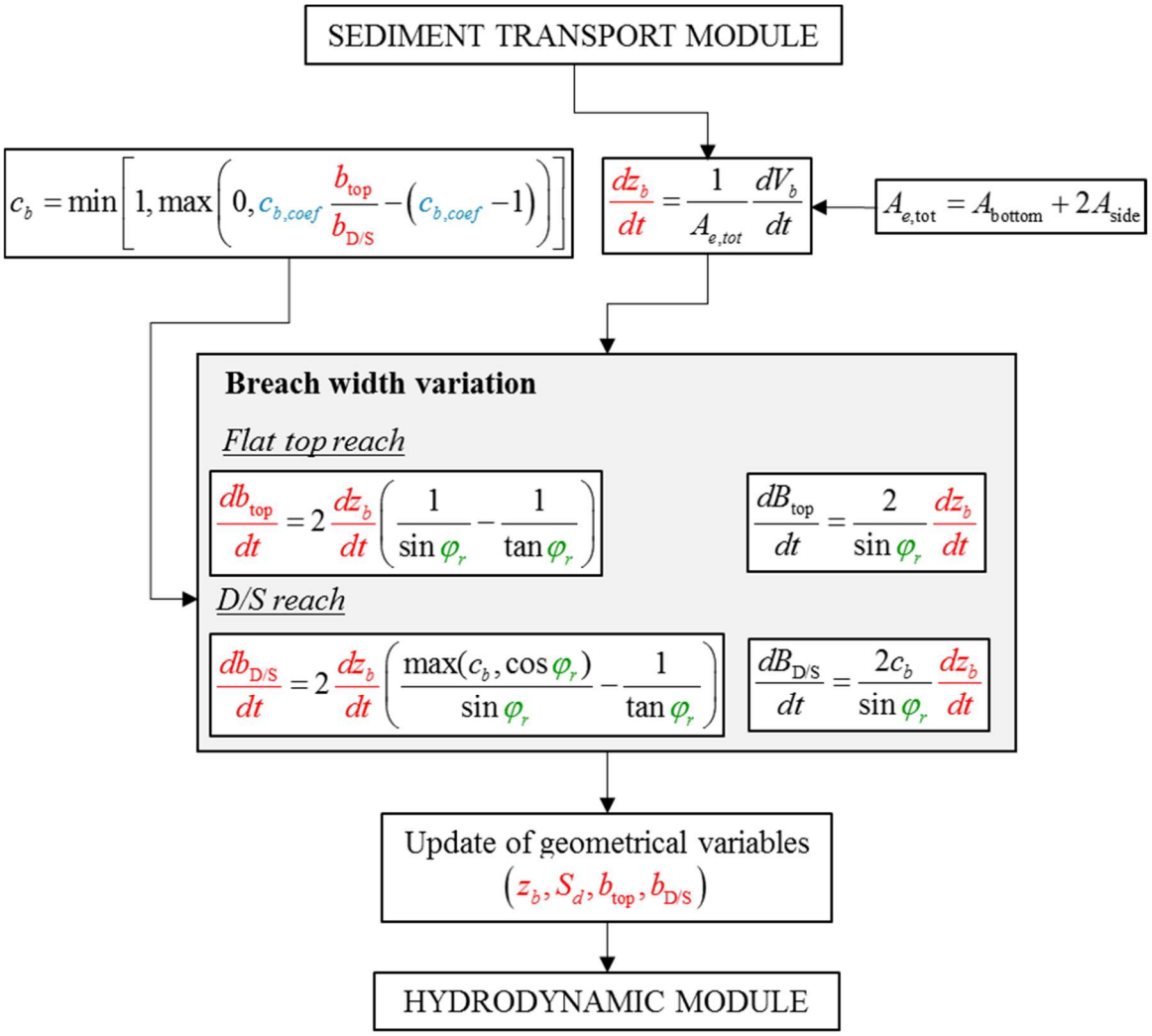


Figure S3. Flow chart of the morphodynamic module.

Table S1. Literature review (A).

Reference	Numerical model	Uncertainty analysis				
		Case study	Inputs type	PDF characterization	Sampling	Method/Indicator
Tsai et al. (2019)	Simplified physically based (0D)	1 field-scale dam	Independent (3 inputs)	Modeler expertise	PEM (7 runs)	Variance-based method Statistical moments of the outputs
Abdedou et al. (2020)	Detailed hydraulic model (1D) No dam model	1 field-scale dam	Independent (3 inputs)	Modeler expertise	Monte-Carlo (10^6) + 2 meta-models (10+100, each)	Confidence interval Outputs PDF
Froehlich and Goodell (2012)	1D hydraulic model 0D empirical breach model	1 field-scale dam	Independent (4 inputs)	Modeler expertise	PEM (16 runs)	Statistical moments of the outputs
Vorogushyn et al. (2011)	1D hydraulic model 0D probabilistic dike breach model 2D storage cell inundation model	1 field-scale fluvial dike case along a river (breaching possible at different places, simultaneously)	Dependent and independent (5 inputs)	MLE	Monte-Carlo (1000 runs)	Mean and percentiles Uncertainty bands
Kalinina et al. (2020)	1D hydraulic model No dam model	1 field-scale dam	Dependent (9 inputs)	MLE	Latin Hypercube (2000 runs of initial model to build meta-model + 10^6 runs of intermediate meta-model)	Sobol indices Outputs PDF
Pheulpin et al. (2020)	1D + 2D hydraulic models Fixed dam geometry	1 field-scale fluvial dike	Independent (3 inputs)	Modeler expertise	Monte-Carlo (200 runs of initial model to build each meta-model (2) + 5000 runs of each meta-model)	Sobol indices Outputs PDF
Goeury et al. (2022)	2D hydrodynamic model 0D breach model (Prescribed evolution)	1 field-scale river with multiple fluvial dikes	Independent (~300 inputs)	Modeler expertise	Monte-Carlo	PFI Delta indicator Outputs PDF

Notations and abbreviations: PEM = Point estimate method; MLE = Maximum likelihood estimation; PFI = Permutation feature importance.

Table S2. Literature review (B).

Reference	Numerical model	Uncertainty analysis				
		Case study	Inputs type	PDF characterization	Sampling	Method/Indicator
Froehlich (2008)	0D empirical breach model	1 field-scale dam	Independent (3 inputs)	MLE	Monte-Carlo (10^5 runs)	Mean and percentiles
Sattar (2014)	0D empirical breach model	General (field-scale dams)	Independent (6 inputs)	MLE	Monte-Carlo (250,000 runs)	MAD Percentiles Least square linearization technique
Ahmadisharaf et al. (2016)	4 different 0D empirical breach models	1 field-scale dam	Independent (3 inputs)	Modeler expertise	Latin Hypercube (10^4 for each model)	MAD Mean and percentiles Importance measure Coefficient of determination
Westoby et al. (2015)	2D hydrodynamic model 0D physically based breach model	1 field-scale dam	Independent (8 inputs)	Modeler expertise	Monte-Carlo (10^3 runs)	Uncertainty bands
Bellos et al. (2020)	1D hydrodynamic model 0D breach model (Prescribed evolution)	1 field-scale dam	Independent (7 inputs)	Modeler expertise	400 runs for MSM + Monte-Carlo (10^4 runs)	MSM Uncertainty bands Outputs PDF
Alhasan et al. (2016)	0D simplified physically based model	1 field-scale fluvial dike	Independent (9 inputs)	Modeler expertise	Latin Hypercube (5.10^6 runs)	MSM Uncertainty bands Outputs PDF
Peter et al. (2018)	Simplified physics-based dam breach model (0D)	1 field-scale dam	Independent (3 inputs)	Bayesian inference	Latin Hypercube (5.10^3)	Outputs PDF
Present study	Physically based (0D)	27 dams (lab + field scales)	Independent and dependent (21 inputs)	Modeler expertise	Monte-Carlo (4.10^3 / test)	Sobol indices Mean and percentiles

Notations and abbreviations: MSM = Morris Screening Method; MAD = Mean Absolute Deviation; B_{avg} = breach average width; MLE = Maximum Likelihood Estimation.

Table S3. Frank's (2016) experimental parameters. L_r = reservoir length; l_r = reservoir width; A_r = reservoir area; h_d = dam height; L_k = dam crest length; $z_{b,ini}$ = initial notch depth; B_{ini} = initial notch width; d_{50} = median grain size; Q_{in} = inflow discharge; h_{cr} = critical flow depth.

	Test ID	L_r (m)	l_r (m)	A_r (m^2)	h_d (mm)	L_k (mm)	$z_{b,ini}$ (mm)	B_{ini} (mm)	d_{50} (mm)	Q_{in} (l/s)	h_{cr} (mm)
Scaling	11	0.61	0.5		150	50	50	200	0.86	2.62	14.1
	10	1.22	1		300	100	100	400	1.75	14.8	28.2
	8	2.44	2		600	200	200	800	3.78	83.6	56.3
	12	0.61	0.5		150	50	50	200	0.86	1.31	8.9
	13	1.22	1		300	100	100	400	1.75	7.4	17.7
	14	2.44	2		600	200	200	800	3.78	41.8	35.4
	15	0.61	0.5		150	50	50	200	0.43	2.62	14.1
	16	1.22	1		300	100	100	400	0.86	14.8	28.2
	17	2.44	2		600	200	200	800	1.75	83.6	56.3
Discharge	18	3.44	2		300	100	100	400	1.75	4.4	7.9
	19									8.8	12.5
	20									18.4	20.5
	21									35.2	31.6
Sediment	22	3.44	2		300	100	100	400	0.86	18.4	20.5
	20								1.75		
	23								3.78		
Initial breach width	20	3.44	2		300	100	100	400	1.75	18.4	20.5
	24							500			
	25							600			
	26							800			
	27							1200			
Crest length	28	3.44	2		300	0	100	400	1.75	18.4	20.5
	20					100					
	29					400					
Reservoir water surface area	40	3.44	2	6.88	300	100	20	80	1.75	0	0
	36			26.88							
	41			33.02							
	42			46.88							
	43			126.88							
	44			206.8							

Table S4. Characteristics of the independent variables considered in the present study. X_{ref} stands for the reference value of the related parameter.

	Symbol	Reference value	Variation range	α
Model parameters	c_{eff}	1	$\left[\frac{1.5}{1.7};\frac{2.2}{1.7}\right]$	2
	A_n	12 (field) 16 (lab)	[10; 20]	2
	A_{ll}	20	[18; 22]	5
	n_{min}	0.016	[0.01; 0.017]	3
	θ_{cr}	0.03	[0.025; 0.06]	3
	$\lambda_{0,a}$ $\lambda_{0,b}$	0.2 0.15	[0.15;0.25] [0.1;0.2]	3
	S_p	0.7	[0.2; 1]	4
	λ	3	[0.1; 4]	2
	$c_{b,coef}$	1.8	[1; 2.5]	2
Input parameters	S_u / S_d	2	[0.9; 1.1] X_{ref}	6
	L_k [m]	Test dependent	[-0.01; +0.01] + X_{ref}	6
	h_d [m]	Test dependent	[-0.01; +0.01] + X_{ref}	4
	φ_r	39.5	[38°; 41°]	4
	ρ_s [kg/m^3]	2600	[2400; 2800]	6
	d_{50} [mm]	Test dependent	[0.9; 1.1] X_{ref}	6
	p	0.44	[0.41; 0.47]	4
	Q_{in} [l/s]	Test dependent	[0.9; 1.1] X_{ref}	4

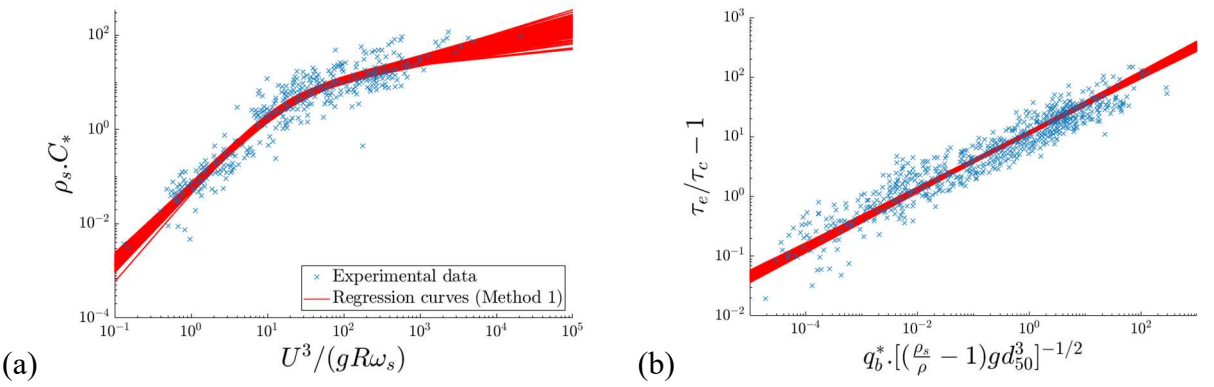


Figure S4. Comparison between experimental data and regression curves based on (a) $\left\{C_a^*, C_b^*, C_c^*, C_d^*\right\}$ and (b) $\left\{q_{b,a}^*, q_{b,b}^*\right\}$ subsets generated using Method 1.

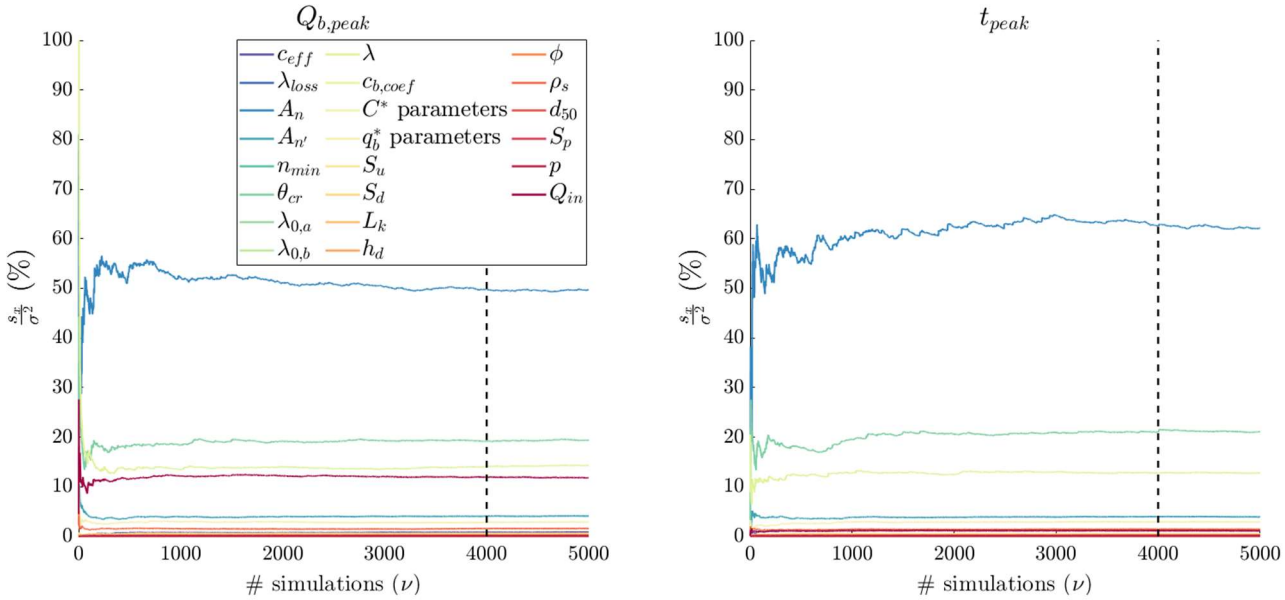


Figure S5. Convergence graphs of Sobol indices of total order for all input variables and both model outputs of interest using Method 1.

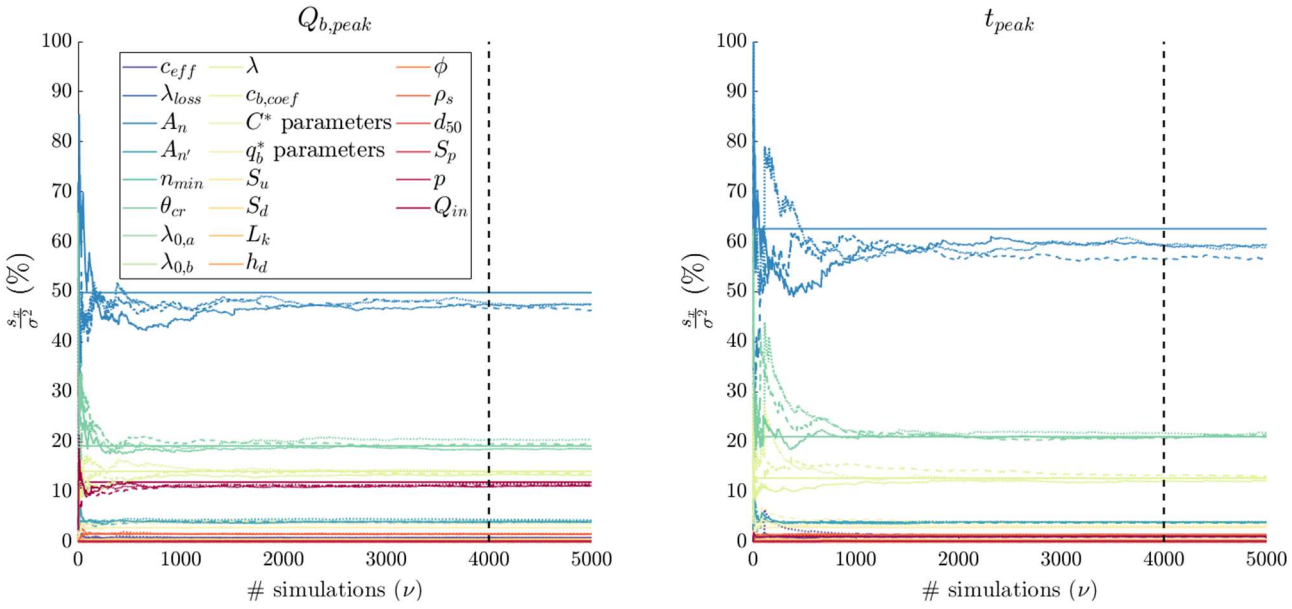


Figure S6. Accuracy and convergence evaluation of Sobol indices of total order obtained using Method 2. Legend: Horizontal plain lines = converged values obtained with Method 1; dashed curves = Method 2 with k = 500; dotted curves = Method 2 with k = 1000; plain curves = Method 2 with k = 2000.

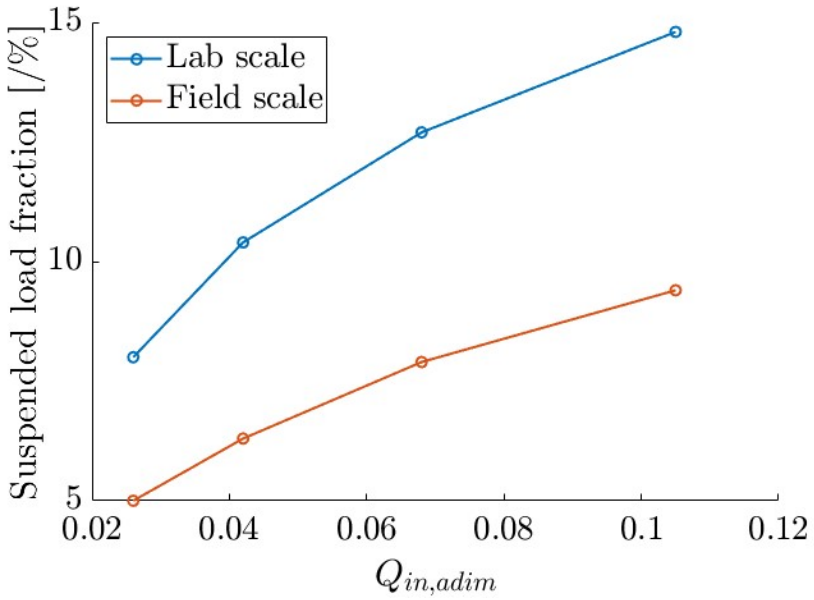


Figure S7. Mean percentage of overall sediment concentration corresponding to suspended load in lab and field -scale cases (Tests 18 to 21).

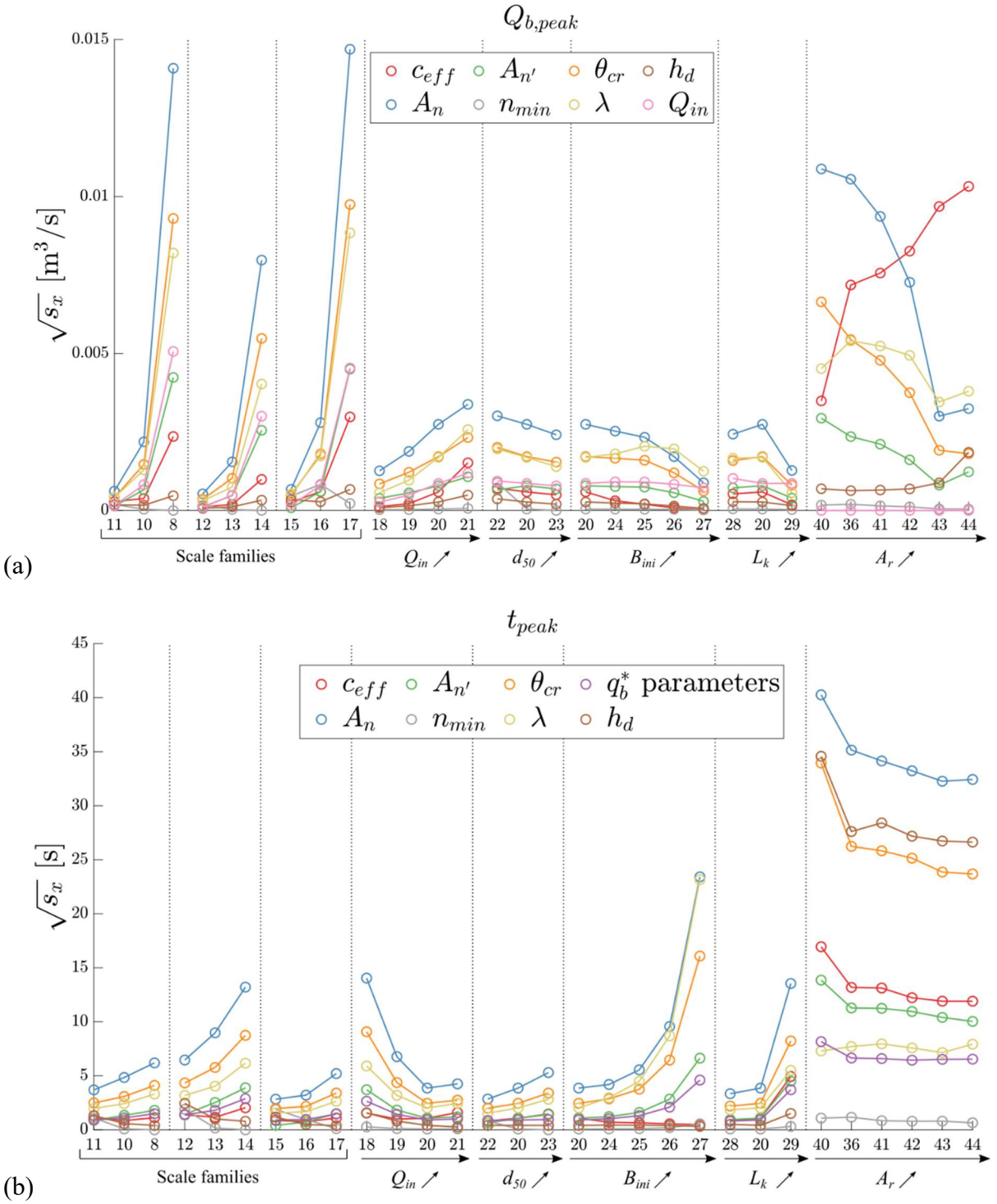


Figure S8. Significance descriptors in laboratory-scale configurations.

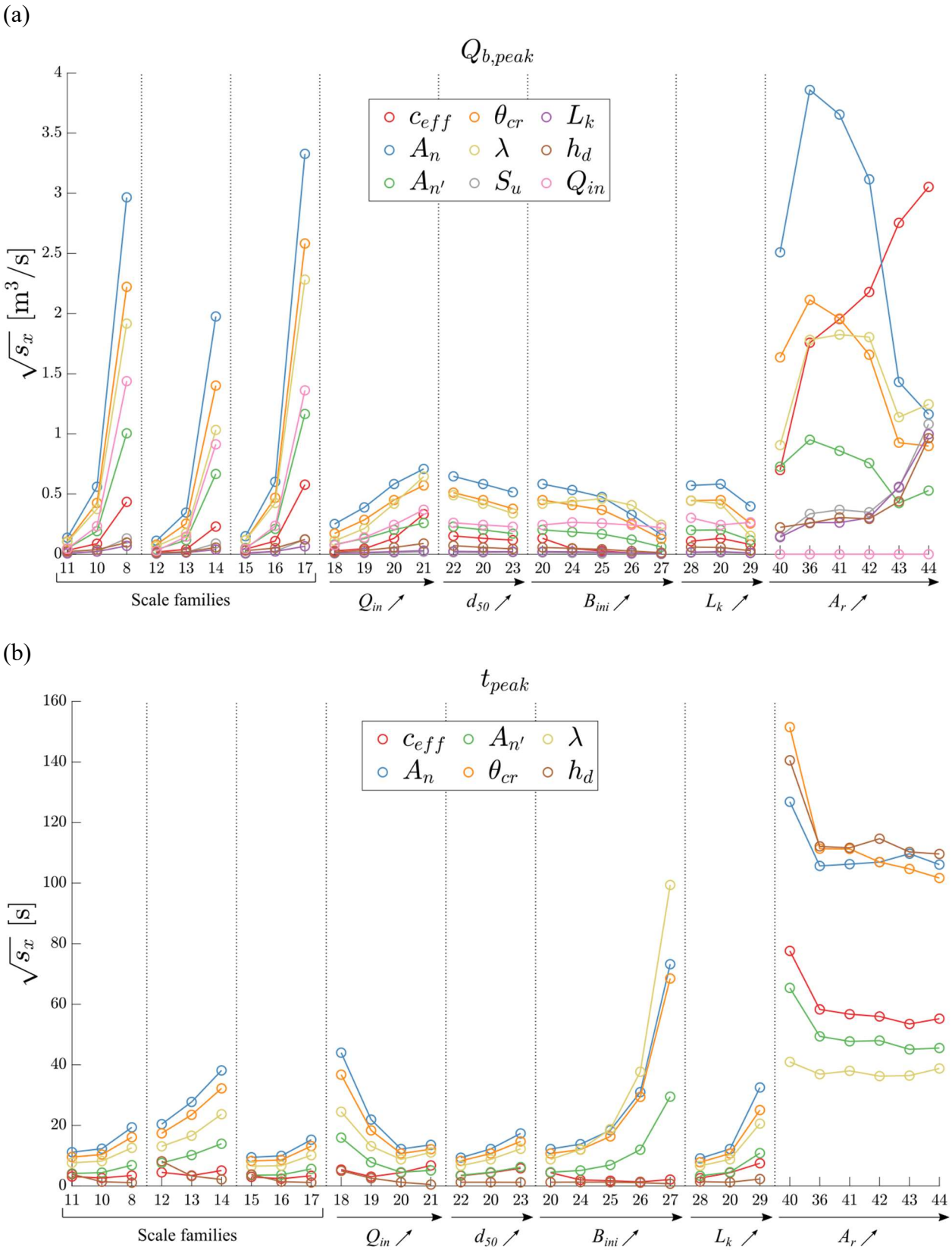


Figure S9. Significance descriptors in field-scale configurations.

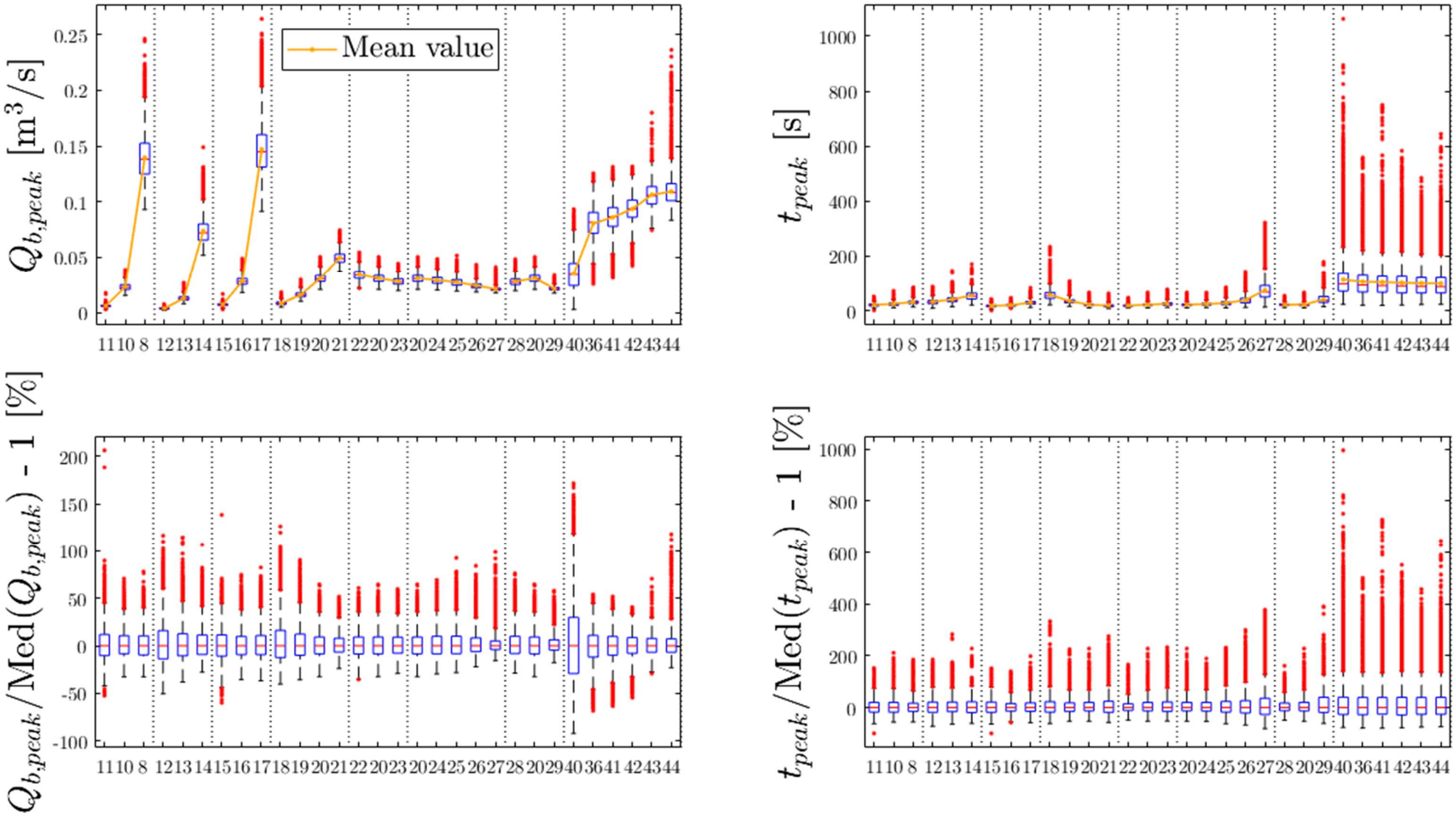


Figure S10. Boxplots of model outputs and their variation with respect to their median value in laboratory-scale configurations.

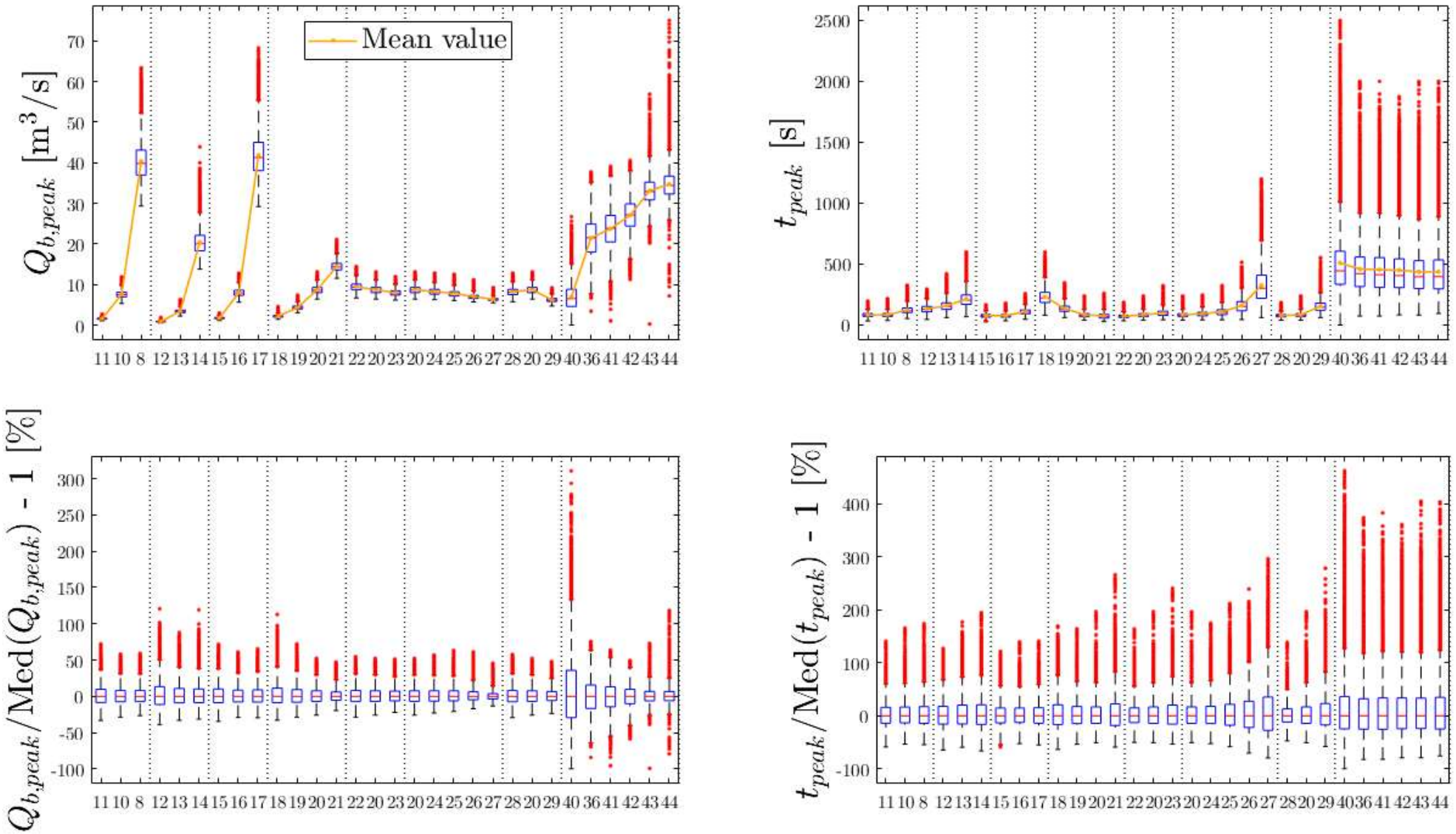


Figure S11. Boxplots of model outputs and their variation with respect to their median value in field-scale configurations.

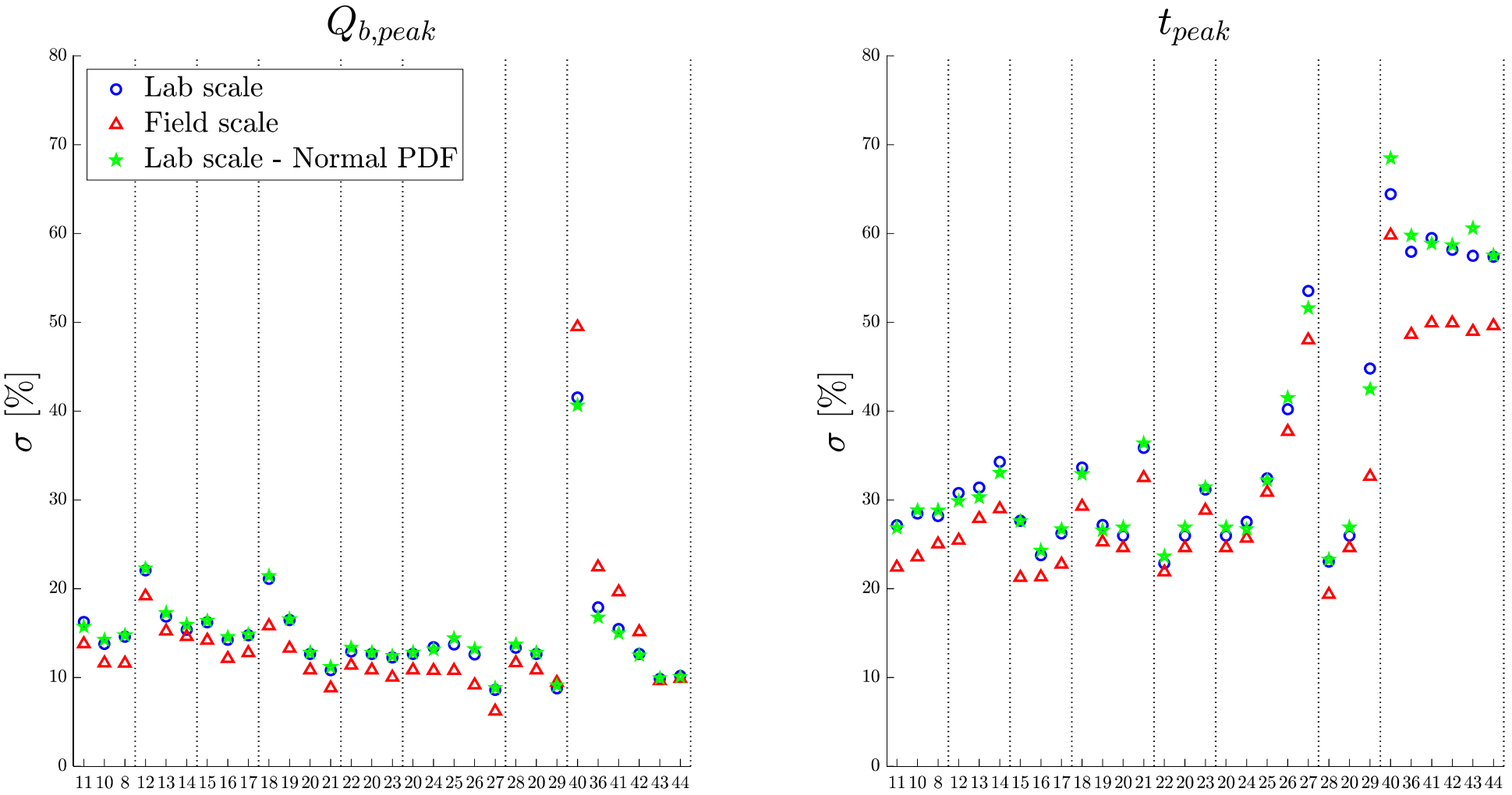


Figure S12. Standard deviation of the normalized and centred results = std(*output*/median(*output*) - 1). Green markers correspond to laboratory-scale configurations with truncated normal distributions associated to parameters c_{eff}, A_n, Q_{in}, θ_{σ}, λ and h_d.

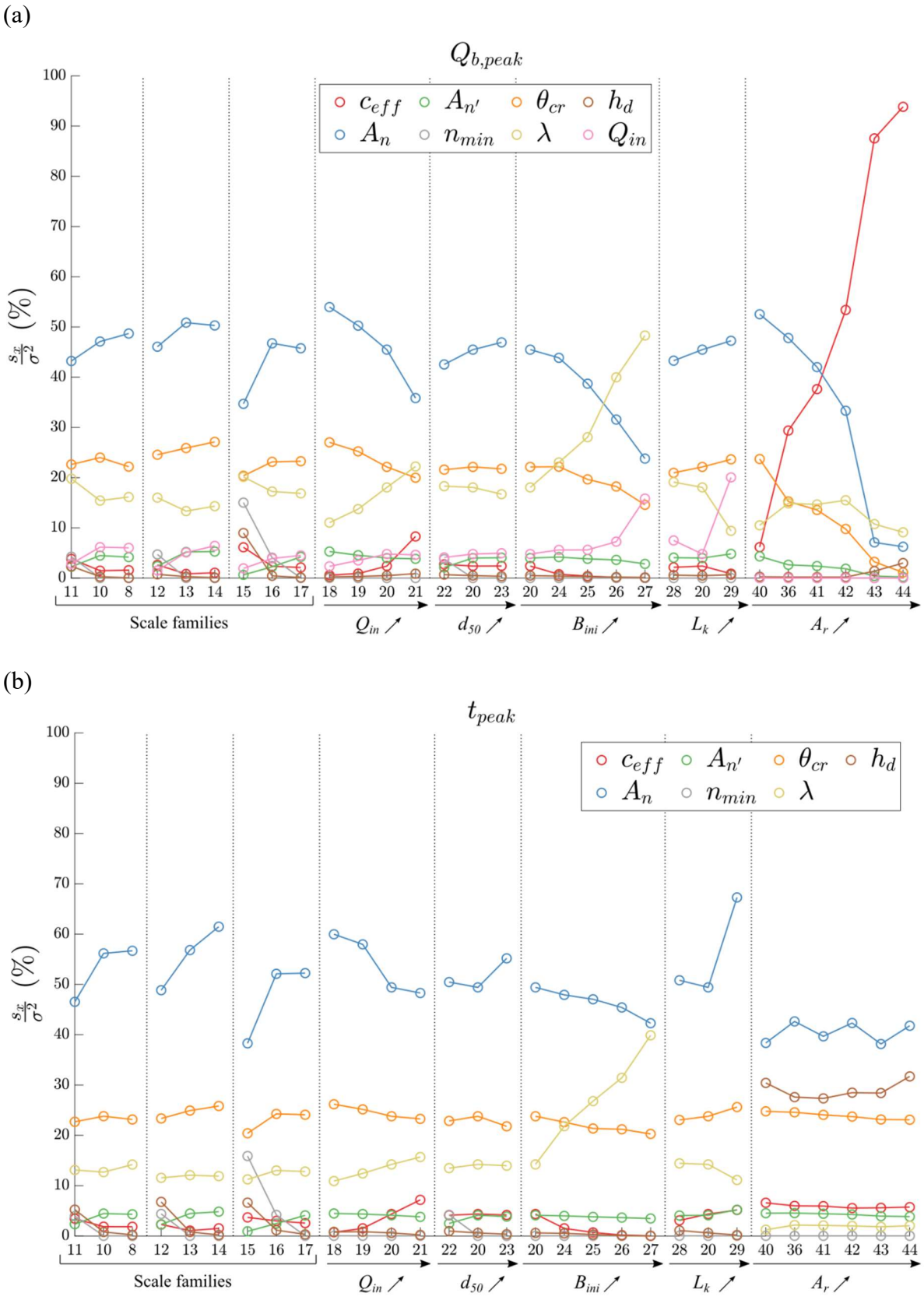


Figure S13. Sobol indices of total order for (a) peak breach discharge, and (b) time to peak in laboratory-scale configurations with truncated normal distributions associated to parameters c_{eff}, A_n, Q_{in}, θ_{cr}, λ and h_d.

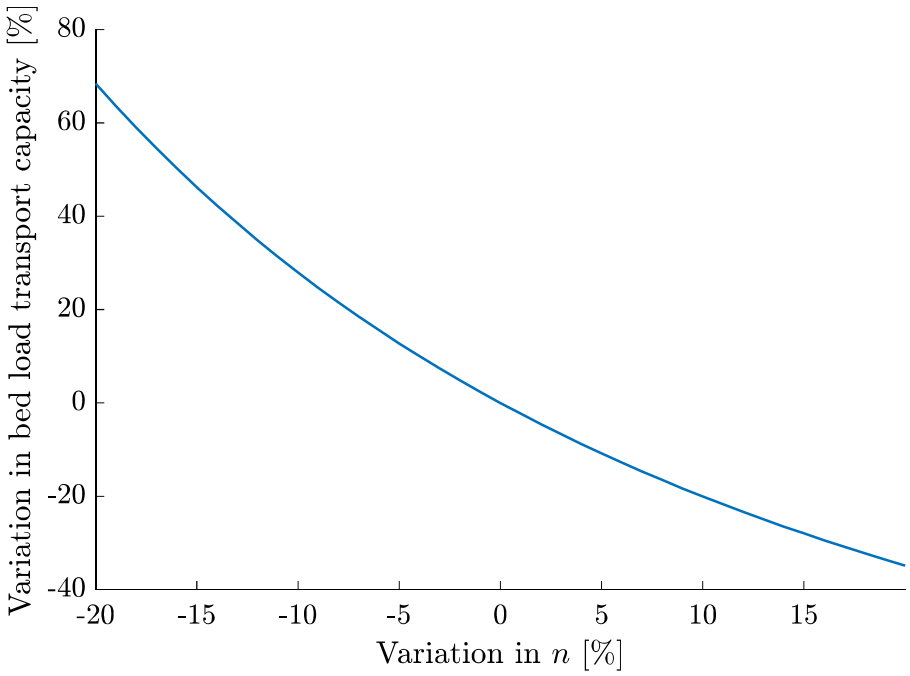


Figure S14. Variation in the bed load transport capacity as a function of a variation in the Manning's coefficient value in Test 10 at laboratory scale. The origin of the axes corresponds to the reference configuration of Test 10.

Chapter 2

Overtopping-Induced Failure of Non–Cohesive Homogeneous Fluvial Dikes: Effect of Dike Geometry on Breach Discharge and Widening

This chapter corresponds to the journal paper "Overtopping-Induced Failure of Non–Cohesive Homogeneous Fluvial Dikes: Effect of Dike Geometry on Breach Discharge and Widening" by V. Schmitz, S. Erpicum, K. El Kadi Abderrezzak, I. Rifai, P. Archambeau, M. Pirotton and B. Dewals, published in 2021 in Water Resources Research. The PhD candidate conducted the analysis of the experimental data, generated the figures, and wrote the manuscript.

Overtopping-Induced Failure of Non–Cohesive Homogeneous Fluvial Dikes: Effect of Dike Geometry on Breach Discharge and Widening

V. Schmitz[1], S. Erpicum[1], K. El kadi Abderrezzak[2,3], I. Rifai[4], P. Archambeau[1], M. Pirotton[1] & B. Dewals[1]

[1] *Research Group of Hydraulics in Environmental and Civil Engineering (HECE), University of Liège, Liège, Belgium.*

[2] *National Laboratory for Hydraulics and Environment (LNHE), EDF R&D, Chatou, France.*

[3] *Saint Venant Laboratory for Hydraulics, Chatou, France.*

[4] *Egis, Guyancourt, France.*

Corresponding author: Vincent Schmitz (V.Schmitz@uliege.be)

Key Points:

- Based on laboratory experiments, the influence of dike geometry (channel-side and floodplain-side slopes and crest length) on breach hydrograph and widening has been assessed.
- Three types of breach hydrographs have been identified and related to the inflow discharge and dike geometry.
- The conditions of occurrence of the three types of breach hydrographs have been clarified based on a simple conceptual model.

Chapter 2: Influence of Dike Geometry

ABSTRACT

Laboratory experiments were conducted to assess the influence of dike geometry on the breaching of non-cohesive homogeneous fluvial dikes. Both the channel-side and floodplain-side dike slopes and the crest length were varied systematically. The time-evolution of the breach discharge and breach width was monitored. Dikes having a larger volume per unit width lead to a more gradual increase in breach discharge and in breach width during the first stage of breach expansion (i.e., phase of rapid erosion). In contrast, the later stage of gradual breach widening is less influenced by the dike geometry. The breach hydrographs were observed to follow three distinct patterns, which are explained based on the relative magnitude of two characteristic time scales and of a normalized form of the dike unit volume.

1. Introduction

1.1. Context

For millennia, fluvial dikes (also called levees) have been built along long stretches of the world's rivers to protect population and property from flooding (Ward et al., 2017). However, failure of fluvial dikes leads to devastating human, economic and environmental consequences worldwide (e.g., Bhattarai et al., 2015; LaRocque et al., 2013; Viero et al., 2013; Vorogushyn et al., 2010). The main causes of failure include flow overtopping (Schmocker & Hager, 2009), piping (Vorogushyn et al., 2009), seepage (Michelazzo et al., 2018b), slope instability, and animal burrowing (Orlandini et al., 2015; Viero et al., 2013). Flow overtopping is by far the most frequent cause of failure (Danka & Zhang, 2015; Özer et al., 2020; Zhang et al., 2016). In the following, we focus on overtopping-induced breaching of fluvial dikes.

The risk caused by dike failure is rising due to lack of maintenance, inadequate rehabilitation works and increasingly frequent hydrological extremes, both floods and droughts. Indeed, during prolonged drought periods, cracks may appear in the dike body, thus increasing the failure probability (Van Lanen et al., 2016; Ward et al., 2020). Consequences of dike failure are often magnified by the increase in exposure and socioeconomic activity typically observed in the protected lowlands. This is known as the levee-effect *(*e.g., Aerts et al., 2018; Di Baldassarre et al., 2018). Therefore, accurately modelling the dike breaching process and the induced flooding is of critical importance for flood risk management. Since field measurements are complex and hazardous whilst experimental tests are relatively costly, computational models simulating the breaching of fluvial dikes are of prime interest. However, those models are not yet mature (e.g., Dou et al., 2014; Elalfy et al., 2018; Faeh, 2007; Kakinuma & Shimizu, 2014), and more experimental research is needed to unravel the underpinning breaching mechanisms.

1.2. State-of-the-art

Compared to the case of fluvial dikes, more experimental research has been conducted on the failure of embankment dams, i.e., earthen or rockfill structures normal to the main river channel (Amaral et al., 2020; Coleman et al., 2002; Bento et al., 2017; Jandora & Říha, 2008; Müller et al., 2016; Pickert et al., 2011; Morris et al., 2007; Sadeghi et al., 2020;

Schmocker & Hager, 2009; Schmocker & Hager, 2012; 2014; among others). In contrast, studies aiming at a deep understanding of the breaching mechanisms of fluvial dikes developed mostly in the last decade (Table 1). The breaching of fluvial dikes exhibits distinct features than those involved in embankment dams (ASCE/EWRI Task Committee on Dam/Levee Breaching, 2011). This includes complex 3D flow structures involving spiral flow in the breach (Michelazzo et al., 2015), non-symmetric breach expansion and impinging jet erosion (e.g., Elalfy et al., 2018; Kakinuma et al., 2013; Rifai et al., 2017).

Field-scale observations of fluvial dike breaching remain scarce, due to the difficulty in setting up large-scale tests, controlling boundary conditions and recording accurate measurements. To our knowledge, only Kakinuma et al. (2013) conducted four field tests of overtopping of 3 m high dikes made of sand and gravel. The dike composition, crest width and inflow discharge in the main channel were varied.

Laboratory experiments have enabled parametric studies and more detailed monitoring of overtopping-induced breaching of dikes (Table 1). In these experiments, the considered dikes were homogeneous, and their height ranged between 0.10 m and 0.20 m, except for the studies by Rifai et al. (2017; 2018; 2021) who used 0.30 m high dikes. Virtually all studies varied the inflow discharge in the main channel. While in all other tests a straight main channel was considered, the dike was placed along the outer bank of a 180°-bend in the experiments of Yu et al. (2013), Wei et al. (2016) and Wu et al. (2018). Islam (2012), Bhattarai et al. (2015) and Michelazzo et al. (2018a) used an erodible bottom in their laboratory setup, enabling the analysis of scouring on the floodplain side of the dike. Rifai et al. (2018) considered a confined floodplain to assess the tailwater effect on breach expansion.

Charrier (2015), Wei et al. (2016) and Wu et al. (2018) studied the failure of dikes made of cohesive material. For non-cohesive material, the influence of the dike composition was analysed by Islam (2012), Bhattarai et al. (2015) and Wu et al. (2018). Results showed generally faster erosion for tests with coarser grains. Tests by Rifai et al. (2021) involved the breaching of dikes made of sand and silt in various proportions (up to 30 % volumetric ratio of silt), indicating a limited influence of the silt content on the overall breach dynamics.

In contrast, the influence of the dike geometry was never investigated systematically. Islam (2012) considered three different dike heights; but these variations were obtained from a change in the main channel bottom elevation with respect to the floodplain level, resulting in a different dike height on the channel-side and on the floodplain-side. Yu et al. (2013) analysed three distinct dike geometries, but several parameters were changed simultaneously (crest length, slope of dike faces as well as dike composition), so that the results do not enable disentangling the effect of individual geometric parameters. Wu et al. (2018) tested two different relative crest lengths. However, Yu et al. (2013) and Wu et al. (2018) placed the dike in a 180° bend, thus hampering the transposition of their findings to the more canonical case of a fluvial dike along a straight main channel.

1.3. Objective

Knowledge gaps remain regarding how the breaching process is influenced by dike geometric parameters, such as the ratio of the crest length to the dike height (L_K / w), or the channel- and floodplain-side slopes (S_u and S_d, respectively, as shown in Figure 1). The influence of these parameters was studied for the frontal configuration (e.g., Müller et al., 2006), but not for a fluvial dike. In this paper, we aim to unveil the impact of variations in the geometry of a non-cohesive dike on the breach widening and outflow hydrograph in case of flow overtopping. Based on the same laboratory setup as used by Rifai et al. (2017; 2018), we present the results of tests covering eight distinct dike geometric configurations, in combination with three inflow discharges in the main channel.

The laboratory setup, test program and measurement techniques are presented in Section 2. Section 3 details the observed evolution of the breach discharge and width. In Section 4, we provide a typology of breach hydrographs and hint at an explanation based on the relative magnitude of two time scales and of the dike normalized unit volume. Conclusions are drawn in Section 5.

Table 1. Previous experimental research on overtopping-induced breaching of homogeneous fluvial dikes.

		Experimental setup				Parameters systematically varied in the tests			
Reference	Laboratory	Dike material	Dike height w (m)	Erodible bottom	Number of tests	Dike composition	Inflow rate Q_{in}	Dike height w	Others
Islam (2012)	Nagoya University	C & NC	0.10 – 0.20	✓	6	✓	✓	✓(1)	MC bottom elevation
Bhattarai et al. (2015)	Kyoto University	NC (sand)	0.15	✓	21	✓			Bottom erodibility
Charrier (2015)	University of Aix-Marseille	C (clay content)	0.10	-	5	✓	✓		
Yu et al. (2013)	Wuhan University	NC (sand or coal particles)	0.15 – 0.18	-	12		✓	✓(2)	Curved MC
Wei et al. (2016)	Wuhan University	C	0.16	-	4	✓	✓		Curved MC
Elalfy et al. (2018)	University of South Carolina	NC (sand)	0.20	-	2		✓		
Michelazzo et al. (2018a)	LWI Braunschweig	NC (sand)	0.25	✓	10		✓		
Wu et al. (2018)	Wuhan University	NC (sand)	0.15 – 0.18	-	9	✓	✓	✓(2)	Curved MC
		C (silt loam)	0.18	-	3	✓	✓		
Rifai et al. (2017) Rifai et al. (2018)	University of Liege	NC (sand)	0.30	-	23		✓		Backwater effects
Rifai et al. (2021)	EDF R&D, LNHE Chatou	NC (sand and silt)	0.30	-	8	✓	✓		
Present study	University of Liege	NC (sand)	0.30	-	26(3)		✓		L_k, S_u, S_d

(1) The dike height varies as a result of changes in the bottom elevation of the main channel, which may differ from the floodplain elevation.
(2) The main channel is curved, forming a 180° bend. (3) As detailed in Section 2.2, these 26 tests include four repetitions to assess the results reproducibility.

Notations and abbreviations: w = dike height; Q_{in} = inflow discharge in the main channel; L_k = length of dike crest; S_d = slope of the floodplain-side of the dike, S_u = slope of the channel-side of the dike; C = cohesive; NC = non-cohesive; BC = boundary condition; MC = main channel.

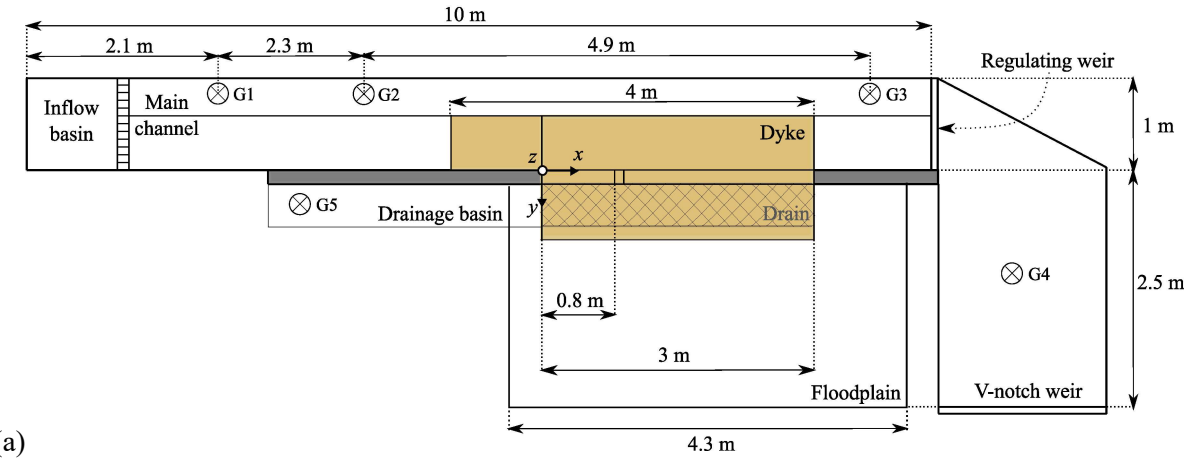
2. Laboratory experiments

2.1. Experimental setup

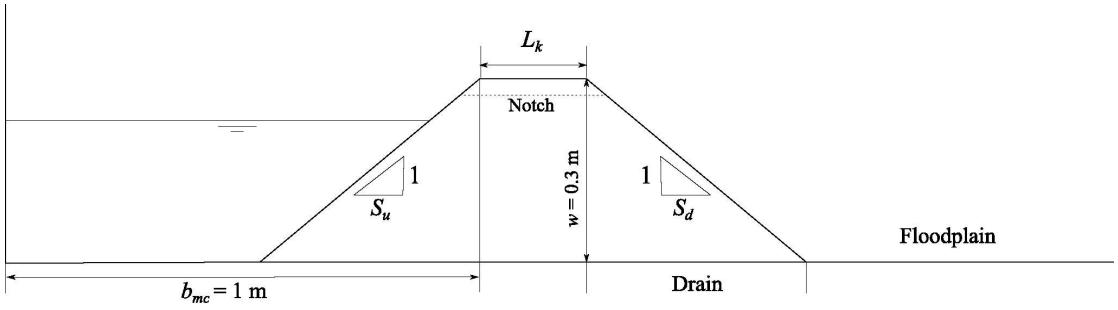
The experimental setup was located in the Laboratory of Hydraulic Engineering of the University of Liege (Figure 1). It consists of a horizontal, straight main channel (10 m × 1 m) of trapezoidal cross-section, with a 3-meter-long lateral opening towards a floodplain (4.3 m × 2.5 m). Along this side opening, a trapezoidal-shaped fluvial dike was built with uniform sand of median diameter d_{50} = 1 mm, following the same procedure as used by Rifai et al. (2017; 2018; 2019). The main channel and floodplain were covered with an impermeable whitewash coating to ensure roughness continuity between the flume, floodplain and sand dike (Rifai et al., 2017). The bottom of the main channel and the floodplain were at the same level. To control seepage flow through the dike body, a drainage system was installed at the dike bottom (Rifai et al., 2017; 2019). Finally, a perforated plate followed by a reservoir with a calibrated weir was placed at the main channel end to control the downstream water level and collect the main channel outflow. A full description of the laboratory setup was provided by Rifai et al. (2017; 2019).

The dike geometry is characterized by its height w as well as by the channel-side slope S_u, the floodplain-side slope S_d and the crest length L_k (Figure 1b). The dike height was kept constant at w = 0.3 m in all tests, while the three other parameters were systematically varied.

Each geometric configuration was tested with three different main channel inflow discharges. The inflow discharge Q_{in} was kept constant and supplied at the main channel upstream end. During the filling phase, the water level rose in the main channel. The perforated plate was adjusted such that, at the beginning of each test, the water level corresponded to the dike crest elevation. To trigger breaching, a 2-cm-deep and 10-cm-wide initial notch was created in the dike crest, at a distance of 0.8 m from its upstream end. Once the water level overtopped the notch, surface erosion and breaching developed.



(a)



(b)

Figure 1. Experimental setup: (a) plane view; (b) dike vertical profile, highlighting main geometric parameters: channel-side and floodplain-side dike slopes (S_u and S_d), height (w) and crest length (L_k).

The primary goal of the present study is to investigate the breaching mechanisms of fluvial dikes, and not to represent a scaled particular field structure. Nonetheless, assuming that material entrainment and failure are dominating processes in the breaching of non-cohesive homogeneous dikes, it is reasonable to consider the similarity of Shields number and of the side slope stability factor in the interpretation of the laboratory tests (El kadi Abderrezzak et al., 2014). Considering the Froude similarity for undistorted moveable model, with a sufficiently high Reynolds number, the grain size scales like the flow depth for alluvial material (i.e., sediment density and angle of repose are almost similar in the model and prototype). Therefore, the current experimental model may be seen as a 1:10 scale model of a 3-m high prototype dike made of medium gravels of 1 cm in diameter.

2.2. Test program

The test program is detailed in Table 2. Eight different dike geometries were considered, as well as three different inflow discharges in the main channel: Q_{in} = 25 l/s, 40 l/s and 55 l/s. A total of 26 experiments were carried out, with four of them aiming at assessing the test reproducibility.

The first configuration (A) leads to non-dimensional geometric parameters identical to those of the reference dike used by Müller et al. (2016): $S_u = S_d = 2$ and $L_k = 0.15$ m (i.e., $L_k / w = 0.5$). This dike geometry is used as a reference here. The seven other geometric configurations were defined by varying one geometric parameter at a time. While the floodplain-side of the dike was varied between $S_d = 1.5$ and $S_d = 3$, the channel-side of the dike was only varied between $S_u = 1.5$ and $S_u = 2.0$ for two reasons. First, tests involving variations in S_u were particularly demanding because they required an adaptation of the main channel cross-section over the whole length of the laboratory setup, whereas varying S_d or L_k implied only a change in the construction of the sandy dike. Second, using $S_u = 3$ would lead to a dike extending over 90 % of the main channel width (Figure 1a), which was deemed unrealistic. Configuration H was only tested with Q_{in} = 55 l/s due to difficulties in triggering breaching for this particularly strong (i.e., with a relatively large volume per unit width) dike ($L_k / w = 2$). Table 2 also provides the value of μ, defined as the dike volume per unit width ($L_k\, w + w^2\, (S_u + S_d) / 2$) normalized by its value for a dike of identical height w but characterized by the following non-dimensional geometric parameters: $S_u = S_d = 2$ and $L_k / w = 0.5$. This indicator was introduced by Müller et al. (2016) for studying the influence of dam geometry on the breaching in frontal configuration.

Table S1 in Supplement details each individual test, highlighting among others small deviations (mean value of 3 %) between the target inflow discharge (as mentioned in Table 2) and the actual one. Indeed, in each test, the inflow discharge was slightly adapted compared to its target value, to ensure that the water level in the main channel just exceeds the initial notch level at the beginning of the experiment.

Table 2. Test program. The four tests corresponding to underlined values of the inlet discharge Q_{in} were repeated twice to assess the reproducibility of the observations.

Conf. ID	S_u (-)	S_d (-)	L_k (m)	μ	Q_{in} (l/s)			F		
A	2.0	2.0	0.15	1	<u>25</u>	<u>40</u>	55	<u>0.083</u>	<u>0.133</u>	0.183
B	1.5	2.0	0.15	0.9	25	40	55	0.071	0.114	0.157
C	2.0	1.5	0.15	0.9	25	40	<u>55</u>	0.083	0.133	<u>0.183</u>
D	2.0	2.5	0.15	1.1	25	40	55	0.083	0.133	0.183
E	2.0	3.0	0.15	1.2	25	40	55	0.083	0.133	0.183
F	2.0	2.0	0.00	0.8	25	<u>40</u>	55	0.083	<u>0.133</u>	0.183
G	2.0	2.0	0.30	1.2	25	40	55	0.083	0.133	0.183
H	2.0	2.0	0.60	1.6	-	-	55	-	-	0.183

2.3. Measurement techniques

The water level was measured at three locations in the main channel (G1, G2 and G3), in the outflow tank (G4) and in the drainage tank (G5) using ultrasonic sensors (accuracy of ± 1 mm) (Figure 1a). The inflow discharge Q_{in} was measured using an electromagnetic flowmeter (accuracy of ± 0.5 % of full scale, i.e., 150 l/s). The outflow discharge was deduced from the discharge passing through a V-notch weir (deduced from water level at G4) and mass balance in the outflow tank. The drainage discharge Q_d was estimated from the water level measured at G5. The breach discharge Q_b was determined from mass balance in the main channel (Rifai et al., 2017; 2019).

The experiments were recorded with a digital video camera (Panasonic GH4) set on Full-HD resolution (1920 × 1080 pixels) and with a recording speed of 60 frames/s. Compared to previous studies (Table 1), a relatively high number of tests was performed here. To monitor the breach widening, we used a simplified version of the laser profilometry technique (LPT) presented by Rifai et al. (2020): the laser sheet was not rotated in our case, which enabled us to reliably monitor the breach widening at the crest level as the laser sheet was aligned with the dike crest, but not to produce 3D reconstructions of the breach evolving geometry.

3. Results

Figure 2 and Figure 3 represent the evolution of the breach discharge and of the breach widening for each tested geometric configuration and each inflow discharge in the main channel. The evolution of the water level in the main channel is shown in Figure S1 in Supplement. Snapshots of experimental tests in Configuration A (S_u = 2, S_d = 2, L_k = 0.15 m) at four different times for three inflow discharges are available in Table S3 in Supplement.

3.1. Stages of breach expansion

As detailed by Rifai et al. (2017, 2018), the breach evolution can be subdivided into three main stages: gradual start of overtopping at the initial notch, with a slow initiation of dike erosion (Stage 0), rapid erosion, leading to a fast increase in the breach size (width and depth) and discharge (Stage 1), quasi-stabilization of the flow with a reduced flow depth in the main channel and, generally, a continuing but slow breach expansion towards downstream (Stage 2). Note that our analyses do not focus on Stage 0, which is affected by the initial notch characteristics and channel filling procedure.

In Figures 2 and 3, Stages 1 and 2 can be clearly distinguished. Almost all curves start with a steep initial increase in the breach discharge and in the downstream widening, followed by a relatively abrupt slope reduction indicating the transition between Stage 1 and Stage 2. In general, the lower the dike unit volume (i.e., lower values of S_u, S_d or L_k, hence also of μ), the more noticeable the transition between Stage 1 and Stage 2. Only in the case of a particularly strong dike (longest considered crest length: L_k = 0.6 m, μ = 1.6) and highest inflow discharge (Q_{in} = 55 l/s), no clear transition can be detected.

In most tests, the breach invert reached the main channel bottom over at least part of the breach width; but this was not the case for the strongest dike (Configuration H, L_k = 0.6 m). A more thorough description of the breaching mechanism is provided by Rifai et al. (2017), and is not repeated here for the sake of brevity. The reproducibility of the tests can be assessed by comparing curves of identical colour in a same panel in Figures 2 and 3. The relative difference in breach discharge between two repeated tests is lower than 10 % of the corresponding inflow discharge, as shown in Figure S2.

3.2. Influence of dike geometry in Stage 1

In several cases, a global or a local maximum can be seen in the breach hydrograph (Figure 2). When the inflow discharge is relatively small (Q_{in} = 25 l/s), all hydrographs exhibit a global maximum, except in Configuration E which corresponds to the strongest dike (S_d = 3, μ = 1.2) tested for this value of Q_{in}. Moreover, the weaker the dike, the more pronounced the magnitude of the peak in the breach discharge: the maximum value of the breach discharge reaches about 115 % of Q_{in} for $\mu \leq 1$, while the ratio breach discharge Q_b to Q_{in} ranges between 105 % and 110 % when $\mu > 1$. For a moderate inflow discharge (Q_{in} = 40 l/s), a local maximum can be noticed only for particularly weak dikes, i.e., in Configuration C (S_d = 1.5, μ = 0.9) and Configuration F (L_k = 0 m, μ = 0.8). No maximum in the breach hydrograph is visible for stronger dikes ($\mu \geq 1.0$), nor when the inflow discharge becomes higher (Q_{in} = 55 l/s).

When the inflow discharge is increased, the breach expansion towards downstream intensifies, whereas the expansion towards upstream is reduced (Figure 3). For the highest tested inflow discharge (Q_{in}= 55 l/s), the breach does not expand towards upstream, irrespective of the dike geometry, except for the strongest dike (L_k = 0.6 m, μ = 1.6). This appears consistent with the higher flow momentum parallel to the dike axis when Q_{in}= 55 l/s, leading to stronger erosion by impinging jet on the downstream part of the breach, and a gradual shift of the breach centreline towards the channel downstream end.

As can be seen in Figure 2, the gradient in breach discharge during Stage 1 is milder for stronger dikes (i.e., higher values of μ) than for weaker dikes. This influence of the dike geometry on the breach hydrograph is magnified when the inflow discharge is higher. Indeed, the differences between the breach hydrographs are stronger in the cases Q_{in} = 55 l/s than they are for Q_{in} = 40 l/s and for Q_{in} = 25 l/s. These observations are well supported by Figure S3 in Supplement, which plots a characteristic rising time of the breach hydrograph in Stage 1, as a function of the geometric parameters S_u, S_d, L_k and μ: the stronger the dike, the shorter the rising time of the breach hydrograph. The effect of changing S_u seems smaller; but it is certainly due to its smaller range of variation (between 1.5 and 2.0) compared to that of S_d (between 1.5 and 3.0). Also, the change in the rising time with the dike geometry is more pronounced for higher values of Q_{in}. Similarly, Figure 3 reveals that the breach expands faster during Stage 1 in the case of a weaker dike, while for stronger dikes it widens more gently in both up- and downstream directions.

3.3. Influence of dike geometry in Stage 2

During Stage 2, the breach expansion rate is less influenced by the dike geometric parameters than it is during Stage 1 (Figure 3). Only variations in the channel-side dike slope S_u seem to have a significant effect on the breach widening rate during Stage 2. This effect is magnified when Q_{in} is increased. However, this may result from the fact that varying the channel-side dike slope S_u while keeping the inflow discharge Q_{in} constant modifies the cross-section in the main channel, and therefore also the flow velocity, the Froude number F and the storage volume in the main channel.

Similarly, the breach discharge in Stage 2 is little affected by the dike geometric characteristics. For relatively low and intermediate inflow discharges in the main channel (Q_{in} = 25 l/s and Q_{in} = 40 l/s), the breach discharge in Stage 2 converges towards a similar value, fairly close to Q_{in}, irrespective of the dike geometry. Figure S4 in Supplement displays the breach hydrographs based on a logarithmic axis for time. This enables a better appraisal of the convergence of the breach discharge towards a quasi-equilibrium value Q_{NE} (Michelazzo et al., 2018a). For Q_{in} = 25 l/s, a quasi-equilibrium breach discharge is reached before the end of the tests, with a value ranging between 90 % and 95 % of the inflow discharge, regardless of the dike geometric parameters. Similar results are obtained for Q_{in} = 40 l/s, whereas for Q_{in} = 55 l/s the tests were stopped before the breach discharge reached a quasi-equilibrium value, because the downstream breach expansion reached the limit of the 3 m long erodible dike section.

Figure S5 in Supplement represents the quasi-equilibrium breach discharge Q_{NE} (normalized by the inflow discharge Q_{in}) as a function of each geometric parameter (S_u, S_d, L_k) and as a function of the dike volume per unit width (μ). No significant trend can be detected, confirming the limited influence of the dike geometry on Q_{NE} and, more generally on Stage 2 of the breach expansion.

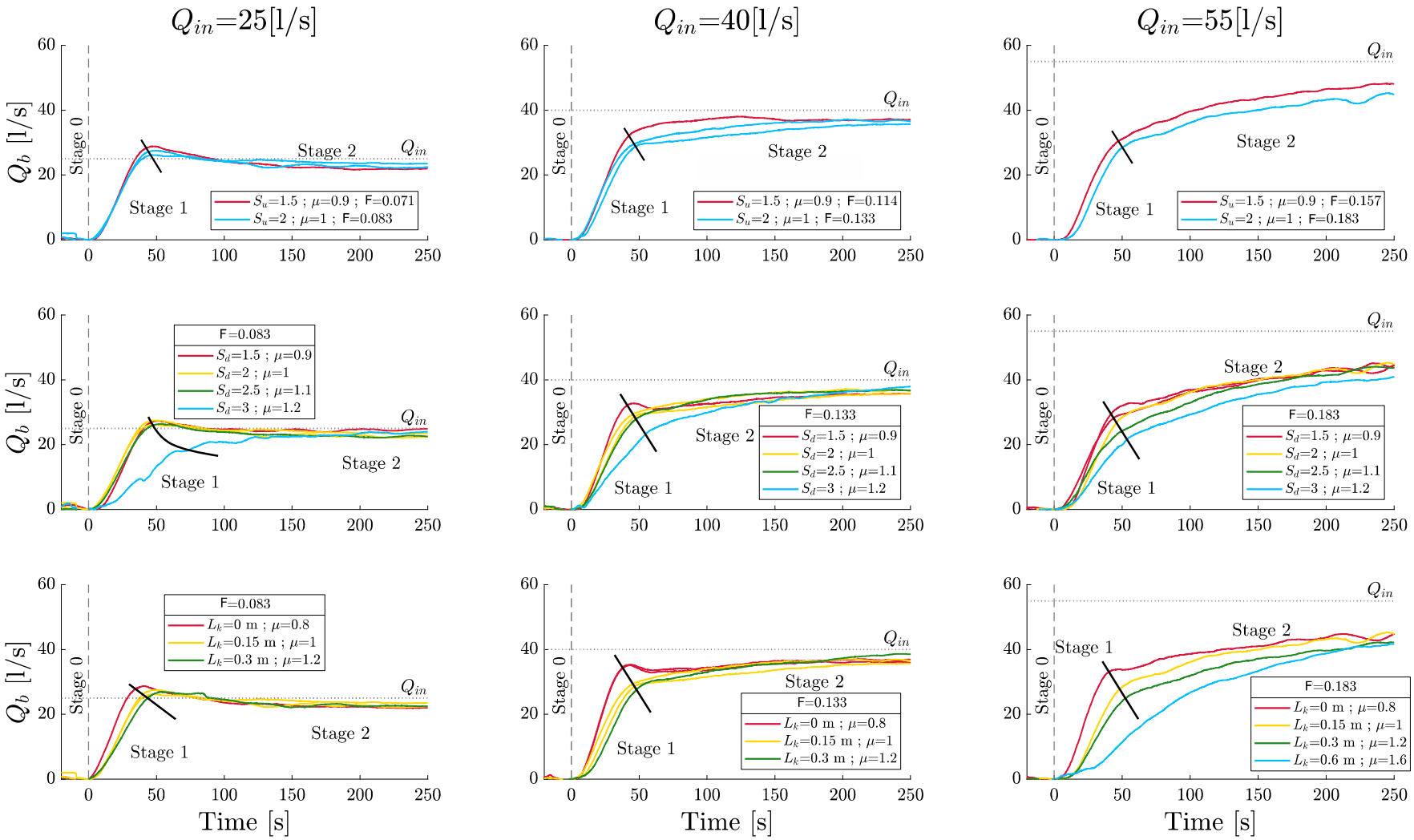


Figure 2. Evolution of the breach discharge Q_b for all tested geometric configurations and inflow discharges. In each plot, values of parameters not specified in the legend are set as in the reference case, i.e., Configuration A: $S_u = 2$, $S_d = 2$, $L_k = 0.15$ m. The black lines highlight the transition between Stage 1 and Stage 2. Origin of time corresponds to the start of Stage 1. Curves of identical colour in a same panel refer to repeated tests.

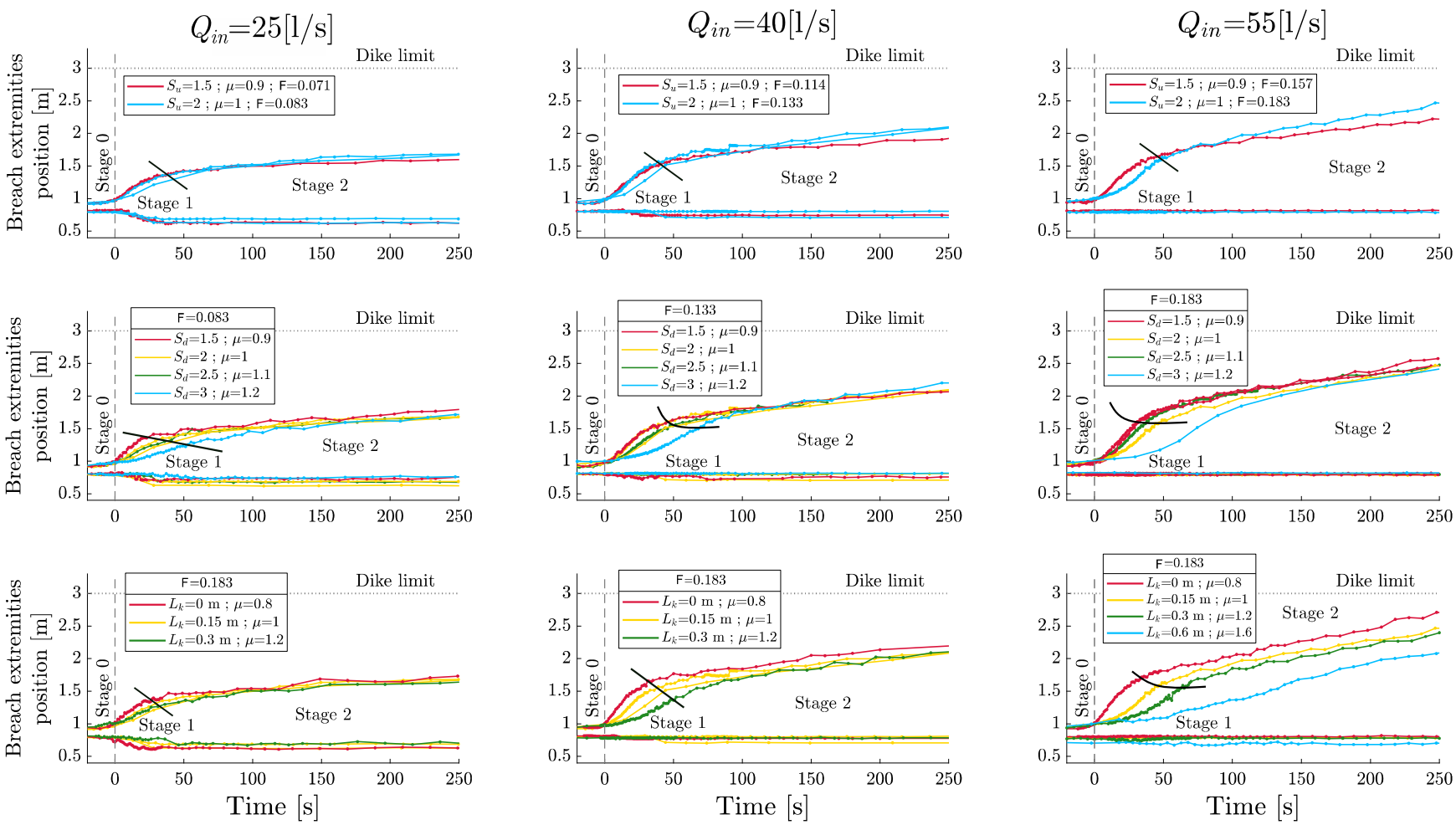


Figure 3. Evolution of the breach extremities location (at the crest level, along the crest centre line), with initial notch at x = 0.8 m, for all tested geometric configurations and inflow discharges. In each plot, parameters not specified in the legend are set as in the reference case, i.e., Configuration A: S_u = 2, S_d = 2, L_k = 0.15 m. The black lines (or curves) highlight the transition between Stage 1 and Stage 2. Origin of time axis corresponds to the start of Stage 1. Curves of identical colour in a same panel refer to repeated tests. Data could not be retrieved for one repeated test (Run 8-R in Table S1) in Configuration F (μ = 0.8) with Q_{in} = 40 l/s.

4. Discussion

4.1. Typology of breach hydrographs

Based on the results presented in Figures 2 and 3, a standard approach consists in normalizing the observations to derive relationships between non-dimensional variables (e.g., Schmocker & Hager, 2012; Schmocker et al., 2014). In line with this approach, Figure S3 in Supplement suggests that the non-dimensional parameter μ introduced by Müller et al. (2016) succeeds in capturing the overall effect of parameters S_u, S_d and L_k. In contrast, it was not possible to normalize the breach hydrographs so that they feature onto a single curve, because they follow distinct patterns.

The observed breach hydrographs (Figure 2) follow three distinct patterns, as sketched in Figure 4:

- Type A hydrograph, showing a rapid rise followed by a global maximum, a decline and a plateau corresponding to the equilibrium breach discharge;
- Type B hydrograph, showing a rapid rise followed by a local maximum, a decline and a gradual rise towards an equilibrium or quasi-equilibrium breach discharge;
- Type C hydrograph, showing a continuously increasing evolution towards an equilibrium or quasi-equilibrium breach discharge.

As detailed in Table S2 in Supplement, Type A hydrographs are mostly observed for a relatively low Froude number in the main channel (i.e., relatively low inflow discharge) and relatively weak dikes. In contrast, Type C hydrographs are obtained for a higher Froude number in the main channel and stronger dikes. This is confirmed by the distribution of the data points in the scatter plot displayed in Figure 5.

Figure 5 displays the results of the present tests and tests conducted in previous studies on the breaching of non-cohesive fluvial dikes along a straight main channel. Based on the available data, it appears that when the product $F \times \mu$ remains below ~ 0.1, the observed breach hydrographs tend to be of Type A. For larger values of $F \times \mu$, the type of hydrograph depends directly on the value of the normalized unit volume μ of the dike: for μ smaller than ~ 0.9, Type B hydrographs are mostly observed, whereas Type C hydrographs are obtained for μ above ~ 0.9. As can be seen in Figure 5, these transitions are not clear-cut, and the criteria based on F and μ should, at this stage of the research, be interpreted as representative of transition zones. The blurred nature of the transitions suggests that additional parameters may also control which type of breach hydrograph is obtained. The validity of the transition criteria beyond the range of values of F and μ tested here should be verified based on future tests exploring a broader range of variation for these parameters.

Overall, data tend to agree with the present classification based on the Froude number and geometric parameter μ. The occurrence of distinct patterns in the breach hydrographs reflects the co-existence of multiple time scales in the processes underpinning fluvial dike breaching. Hereafter, a simple conceptual model is introduced to explain the occurrence of the three types of breach hydrographs.

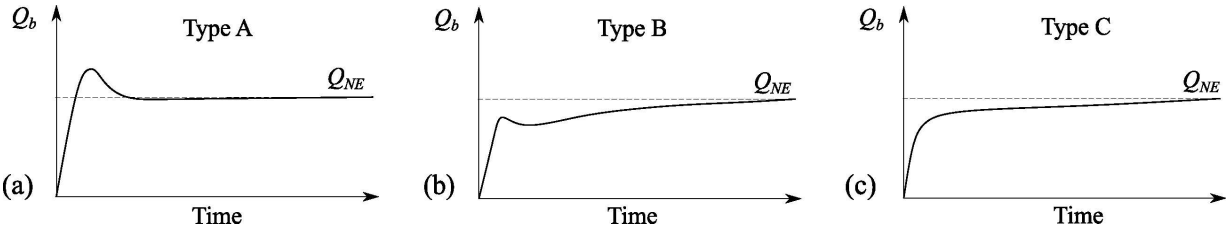


Figure 4. Breach hydrograph types: breach hydrograph with an absolute (a) or relative maximum (b), and a monotonously increasing breach hydrograph (c).

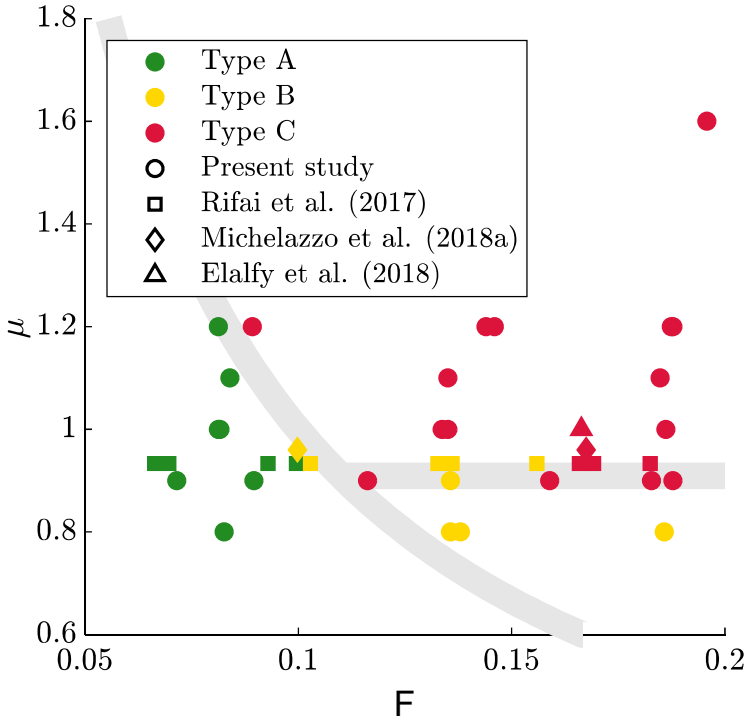


Figure 5. Occurrence of Types A, B and C breach hydrographs in the present study as well as in experimental tests conducted by Rifai et al. (2017), Michelazzo et al. (2018a) and Elalfy et al. (2018) at laboratory scale. F represents the Froude number in the main channel. Colours refer to the breach hydrograph type, while the symbol shape indicates the data source. The grey-shaded lines represent the transition zones corresponding to $F \times \mu = 0.1$, and $\mu = 0.9$ ($F \times \mu > 0.1$).

4.2. Multiplicity of time scales

The hydrodynamic process at stake during the failure of a fluvial dike can be seen as a combination of two extreme situations: the failure of an embankment dam, releasing the water from an upstream reservoir; and an open-channel system. We argue that a competition between these two aspects controls the pattern of the breach hydrograph. This competition depends on the relative magnitude of two characteristic time scales and on the dike strength (e.g., through parameter μ). In line with the concept of theoretical exposition depicted by Squazzoni et al. (2020), we describe hereafter a simple conceptual model, which has as sole objective to explore this theoretical explanation. The model was designed to be the simplest possible, while being

able to reproduce the breach hydrograph typology. It is not intended to make predictions for real-world cases, since it is not able to grasp effects of many parameters, such as the main channel and floodplain geometry, the dike material and non-homogeneous features of real-world structures (lining, core), or the propagation of waves in the main channel.

Assuming the breach hydraulics is similar to that of a rectangular broad-crested weir, the mass balance for the main channel reads:

$$A\frac{dh}{dt} = \underbrace{Q_{in}}_{①} \underbrace{-C_b\, b\sqrt{g}\left(h-z\right)^{3/2}}_{②} \underbrace{-C_d\, b_{mc}\sqrt{g}\, h^{3/2}}_{③}, \tag{1}$$

with A the horizontal area of the main channel, h the water level in the main channel, C_b a discharge coefficient representative of the breach discharge capacity ($C_b \sim 0.5$), b the breach width (time-dependent), g the gravity acceleration, z the breach invert level (time-dependent), C_d a discharge coefficient representing the channel downstream boundary condition, and b_{mc} the main channel width. The drain discharge is neglected in the present theoretical approach.

If terms ① and ③ are disregarded in Eq. (1), then the left-hand-side together with the remaining term ② in the right-hand-side correspond to mass balance in case of a dam failure without inflow into the reservoir. In this case, a characteristic time scale of the process is $t_{c1} = A\ (\ g\ w^3\ )^{-1/2}$, since it enables writing the mass balance in the following simple form:

$$\frac{dH}{dT} = -C_b\, B\left(H-Z\right)^{3/2}, \tag{2}$$

with $H = h\ /\ w$ the non-dimensional water level, $B = b\ /\ w$ the non-dimensional breach width, $Z = z\ /\ w$ the non-dimensional breach invert level and $T = t\ /\ t_{c1}$ the non-dimensional time.

Similarly, if term ② is disregarded in Eq. (1), and only terms ① and ③ are kept together with the left-hand-side, the remaining equation describes an open-channel system with inflow and outflow. In this case, another time scale emerges: $t_{c2} = A\ w\ /\ Q_{in}$, which is the residence time in main channel. It enables writing the mass balance in the following non-dimensional form:

$$\frac{dH}{dT'} = 1 - C_d\,\frac{\beta}{\alpha}\,H^{3/2}, \tag{3}$$

where $T' = t\ /\ t_{c2}$ is an alternate non-dimensional time, while $\beta = b_{mc}\ /\ w$ is a geometric parameter and $\alpha = Q_{in}\ /\ (\ g\ w^5\ )^{1/2}$ is a non-dimensional form for the main channel inflow discharge.

Based on these considerations, Eq. (1) may be written in the following non-dimensional form:

$$\frac{dH}{dT} = \alpha - C_b\, B\left(H-Z\right)^{3/2} - C_d\ \beta\ H^{3/2}. \tag{4}$$

Note that parameter α is the ratio of time scales t_{c1} and t_{c2}. It reflects directly the relative magnitude of the time scales related to the "dam failure" component and to the "open-channel system" component of the flow processes of interest here. To solve Eq. (4), we introduced a plausible assumption for the evolution of the non-dimensional breach invert level: $Z(T) = 1 - (\ 1 + T^{-n}\ )^{-1/n}$, as illustrated in Figure S7a in Supplement. The parameter n is

adjusted to mimic the failure of dikes of various strengths (i.e., varying values of parameter μ). The breach width was assumed to grow as a multiple of the breach depth $1 - Z$ (Stage 1) and as a slowly-increasing linear function of time (proxy for Stage 2): $B = m\,(1 - Z) + \gamma T$, with m and γ two model parameters. The evolution of B is illustrated in Figure S7b in Supplement. The representativity of the expressions used for $Z(T)$ and $B(T)$ was confirmed by a comparison with experimental data collected by Rifai et al. (2018; 2021), as presented in Figure S8 in Supplement. Equation (4) was solved with an explicit Runge-Kutta scheme and considering $H = 1$ as initial condition (i.e., water level at the dike crest).

To ensure flow equilibrium before breach development, $C_d\,\beta$ was set to the same value as α, in line with the "adjustment" of the boundary condition performed in the laboratory experiments to ensure that the water level in the main channel reaches the dike crest for each inflow discharge before breaching starts (Rifai et al., 2017; 2018). We set the remaining parameters to plausible values: $C_b = 0.5$, $m = 2$ and $\gamma = 0.05$, whereas parameter α was calculated from the three inflow discharges $Q_{in} = 25$ l/s, 40 l/s and 55 l/s, and from the dike height $w = 0.3$ m.

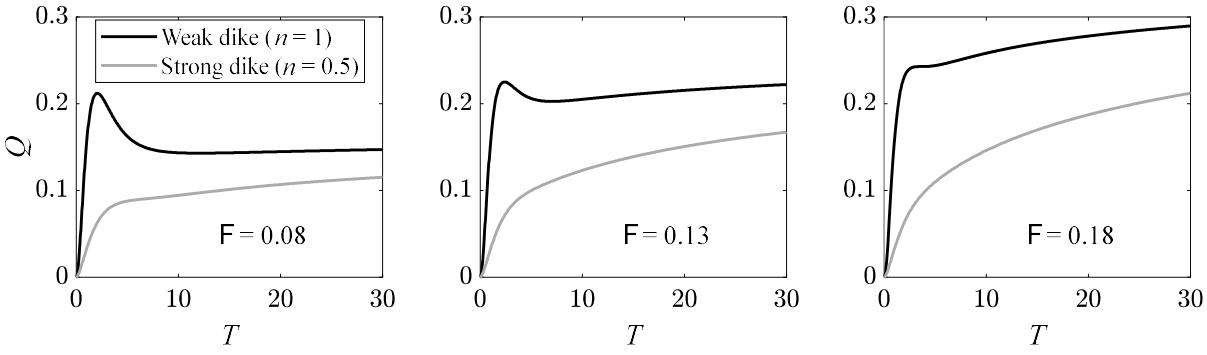


Figure 6. Non-dimensional breach hydrographs $Q(T) = C_b\,B\,(H - Z)^{3/2}$, computed by the conceptual model Eq. (4) for the three inflow discharges considered in the laboratory experiments, corresponding to indicated initial Froude numbers F in the main channel.

Figure 6 displays results for the three inflow discharges Q_{in}, leading to the following initial Froude numbers in the main channel: F = 0.08, F = 0.13 and F = 0.18 (see Table S1 in Supplement). It is amazing to see the similarity between the results of the simple conceptual model based on Eq. (4) and the laboratory observations displayed in Figure 2. The different patterns followed by the breach hydrographs (e.g., occurrence or not of an absolute or relative maximum in the breach discharge) are fairly well reproduced by the conceptual model by changing only the inflow discharge, hence parameter α, and the strength of the dike (through parameter n). We believe that these results illustrate that the typology of breach hydrographs introduced in Section 4.1 is entirely controlled by the relative magnitude of the time scales related to the "dam failure" and the "open-channel system" flow processes, and by the dike strength. Hence, Type A hydrographs correspond to a situation where the reservoir feature is prevalent whilst Type C hydrographs correspond to an open-channel system. For Type B hydrographs, both influences compete.

5. Conclusion

This paper examined the influence of dike geometry on the breach hydrograph and breach widening in the case of failure of homogeneous non-cohesive fluvial dikes due to overtopping at laboratory scale. The dike channel-side slope S_u, floodplain-side slope S_d and crest length L_k were varied systematically. The normalized unit dike volume μ was found to capture successfully the overall effect of these parameters.

The influence of dike geometry differs between Stage 1 (fast erosion) and Stage 2 (gradual evolution towards an equilibrium or quasi-equilibrium state) of breach expansion. During Stage 1, the gradient in breach discharge is milder for stronger dikes (i.e., larger μ), and this effect is magnified when the flow in the main channel is faster. Similarly, the breach expands more gently in both up- and downstream directions in the case of a stronger dike. In contrast, during Stage 2, the rate of breach expansion and outflow discharge are less influenced by the dike geometric parameters.

As the observed breach hydrographs follow three distinct patterns, we introduced a typology of breach hydrographs based on the occurrence or not of an absolute or relative maximum in the breach discharge. Such a maximum occurs mainly for relatively weak dikes (lower values of μ) and a relatively slow flow in the main channel. Our classifications based on the Froude number in the main channel and on the dike strength (through parameter μ) seems to agree with previous studies. Based on a simple conceptual model, we linked the breach discharge typology to the relative magnitude of time scales related to “dam failure” and “open-channel system” flow processes, and to the dike strength. Non-dimensional parameters that emerge from this model may prove useful to normalize breach hydrographs for specific stages of breach expansion. For quantitative predictions of real-world cases, more sophisticated models must be used (e.g., Elalfy et al., 2018; Wu, 2013).

In our tests, when we varied the channel-side slope of the dike, while keeping the inflow discharge unchanged in the main channel, we actually modified the flow velocity, Froude number and hence the momentum in the main channel. In the future, the specific effect of varying the channel-side slope of the dike could be isolated by means of experimental tests designed to preserve the Froude number in the main channel by adjusting the inlet discharge when the channel-side slope of the dike is varied.

Limitations of this study include a relatively narrow range of investigated Froude numbers in the main channel ($0.07 < \mathsf{F} < 0.2$) and normalized unit dike volume (mostly $0.8 \leq \mu \leq 1.2$). The influence of main channel roughness and bottom erodibility was also not investigated, although the latter may play an important role in the breaching process. Though our experiments used comparatively larger dikes (0.30 m in height) than in most previous research (0.10 to 0.20 m in height), scale effects may nonetheless affect the laboratory observations. Therefore, larger scale studies are needed, as well as laboratory experiments aiming at grasping the effects of other parameters such as non-homogeneous features of real-world dikes (lining, core).

Acknowledgement

The authors gratefully acknowledge B. Boucher, B. Coroenne and N. Rousseau for their contribution to the experimental tests. They also express their gratitude to J. Walder and two anonymous Reviewers, whose detailed analysis and insightful comments have greatly improved this paper.

Supporting Information

The supplement related to this article, including texts, tables and figures, is available in the "Supplemental Material" file.

Data used in this article is available in the "Dataset" file (.h5 format).

References

Aerts, J. C., Botzen, W. J., Clarke, K. C., Cutter, S. L., Hall, J. W., Merz, B., ... & Kunreuther, H. (2018). Integrating human behaviour dynamics into flood disaster risk assessment. *Nature Climate Change*, *8*(3), 193-199.

Amaral, S., Caldeira, L., Viseu, T., & Ferreira, R. M. (2020). Designing experiments to study dam breach hydraulic phenomena. *Journal of Hydraulic Engineering*, *146*(4), 04020014.

ASCE/EWRI Task Committee on Dam/Levee Breaching. (2011). Earthen embankment breaching. *Journal of Hydraulic Engineering*, *137*(12), 1549-1564.

Bento, A. M., Amaral, S., Viseu, T., Cardoso, R., & Ferreira, R. M. (2017). Direct estimate of the breach hydrograph of an overtopped earth dam. *Journal of Hydraulic Engineering*, *143*(6), 06017004.

Bhattarai, P. K., Nakagawa, H, Kawaike, K, & Zhang, H. (2015). Study of breach characteristics and scour pattern for overtopping induced river dyke breach. In *E-roceedings of the 36th IAHR World Congress*.

Charrier, G. (2015). *Etude expérimentale des ruptures de digues fluviales par surverse*. PhD thesis, University of Aix-Marseille. (in french)

Coleman, S. E., Andrews, D. P., & Webby, M. G. (2002). Overtopping breaching of noncohesive homogeneous embankments. *Journal of Hydraulic Engineering*, *128*(9), 829-838.

Danka, J., & Zhang, L. M. (2015). Dike failure mechanisms and breaching parameters. Journal of Geotechnical and Geoenvironmental Engineering, *141*(9), 04015039.

Di Baldassarre, G., Kreibich, H., Vorogushyn, S., Aerts, J., Arnbjerg-Nielsen, K., Barendrecht, M., ... & Marchi, B. D. (2018). Hess opinions: An interdisciplinary research agenda to explore the unintended consequences of structural flood protection. *Hydrology and Earth System Sciences*, *22*(11), 5629-5637.

Dou, S. T., Wang, D. W., Yu, M. H., & Liang, Y. J. (2014). Numerical modeling of the lateral widening of levee breach by overtopping in a flume with 180 bend. *Natural Hazards and Earth System Sciences*, *14*(1), 11.

Elalfy, E., Tabrizi, A. A., & Chaudhry, M. H. (2018). Numerical and experimental modeling of levee breach including slumping failure of breach sides. *Journal of Hydraulic Engineering*, *144*(2), 04017066.

El kadi Abderrezzak, K., Die Moran, A., Mosselman, E., Bouchard, J. P., Habersack, H., Lebert, F., & Aelbrecht, D. (2014). A physical, movable-bed model for non-uniform sediment transport, fluvial erosion and bank failure in rivers. *Journal of Hydro-environment Research*, *8*(2), 95-114.

Faeh, R. (2007). Numerical modeling of breach erosion of river embankments. *Journal of Hydraulic Engineering*, *133*(9), 1000-1009.

Islam, M. (2012). *Study on levee breach and successive disasters in low-land through numerical and experimental approaches*. PhD thesis. Nagoya university (Japan).

Jandora, J., & Říha, J. (2008). *The failure of embankment dams due to overtopping*. Vutium.

Kakinuma, T., & Shimizu, Y. (2014). Large-scale experiment and numerical modeling of a riverine levee breach. *Journal of Hydraulic Engineering*, *140*(9), 04014039.

Kakinuma, T., Tobita, D., Yokoyama, H., & Takeda, A. (2013). Levee breach observation at Chiyoda experimental flume. In Advances in River Sediment Research, pages 1013–1020, Kyoto, Japan. ISBN 9781138000629.

LaRocque, L. A., Elkholy, M., Hanif Chaudhry, M., & Imran, J. (2013). Experiments on urban flooding caused by a levee breach. *Journal of Hydraulic Engineering*, *139*(9), 960-973.

Michelazzo, G., Oumeraci, H., & Paris, E. (2015). Laboratory study on 3D flow structures induced by zero-height side weir and implications for 1D modeling. *Journal of Hydraulic Engineering*, *141*(10), 04015023.

Michelazzo, G., Oumeraci, H., & Paris, E. (2018a). New hypothesis for the final equilibrium stage of a river levee breach due to overflow. *Water Resources Research*, *54*(7), 4277-4293.

Michelazzo, G., Paris, E., & Solari, L. (2018b). On the vulnerability of river levees induced by seepage. *Journal of Flood Risk Management*, *11*, S677-S686.

Morris, M. W., Hassan, M. A. A. M., & Vaskinn, K. A. (2007). Breach formation: Field test and laboratory experiments. *Journal of Hydraulic Research*, *45*(sup1), 9-17.

Müller, C., Frank, P. J., & Hager, W. H. (2016). Dyke overtopping: effects of shape and headwater elevation. *Journal of Hydraulic Research*, *54*(4), 410-422.

Orlandini, S., Moretti, G., & Albertson, J. D. (2015). Evidence of an emerging levee failure mechanism causing disastrous floods in I taly. *Water Resources Research*, *51*(10), 7995-8011.

Özer, I. E., van Damme, M., & Jonkman, S. N. (2020). Towards an International Levee Performance Database (ILPD) and Its Use for Macro-Scale Analysis of Levee Breaches and Failures. *Water*, *12*(1), 119.

Pickert, G., Weitbrecht, V., & Bieberstein, A. (2011). Breaching of overtopped river embankments controlled by apparent cohesion. *Journal of Hydraulic Research*, *49*(2), 143-156.

Rifai, I., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., El Kadi Abderrezzak, K., & Dewals, B. (2017). Overtopping induced failure of noncohesive, homogeneous fluvial dikes. *Water Resources Research*, *53*(4), 3373-3386.

Rifai, I., El Kadi Abderrezzak, K., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., & Dewals, B. (2018). Floodplain backwater effect on overtopping induced fluvial dike failure. *Water Resources Research*, *54*(11), 9060-9073.

Rifai, I., El Kadi Abderrezzak, K., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., & Dewals, B. (2019). Flow and detailed 3D morphodynamic data from laboratory experiments of fluvial dike breaching. *Scientific Data*, *6*, 53.

Rifai, I., Schmitz, V. Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., Dewals, B. & El Kadi Abderrezzak, K. (2020). Continuous Monitoring of Fluvial Dike Breaching by a Laser Profilometry Technique. *Water Resources Research*, 10.1029/2019WR026941.

Rifai, I., El Kadi Abderrezzak, K., Hager, W. H., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., & Dewals, B. (2021). Apparent cohesion effects on overtopping induced fluvial dike breaching. *Journal of Hydraulic Research*, *59*(1), 75-87.

Sadeghi, S., Hakimzadeh, H., & Babaeian Amini, A. (2020). Experimental Investigation into Outflow Hydrographs of Nonhomogeneous Earth Dam Breaching due to Overtopping. *Journal of Hydraulic Engineering*, *146*(1), 04019049.

Schmocker, L., & Hager, W. H. (2009). Modelling dike breaching due to overtopping. *Journal of Hydraulic Research*, *47*(5), 585-597.

Schmocker, L., & Hager, W. H. (2012). Plane dike-breach due to overtopping: Effects of sediment, dike height and discharge. *Journal of Hydraulic Research*, *50*(6), 576-586.

Schmocker, L., Frank, P. J., & Hager, W. H. (2014). Overtopping dike-breach: Effect of grain size distribution. *Journal of Hydraulic Research*, *52*(4), 559-564.

Squazzoni, F., Polhill, J. G., Edmonds, B., Ahrweiler, P., Antosz, P., Scholz, G., ... & Gilbert, N. (2020). Computational models that matter during a global pandemic outbreak: A call to action. *Journal of Artificial Societies and Social Simulation*, *23*(2).

Van Lanen, H. A., Laaha, G., Kingston, D. G., Gauster, T., Ionita, M., Vidal, J. P., ... & Delus, C. (2016). Hydrology needed to manage droughts: the 2015 European case. *Hydrological Processes*, *30*(17), 3097-3104.

Viero, D. P., D'Alpaos, A., Carniello, L., & Defina, A. (2013). Mathematical modeling of flooding due to river bank failure. *Advances in Water Resources*, *59*, 82-94.

Vorogushyn, S., Merz, B., & Apel, H. (2009). Development of dike fragility curves for piping and micro-instability breach mechanisms. *Natural Hazards and Earth System Sciences (NHESS)*, *9*(4), 1383-1401.

Vorogushyn, S., Merz, B., Lindenschmidt, K. E., & Apel, H. (2010). A new methodology for flood hazard assessment considering dike breaches. *Water resources research*, *46, W08541*.

Ward, P. J., Jongman, B., Aerts, J. C., Bates, P. D., Botzen, W. J., Loaiza, A. D., ... & Winsemius, H. C. (2017). A global framework for future costs and benefits of river-flood protection in urban areas. Nature climate change, *7*(9), 642-646.

Ward, P. J., de Ruiter, M. C., Mård, J., Schröter, K., Van Loon, A., Veldkamp, T., ... & Capewell, L. (2020). The need to integrate flood and drought disaster risk reduction strategies. *Water Security*, *11*, 100070.

Wei, H., Yu, M., Wang, D., & Li, Y. (2016). Overtopping breaching of river levees constructed with cohesive sediments. *Natural Hazards and Earth system sciences*, *16*(7), 1541-1551.

Wu, S., Yu, M., Wei, H., Liang, Y., & Zeng, J. (2018). Non-symmetrical levee breaching processes in a channel bend due to overtopping. *International Journal of Sediment Research*, *33*(2), 208-215.

Wu, W. (2013). Simplified physically based model of earthen embankment breaching. *Journal of Hydraulic Engineering*, *139*(8), 837-851.

Yu, M. H., Wei, H. Y., Liang, Y. J., & Zhao, Y. (2013). Investigation of non-cohesive levee breach by overtopping flow. *Journal of Hydrodynamics, Ser. B*, *25*(4), 572-579.

Zhang, L., Peng, M., Chang, D., & Xu, Y. (2016). Dam failure mechanisms and risk assessment. John Wiley & Sons.

Supplement to "Overtopping-Induced Failure of Non–Cohesive Homogeneous Fluvial Dikes: Effect of Dike Geometry on Breach Discharge and Widening"

Run ID	Conf. ID	Target Q_{in} (l/s)	Actual Q_{in} (l/s)	F (-)	S_u (-)	S_d (-)	L_k (m)	L_k / w	μ (-)
1	A	25	24.6	0.082	2	2	0.15	0.5	1
1-R	A	25	24.5	0.081	2	2	0.15	0.5	1
2	B	25	25.1	0.071	1.5	2	0.15	0.5	0.9
3	C	25	27.0	0.090	2	1.5	0.15	0.5	0.9
4	D	25	25.3	0.084	2	2.5	0.15	0.5	1.1
5	E	25	26.9	0.089	2	3	0.15	0.5	1.2
6	F	25	24.9	0.083	2	2	0	0	0.8
7	G	25	24.5	0.081	2	2	0.3	1	1.2
8	A	40	40.3	0.134	2	2	0.15	0.5	1
8-R	A	40	40.7	0.135	2	2	0.15	0.5	1
9	B	40	40.8	0.116	1.5	2	0.15	0.5	0.9
10	C	40	40.9	0.136	2	1.5	0.15	0.5	0.9
11	D	40	40.7	0.135	2	2.5	0.15	0.5	1.1
12	E	40	44.0	0.146	2	3	0.15	0.5	1.2
13	F	40	40.9	0.136	2	2	0	0	0.8
13-R	F	40	41.6	0.138	2	2	0	0	0.8
14	G	40	43.4	0.144	2	2	0.3	1	1.2
15	A	55	56.1	0.186	2	2	0.15	0.5	1
16	B	55	55.8	0.159	1.5	2	0.15	0.5	0.9
17	C	55	55.1	0.183	2	1.5	0.15	0.5	0.9
17-R	C	55	56.6	0.188	2	1.5	0.15	0.5	0.9
18	D	55	55.7	0.185	2	2.5	0.15	0.5	1.1
19	E	55	56.5	0.187	2	3	0.15	0.5	1.2
20	F	55	56.0	0.186	2	2	0	0	0.8
21	G	55	56.6	0.188	2	2	0.3	1	1.2
22	H	55	59.0	0.196	2	2	0.6	2	1.6

Table S1. List of experiments Q_{in} = inflow discharge in the main channel; F = initial Froude number in the main channel; 1:S_u = channel-side dike slope; 1:S_d = floodplain-side dike slope; L_k = dike crest length; w = dike height; μ = standardized dike volume per unit width.

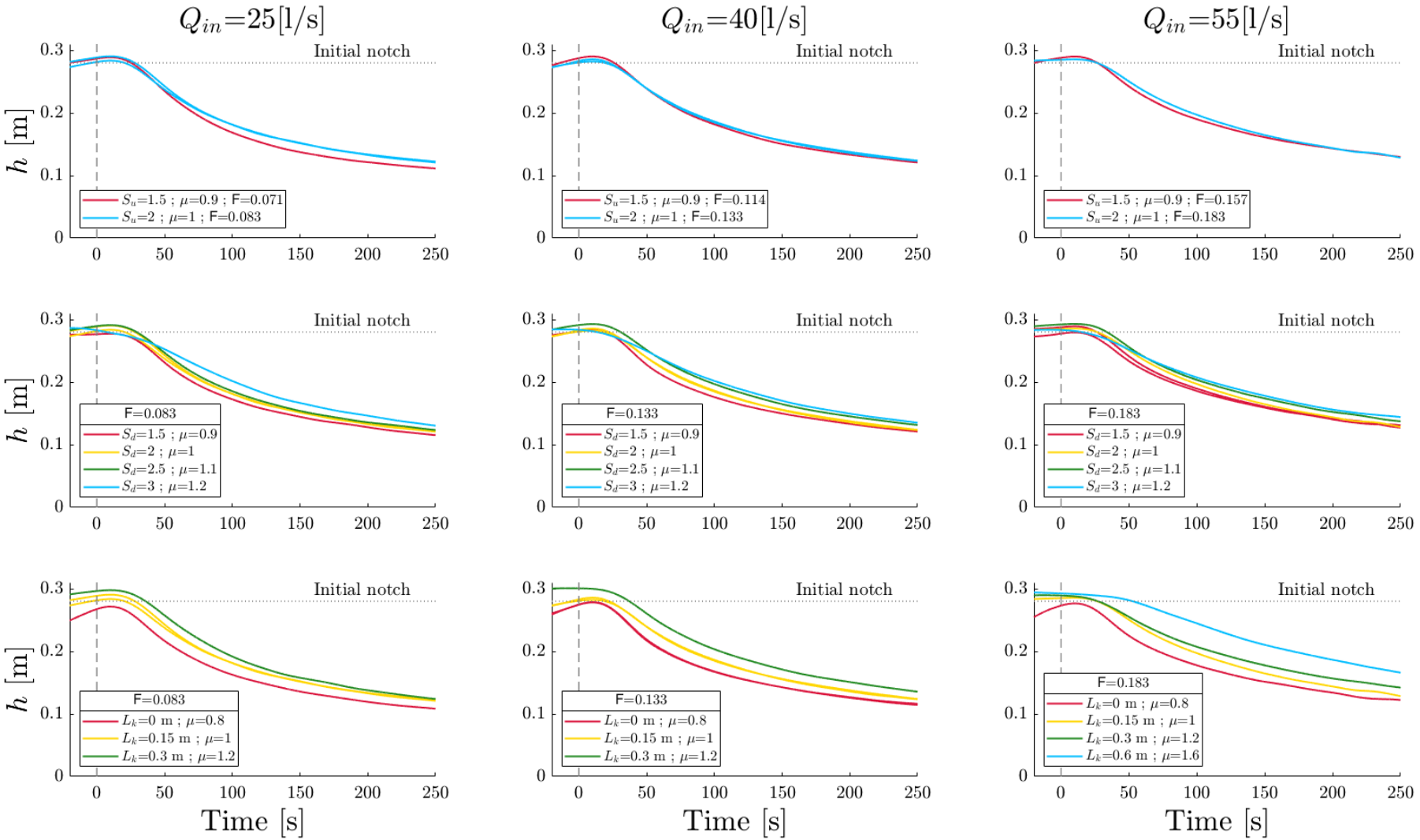


Figure S1. Evolution of the water level in the main channel for all tested geometric configurations and inflow discharges. The represented water level is a weighted-average of the measurements at gauges G1, G2 and G3 (Figure 1a), as detailed in Section 3.1 of Rifai et al. (2017). In each plot, parameters not specified in the legend are set as in the reference Configuration A (S_u = 2, S_d = 2, L_k=0.15m). Origin of time axis corresponds to the start of Stage 1. Curves of identical colour in a same panel refer to repeated tests.

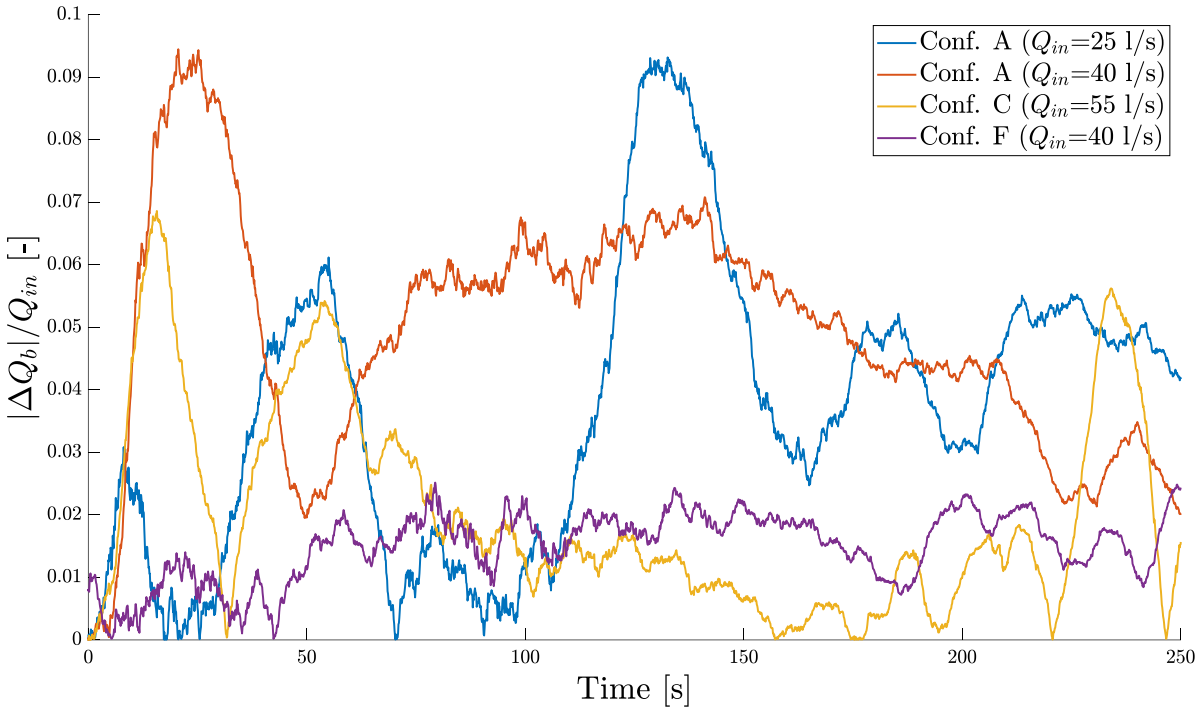


Figure S2. Difference in the breach discharge Q_b between repeated tests, normalized by the inflow discharge Q_{in}.

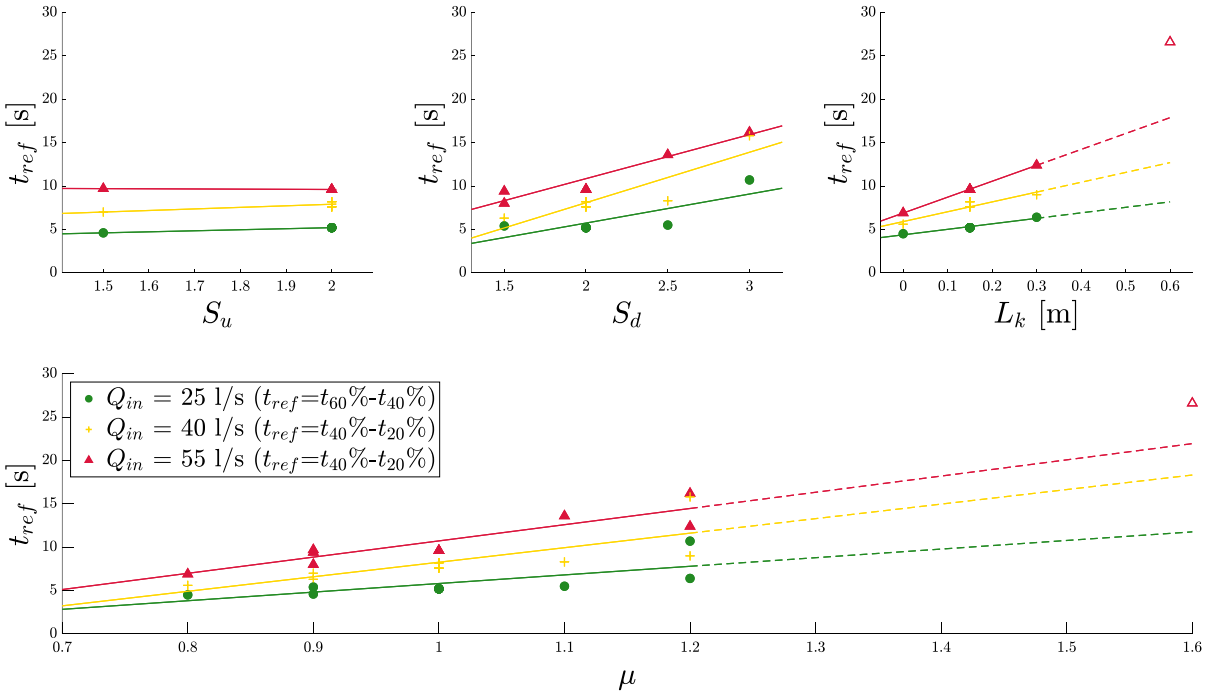


Figure S3. Characteristic rising time of the breach hydrograph as a function of dike geometric parameters. See Text S1 for more information about the reference time definition. Note that Configuration H (L_k = 0.6m, μ = 1.6) was disregarded for the linear interpolations, as the corresponding results are off-scale compared to all other tests. A change in the slope occurs probably for 0.3 m < L_k < 0.6 m (i.e., 1.2 < μ < 1.6); but the present data is not sufficient to properly capture it.

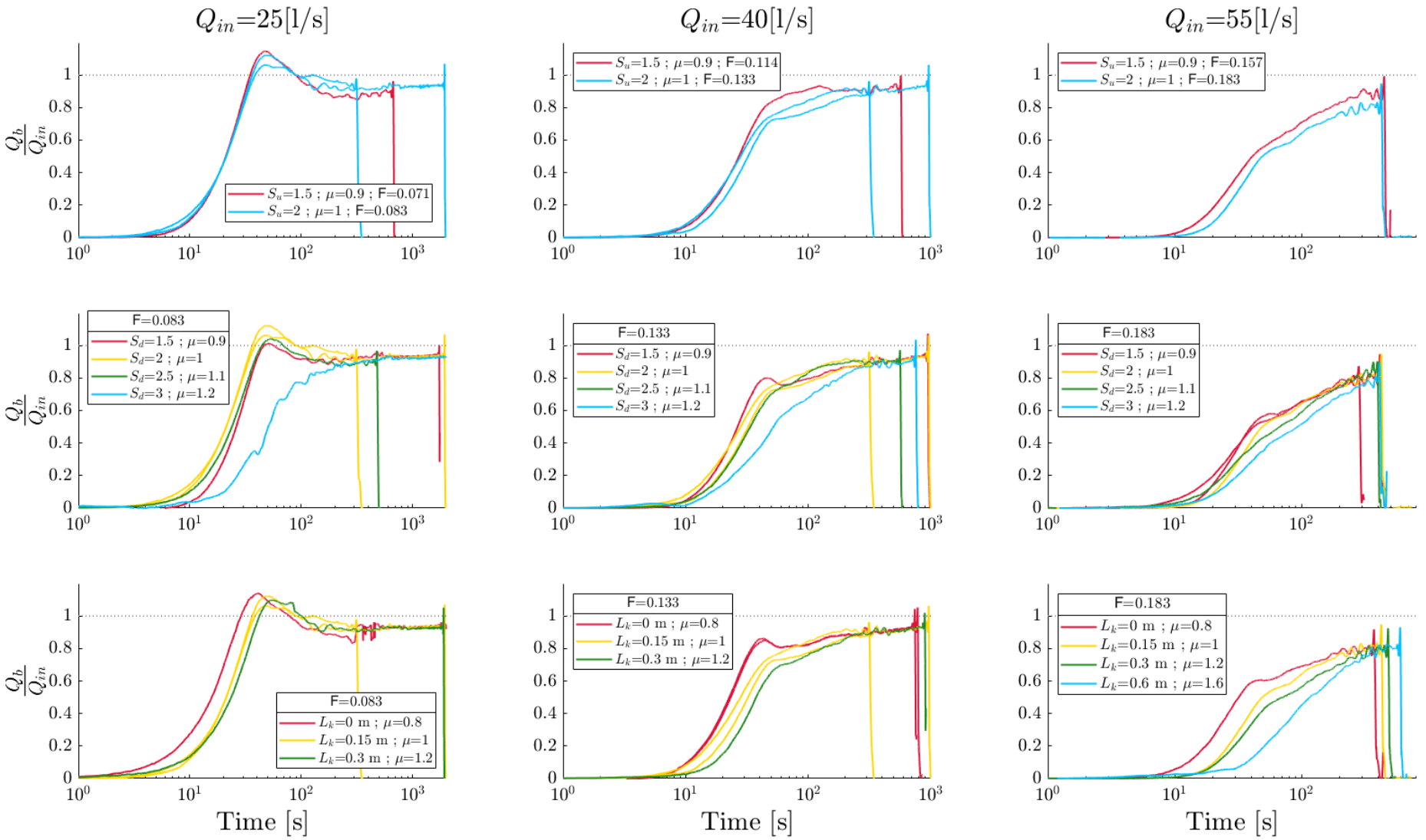


Figure S4. Evolution breach discharge Q_b to inflow discharge Q_{in} ratio. Time is represented on a logarithmic axis, enabling a better appraisal of quasi-equilibrium breach discharge. Sudden drops in the breach discharge correspond to the end of the tests (the pump being stopped). Origin of time axis corresponds to the start of Stage 1. Curves of identical colour in a same panel refer to repeated tests.

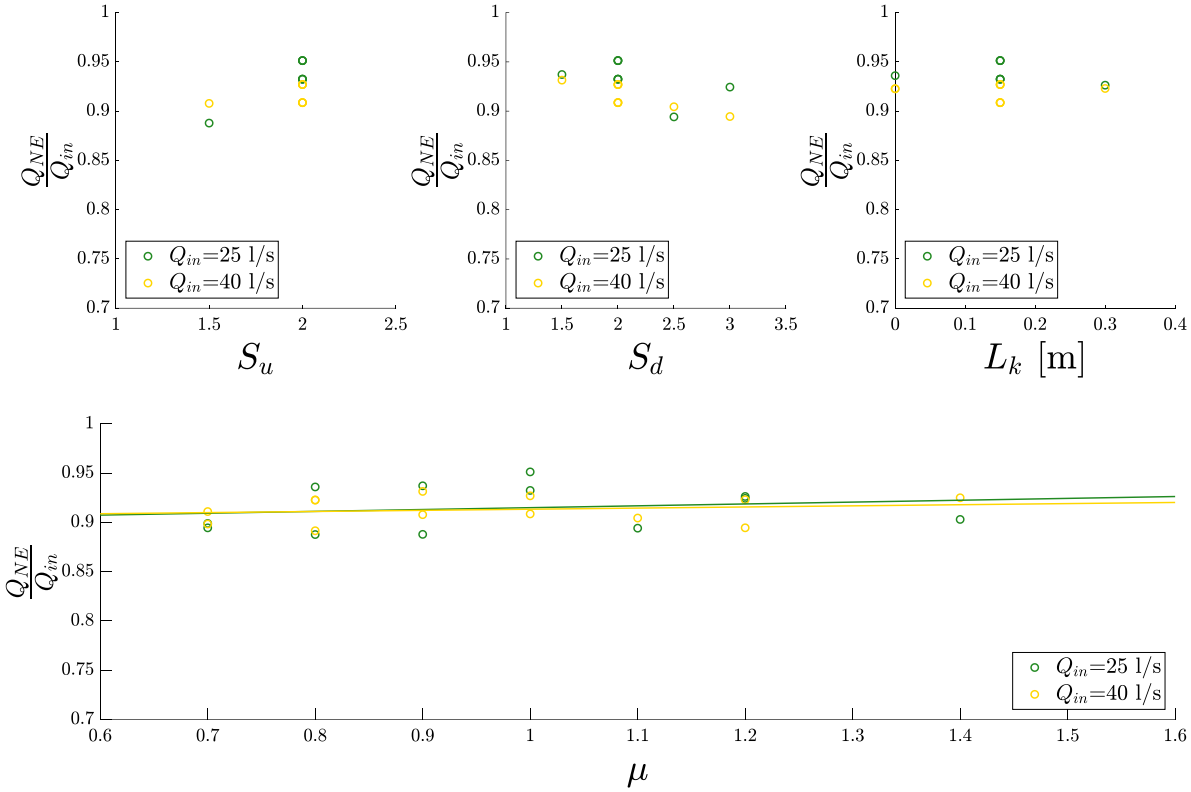


Figure S5. Quasi-equilibrium breach discharge Q_{NE} normalized by the inflow discharge Q_{in} as a function of the geometric parameters for lower and intermediate inflow discharges (Q_{in} = 25 l/s and Q_{in} = 40 l/s). See Text S1 for more information about the reference time definition.

Text S1

The characteristic time t_{ref} displayed in Figure S2 is defined as the time needed for the relative breach discharge (Q_b / Q_{in}) to increase by 20 % during Stage 1. To account for different patterns in the breach hydrographs (e.g. the transition between Stage 1 and Stage 2 occurs earlier for larger values of Q_{in}), the following specific definitions were used:

- $t_{ref} = t_{60\%} - t_{40\%}$ for Q_{in} = 25 l/s,
- $t_{ref} = t_{40\%} - t_{20\%}$ when Q_{in}=40 l/s and 55 l/s

where notation $t_{X\%}$ refers to the time at which Q_b / Q_{in} reaches X %. For each test, Figure S5 shows the values in the breach discharge evolution which were used to estimate the characteristic time t_{ref}, in accordance with the definitions given above.

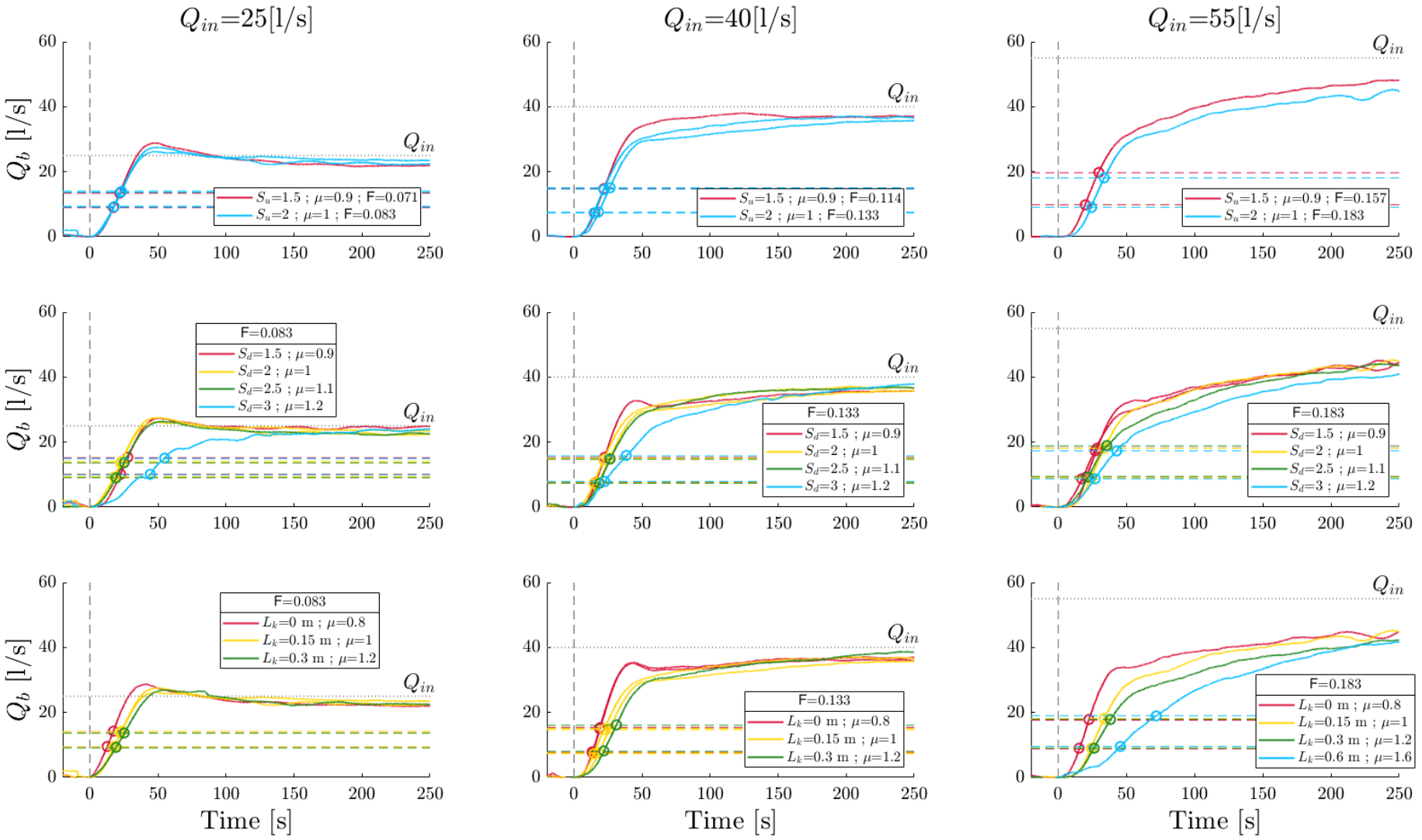
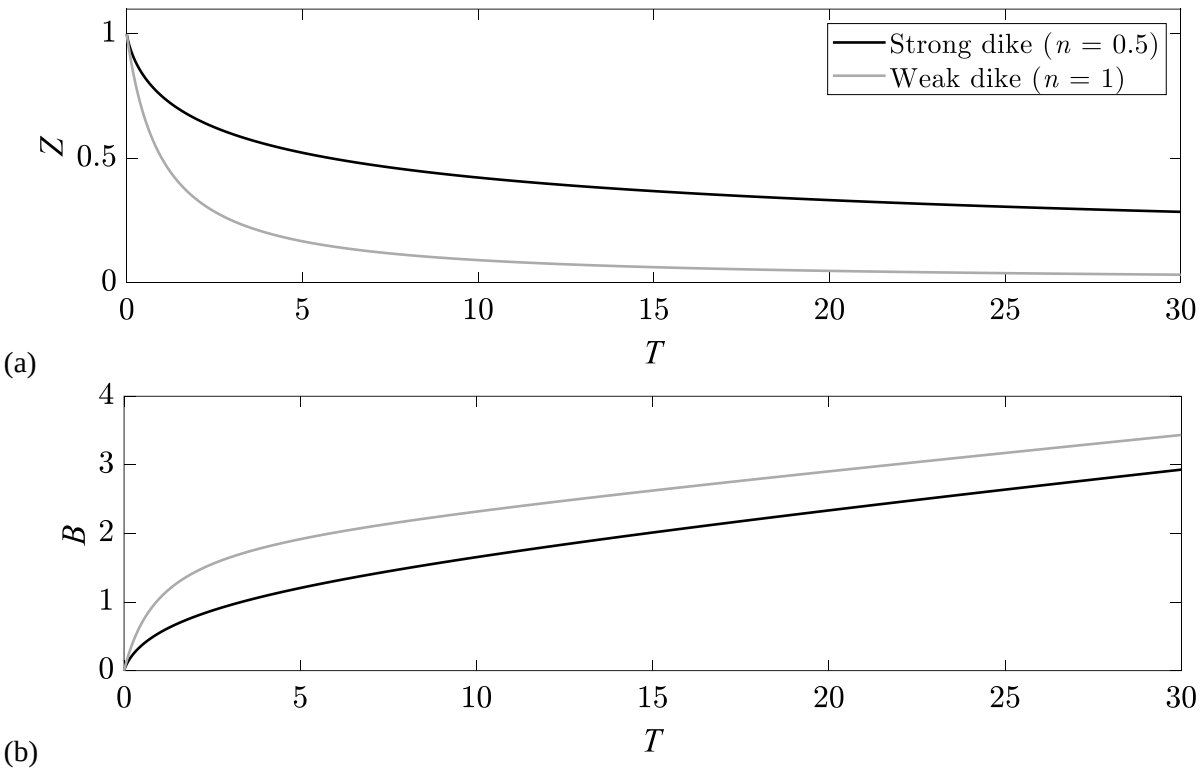


Figure S6. Evolution of the breach discharge Q_b for all tested geometric configurations and inflow discharges. Values of breach discharge which were used to estimate the characteristic time t_{ref} are highlighted by horizontal dashed lines and by circles on the breach hydrographs. Origin of time axis corresponds to the start of Stage 1. Curves of identical color in a same panel refer to repeated tests

		Configuration ID	Q_{in} = 25 [l/s]	Q_{in} = 40 [l/s]	Q_{in} = 55 [l/s]
S_u (-)	1.5	B	A	C	C
	2	A	A	C	C
S_d (-)	1.5	C	A	B	C
	2	A	A	C	C
	2.5	D	A	C	C
	3	E	C	C	C
L_k (m)	0	F	A	B	B
	0.15	A	A	C	C
	0.3	G	A	C	C
	0.6	H	-	-	C
μ	0.8	F	A	B	B
	0.9	B & C	A	B | C	C
	1	A	A	C	C
	1.1	D	A	C	C
	1.2	E & G	A | C	C	C
	1.6	H	-	-	C

Table S2. Type of breach hydrograph obtained for each tested configuration.



(a)

(b)

Figure S7. Evolution of the non-dimensional breach invert level Z (a) and non-dimensional breach width B (b) as a function of the non-dimensional time T in the simplified model.

	Q_{in} = 25 l/s	Q_{in} = 40 l/s	Q_{in} = 55 l/s
t = 0 s			
t = 40 s			
t = 100 s			
t = 250 s			

Table S3. Snapshots of experimental tests in Configuration A (S_u = 2, S_d = 2, L_k = 0.15 m) for three inflow discharges and at four different times.

Chapter 3

Main Channel Width Effects on Overtopping-Induced Non-Cohesive Fluvial Dike Breaching

This chapter corresponds to the journal paper "Main channel width effects on overtopping-induced non-cohesive fluvial dike breaching" by V. Schmitz, I. Rifai, L. Kheloui, S. Erpicum, P. Archambeau, D. Violeau, M. Pirotton, K. El Kadi Abderrezzak, and B. Dewals, published in 2023 in Journal of Hydraulic Research. The PhD candidate conducted the analysis of the experimental data, proposed a new numerical model, generated the associated data and figures, and wrote the manuscript.

Main Channel Width Effects on Overtopping-Induced Non-Cohesive Fluvial Dike Breaching

V. Schmitz[1], I. Rifai[2], L. Kheloui[3], S. Erpicum[1], P. Archambeau[1], D. Violeau[3,4], M. Pirotton[1], K. El kadi Abderrezzak[3,4], and B. Dewals[1]

[1]*Research Group of Hydraulics in Environmental and Civil Engineering (HECE), University of Liège, Liège, Belgium.*

[2] *Hydro-Québec, Montréal, Canada.*

[3] *National Laboratory for Hydraulics and Environment, Research & Development Division, Electricité de France, Chatou, France.*

[4] *Saint-Venant Laboratory for Hydraulics, Chatou, France.*

Corresponding author: Vincent Schmitz (V.Schmitz@uliege.be)

Chapter 3: Influence of the Main Channel Width

ABSTRACT

Laboratory experiments were conducted on the breaching of homogeneous non-cohesive sandy fluvial dikes induced by flow overtopping. Tests were conducted using a main channel, an erodible lateral dike, and a floodplain. The main channel width and Froude number prior to overtopping were systematically varied. Breach discharge was deduced from water level measurements and mass conservation. High-resolution 3D reconstructions of the evolving breach geometry were obtained using a non-intrusive laser profilometry technique. The main channel width and Froude number show significant influence on the breach expansion and hydrograph. Breach hydrographs are divided into three types, depending on the Froude number and a non-dimensional main channel width. An adapted fluvial dike breaching model based on the concept of "effective breach width" is proposed. Using the laboratory data, the computed breach discharge is found extremely satisfactory, although the breach downstream expansion is not accurately reproduced by the model.

1 Introduction

Failure of fluvial dikes (i.e., levees) induced by overtopping flows can lead to devastating floods, causing significant loss of lives and damages to infrastructure and environment (Di Baldassare et al., 2015; Ward et al., 2017). Accurate prediction of the breach expansion and outflow discharge is crucial to achieve a sound assessment of the inundation risk and to design appropriate countermeasures (ASCE/EWRI Task Committee on Dam/Levee Breaching, 2011; LaRocque et al., 2013; Onda et al., 2019). A comprehensive understanding of dike breaching processes and how various factors (e.g., soil, hydraulic loading conditions, dike geometry, overtopping duration) affect the breach expansion is therefore required. In the following, we focus on non-cohesive homogeneous fluvial dikes.

Dike breaching has received renewed attention in recent years, as evidenced by the increasing number of experimental research studies (see review by Rifai et al. (2017) and Schmitz et al. (2021)). Along with the non-symmetrical expansion of the breach in fluvial dike, the complexity of the breach problem is attributed to the multiplicity of involved physical processes, such as surface erosion and side sloping, among others (Powledge et al, 1989; Schmocker et al., 2014). Available experimental works have attempted to clarify the role of some parameters on the breach expansion, including main channel inflow and downstream boundary condition (Michelazzo et al., 2018; Rifai et al., 2017), floodplain backwater (i.e., presence of water in the floodplain or leveed area prior to dike breaching) (Rifai et al., 2018), grain size of the dike material (Islam, 2012; Kakinuma et al., 2013, Rifai et al., 2021), fine sediment inducing apparent cohesion in the dike material (Rifai et al., 2021), and dike cross-sectional shape (Schmitz et al. 2021). However, other factors still require investigation, such as the main channel width.

To our knowledge, no experimental works were conducted on the effect of channel size for fluvial dike breaching. This issue is of particular interest as the channel size may influence the water level evolution following the breaching. In narrower channels, water level responds rapidly to changes in the boundary conditions, in comparison to wider channels, which are characterized by a greater inertia in water level evolution. This effect interacts directly with

the breach flow, and therefore, the breach expansion dynamics. On the other hand, the channel size influences the flow velocity in the near field of the breach with subsequent effects on the breach evolution.

The first aspect, i.e., water level response, was addressed for dam breaching. Wallner (2014) performed laboratory experiments on overtopping induced spatial breaching of dams under different reservoir total volume and shape conditions. Results showed that large storage volumes induce larger breach hydrographs and, for the same storage volume but different reservoir shapes, the highest peak breach discharges were observed for non-linear reservoir characteristics, e.g., a V-shaped reservoir induces a higher peak breach discharge than a rectangular shaped reservoir with the same storage capacity. These results were confirmed experimentally by Frank (2016). Different reservoir sizes were tested by a real time adaptation of the reservoir inflow discharge to water level measurements. Results showed that larger reservoir water volumes lead to larger breach hydrographs. Two tests with equal initial storage surface area, but of different reservoir shapes (rectangular *vs* V-shaped), were compared. With the same surface area, the V-shaped reservoir had smaller overall storage volume. The same peak discharge was observed for both tests and differences between the breach hydrographs appeared only past the peak occurrence. Froehlich (2008) related the final average width of an embankment trapezoidal breach as a function of the reservoir volume at the time of failure. Larger reservoir volumes induced wider breaches and longer breach formation times. Using a data set of 74 cases, Froehlich (2008) showed that reservoir volume was at the forefront of parameters intervening in the breaching process.

The second aspect, i.e., flow velocity field, was briefly analysed by Charrier (2015). Based on water surface velocity measurements, he defined a partition channel width that separates the flow deviated in the breach from that remaining in the main channel. Charrier (2015) showed, that for identical breach width, the higher the channel inflow discharge the smaller the partition width and, for the same inflow discharge, the partition width increased with the breach width. Charrier (2015) assumed that no alteration of the velocity field is induced if the partition width is smaller than the channel width, so that the breach discharge is expected to be the same for wider channels.

The present work is part of an ongoing experimental research program on the breaching of non-cohesive sandy fluvial dikes due to flow overtopping. The main objective of this paper is to assess the effects of main channel width on the breach expansion and discharge. Laboratory observations include time series of water levels in the main channel, time series of flow discharges in the main channel and across the breach, and high-resolution 3D reconstructions of the evolving dike geometry by the Laser Profilometry Technique (LPT). The paper is organised as follows: in Section 2, a description of the experimental setup, measurements, and test program is provided. General results are presented in Section 3 and discussed in Section 4. Conclusions are drawn in Section 5.

2. Laboratory experiments

2.1. Laboratory setup

The experiments were conducted at the National Laboratory of Hydraulics and Environment of the Research and Development (R&D) division of Electricité de France (EDF) using the same experimental set-up as Rifai et al. (2019) (Figure 1a). The experimental setup consisted of a 16.9 m long horizontal, straight main channel of trapezoidal cross-section. Its bed width, w_{mc}, could be varied using movable lateral wooden panels, as depicted in Figure 1b. In this work, four channel widths were considered: 1 m, 1.4 m, 1.8 m, and 2.25 m. A 7 m long trapezoidal dike was built along the right side of the main channel toward a 1×7 m floodplain. The dike material was uniform non-cohesive sand of median diameter $d_{50} = 1$ mm. The dike height, h_d, and crest width, L_k, were set at 0.3 m and 0.1 m, respectively, while both side slopes were fixed to 1:2 (V:H). The main channel and floodplain were at the same level and covered with an impermeable whitewash coating to ensure roughness continuity between the flume, floodplain, and dike (Rifai et al., 2017). A drainage system was placed under the dike to prevent seepage flow. As shown in Figure 1a, a honeycomb straightener was placed upstream of the main channel. A perforated plane was used as regulating weir at the downstream end of the main channel to control water level. Holes were evenly distributed to favour a quasi-uniform velocity distribution over the cross section. Downstream from the perforated plane, a reservoir collected the outflow discharge. Prior to each test, a 2 cm deep and 10 cm wide notch was dug at the dike crest, 2.5 m from its upstream end, to trigger breaching at this specific location.

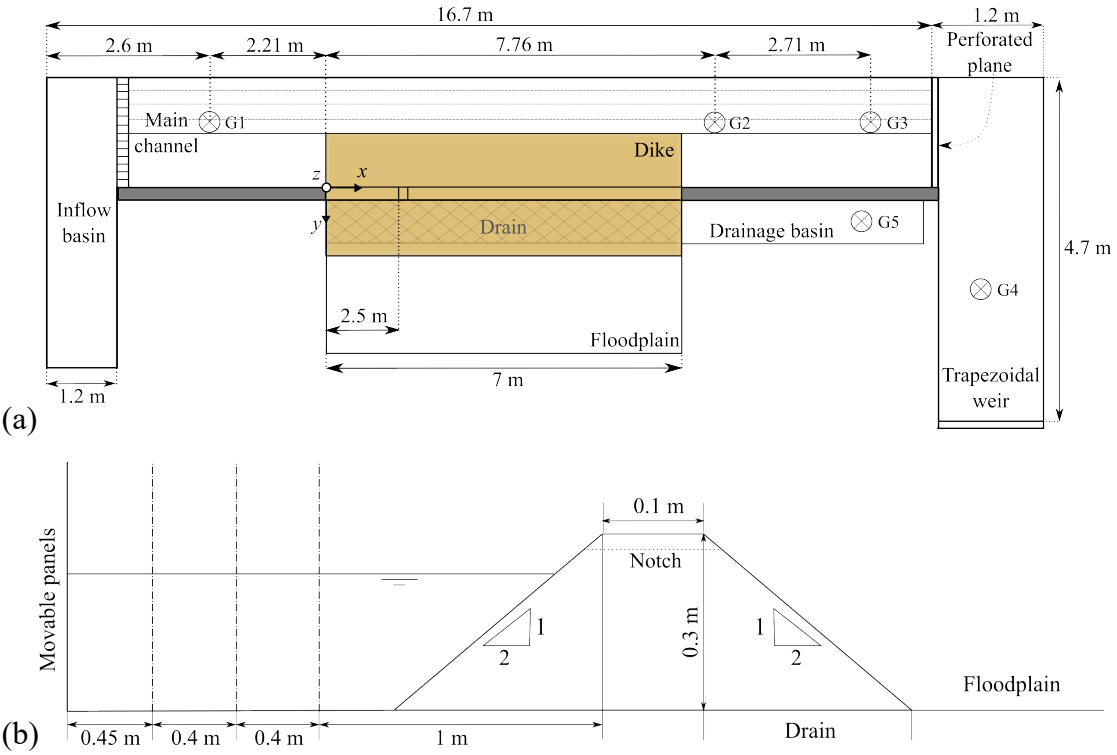


Figure 1. Experimental setup: (a) plane view; (b) dike cross-section, highlighting modular channel width.

2.2. Test program and test procedure

Table 1 presents the overall experimental program with the tested inflow discharges, Q_{in}, and the corresponding inlet initial Froude numbers computed before overtopping, which writes

$$\mathrm{F} = \frac{Q_{in}}{A_h \sqrt{g A_h / w_{mc}}}, \tag{0.1}$$

with A_h the flow cross section and w_{mc} the main channel width. Four different channel widths were considered, with varied inflow discharge. In several previous studies (Rifai et al., 2018; Schmitz et al., 2023), the dike height, h_d, was used as a reference for normalizing geometric quantities. Hence, here also we used h_d as a scaling parameter to define a non-dimensional channel width, which can be interpreted as an aspect ratio of the main channel cross section. In total, 10 configurations were tested. Repeatability and reliability of measurements were verified by performing two or three identical experiments for the selected configurations (i.e., configurations C2, D1, and D2).

For each test, the perforated plane was calibrated so that the water level in the main channel, z_w, reached the dike height for the design inflow discharge. Overtopping started at the initial notch location. Breach discharge was freely released from the floodplain without any storage change nor tailwater effects. Tests were stopped when the breach downstream side closely approached the concrete part of the side opening.

Table 1. Test program. Configurations ID ending with 1, 2 and 3 correspond to lower, intermediate, and higher Froude numbers, respectively.

Conf. ID	Test ID from Rifai et al. (2019)	w_{mc} [m]	Q_{in} (10^{-3} [m^3/s])	w_{mc} / h_d	F
A2	32	1	48	3.33	0.153
B1	33		54		0.108
B2	34	1.4	74	4.67	0.147
B3	35		92		0.184
C1	36		72		0.102
C2	37/38	1.8	98/99	6	0.139/0.140
C3	39		125		0.177
D1	40/41		98/101		0.105/0.108
D2	42/43/44	2.25	139/140/141	7.5	0.149/0.150/0.151
D3	45		160		0.171

2.3. Measurements

Ultrasonic sensors continuously measured the water level (accuracy of ± 1 mm) at three locations in the main channel (G1, G2 and G3), in the outflow tank (G4), and in the drainage tank (G5). An electromagnetic flowmeter measured the inflow discharge, Q_{in} (accuracy of ± 4 %). The outflow discharge (accuracy of ± 3.5 % of Q_{in}) and drainage discharges were deduced via water level variation at G4 and G5, respectively. The breach discharge, Q_b, was determined from mass balance in the main channel (Rifai et al., 2017). The uncertainty affecting the computed breach discharge may be appreciated by comparing breach discharge estimates in repeated tests. Considering the repetitions of Tests D1, C2 and D2 (Fig. 2), it could be estimated that the uncertainty affecting the breach discharge is of the order of 2 % of the inflow discharge on average, while never exceeding 8 %. For more details on the monitoring of water levels and flow discharges, readers may refer to Rifai et al. (2018, 2019, 2021).

The 3D breach evolution was monitored continuously by a non-intrusive profilometry technique consisting of a sweeping laser plan (Rifai et al., 2021). Using a digital camera set on a full HD resolution (1920 × 1080 pixels), the recording was performed at 60 frames per second. The 3D reconstruction algorithm of the dike geometry includes optical distortion and refraction correction modules for submerged dike portions. Further details on the breach geometry reconstruction are given in Rifai et al. (2020).

3. Results

The evolution of the ratio of the breach discharge to the inflow discharge, Q_b / Q_{in}, is presented in Figure 2a for the first 1000 s of each experiment. Figure 2b shows the time-evolution of the breach extremities location at the crest level. The water level in the main channel is displayed in Supplement (Fig. S1).

3.1. Overall breaching dynamics

The breaching process in non-cohesive homogeneous fluvial dikes can be categorized into three stages (Rifai et al., 2017; Michelazzo et al., 2018):

- Stage 0: Overtopping starts at the initial notch and breaching initiates. The breach expands relatively slowly due to low overtopping flow depth and velocity.
- Stage 1: As the overtopping flow depth and velocity increase, erosion intensifies. Both breach deepening and widening are promoted with a shift of the breach centreline toward the channel downstream end. The breach sides encounter repeated sudden collapses. The breach discharge and the water level in the main channel increases and decreases sharply, respectively.
- Stage 2: The water level in the main channel and the breach discharge tend to stabilize. Breach stops deepening while keeping on widening in the downstream direction only, at a slower pace. The expansion is controlled by side slope failures.

In the following analyses, Stage 0 has been systematically discarded due to the overwhelming influence of the initial notch characteristics during this stage.

Both Stages 1 and 2 are clearly visible in Figure 2 for all tests. The non-dimensional breach discharge sharply increases for all curves during Stage 1 and tends to stabilize during Stage 2 (Figure 2a). Similarly, the breach widens quickly during Stage 1 in both upstream and downstream directions (faster toward downstream) while deepening until reaching the main channel bottom (Figure 2b). During Stage 2, the breach expands only toward downstream at a quasi-constant, slower pace. In general, the transition between both stages becomes more distinct for narrower channels and low Froude numbers, e.g., configurations B1 and C1.

3.2. Influence of main channel width during Stage 1

Results show that part of the hydrographs exhibits a global maximum (Figure 2a), which gets more distinct when the initial Froude number in the main channel is low and the main channel is narrow. Conversely, it completely fades out for larger Froude numbers and channel widths. Indeed, with a wider channel and for a given breach width, a larger fraction of the main channel flow is not deviated toward the breach, leading to a smaller non-dimensional breach discharge (Charrier, 2015). Another interesting feature is the initial increase rate in the breach discharge that rises when considering a smaller main channel width. This trend is magnified when the Froude number in the main channel is larger.

During Stage 1, erosion is particularly intense, and the dike widens and deepens sharply (Figure 2b). In all experiments, the downstream widening is significantly larger than the upstream expansion. Though, upstream erosion intensifies when the Froude number in the main channel is reduced. Conversely, the downstream erosion rate rises with the Froude number. This is consistent with earlier observations by Rifai et al. (2017); who showed a substantial increase in the angle between the main channel axis and the breach flow main direction when the Froude number was increased from 0.066 to 0.166. For higher Froude numbers, the flow follows a sharper curve to enter the breach (Charrier, 2015; Rifai et al., 2017). Hence, the load applied by the flow on the downstream breach extremity to modify its direction is higher. For a larger inlet Froude number and flow momentum along the main channel, the velocity component normal to the main channel axis becomes smaller close to the upstream breach extremity. Consequently, the erosion intensity is more limited in this area. This leads to a gradual shift of the breach centreline towards the channel downstream end as the initial Froude number is increased.

Enlarging the main channel width also influences the breaching dynamics. For similar Froude numbers, increasing the main channel width leads to stronger erosion on the breach upstream extremity. The beginning of Stage 1 corresponds to a sudden breach expansion and release of water volume from the main channel into the floodplain. When the main channel is larger, the water level decreases slightly more slowly due to the greater water volume stored in the channel (Fig. S1 in Supplement). Additionally, the breach mean depth, i.e., ratio between breach area and breach width at crest level, evolves in the same way in all tests during Stage 1 (Fig. S2 in Supplement). A larger water volume in the main channel leads to an increased

difference between the main channel water level and the breach bottom. This induces a greater breach discharge, which fosters breach upstream erosion process.

3.3. Influence of main channel width during Stage 2

During Stage 2, the breach discharge in all cases tends to a quasi-stabilized value (around 80% of the inflow discharge), which is hardly affected by the main channel initial Froude number (Figure 2a). This is in good agreement with conclusions drawn by Schmitz et al. (2021).

Here, we additionally highlight the relatively low influence of the main channel width on the breach discharge during Stage 2. During this period, breach deepening almost completely ceases, as shown in Fig. S2 in Supplement. With the increase in breach width, erosion on the breach upstream extremity stops as well (Figure 2b). However, erosion resumes around 800 s for intermediate and high Froude numbers. This phenomenon appears due to the large velocity in the main channel induced by a decrease in the water level (Fig. S1 in Supplement), i.e., flow section, whilst the inflow discharge is kept constant.

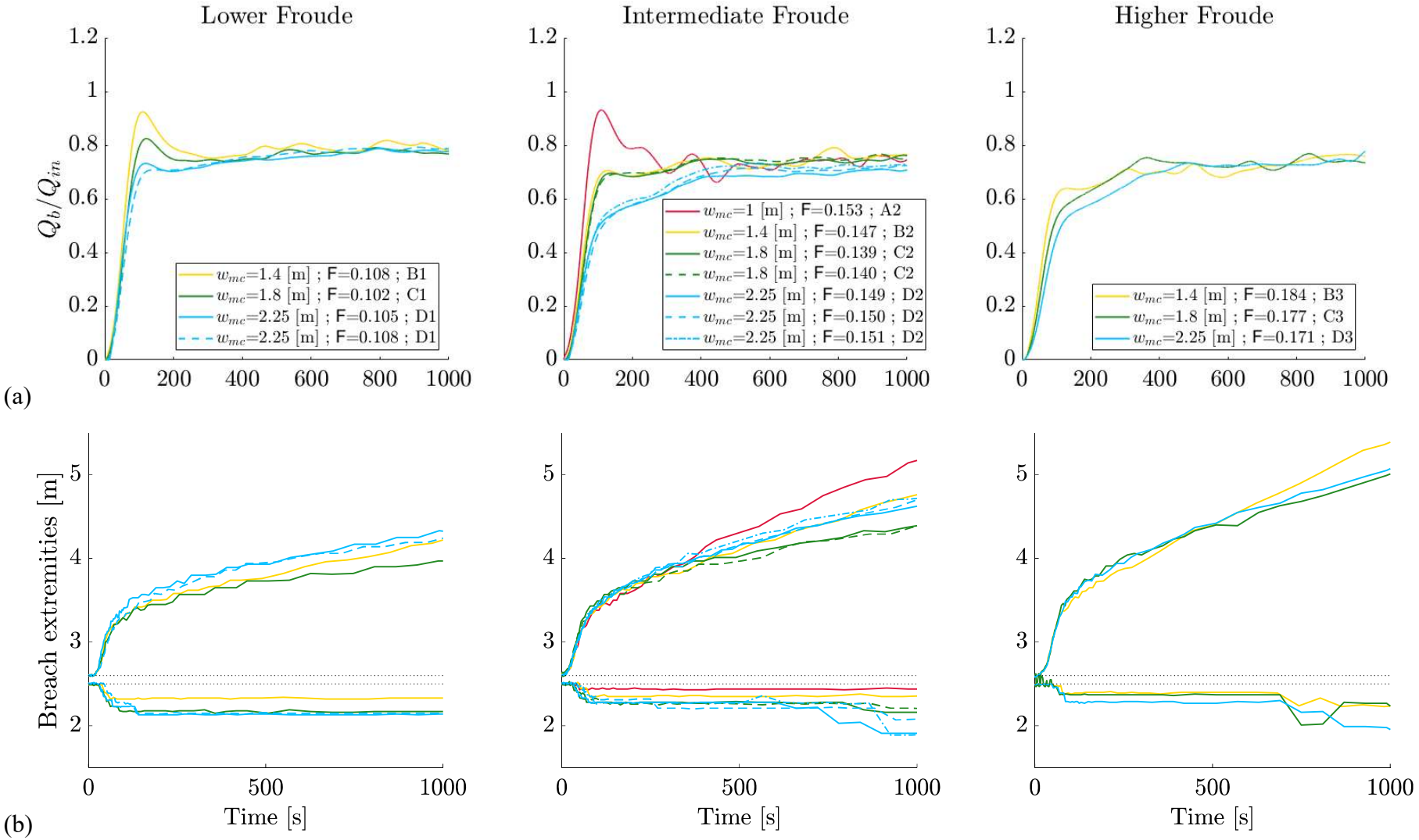


Figure 2. (a) Non-dimensional breach discharge and (b) location of breach extremities (at crest level, along centre line), with initial notch extremities located at 2.5 m and 2.6 m from the dike upstream extremity (dotted lines).

The resulting Froude number in the main channel is significantly increased, which favours erosion in this area. After some time, this lateral erosion reaches the dike centreline and becomes visible in Figure 2b.

The velocity of the flow impacting the breach downstream extremity grows with the Froude number in the main channel. This induces more pronounced breach erosion and downstream expansion. Varying the main channel width has a non-monotonic impact on the breach downstream expansion rate. To understand this feature, two phenomena should be considered (Figure 3). On one hand, the quasi-stabilized non-dimensional breach discharge is similar in all tests (Figure 2a) meaning that, for a given Froude number, the breach discharge, Q_b, increases with the main channel width. For a given breach width, the breach discharge is thus larger, so is the flow velocity through the breach, U_b. This increases erosion intensity and breach expansion rate. On the other hand, when the channel width is reduced, the volume stored in the channel decreases and the water level drops more quickly, as shown in Fig. S1 in Supplement. For a given Froude number, the resulting flow velocity in the main channel, U_{mc}, is significantly increased, leading to a stronger impinging jet on the breach downstream extremity (Charrier, 2015; Rifai et al., 2017).

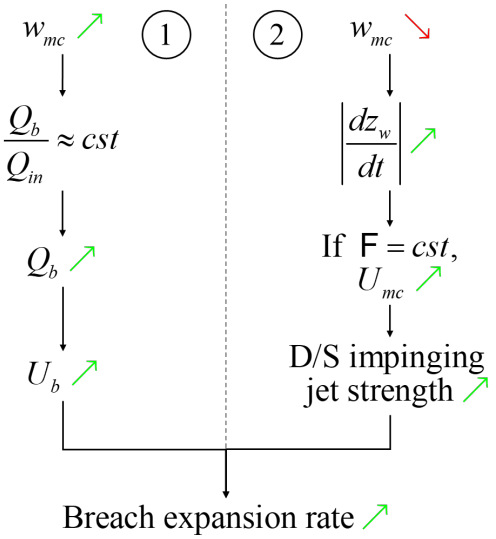


Figure 3. Summary of the impact of variations in the main channel width on the breach expansion rate.

Though, the relative weight of each phenomenon depends on the Froude number in the main channel. For low Froude numbers, Figure 2b shows that erosion is maximum for the widest channel (w_{mc} = 2.25 m, test D1). By contrast, the narrowest channel induces the strongest erosion rate when the Froude number is large (w_{mc} = 1.4 m, test B3). It suggests that increasing the Froude number favours the contribution of the second erosion phenomenon listed above.

4. Discussion

4.1. Classification of hydrographs

Three main hydrograph types emerged, as illustrated in Figure 4. The first type (Type A) presents a global maximum during Stage 1, referring to tests B1, C1 and A2. The second type (Type B) exhibits a local maximum discharge at the end of Stage 1. Experiments D1, B2 and C2 correspond to Type B. All remaining tests belong to Type C, in which no maximum is observed during Stage 1. During Stage 2, all hydrographs tend to a quasi-equilibrium value, irrespectively of their type.

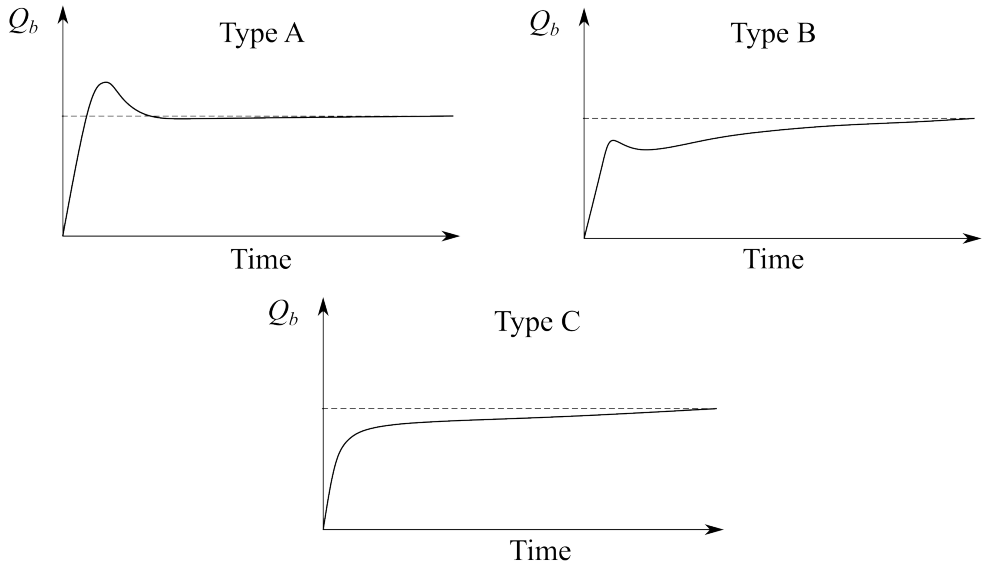


Figure 4. Three main breach hydrograph types (A, B and C).

These three types of hydrographs (A, B and C) were spotted by Schmitz et al. (2021) who suggested that the hydrodynamic process related to fluvial dike failure can be described as a combination of an embankment dam failure, i.e., a sudden water release from a reservoir, and the water flow in a straight open-channel system. Type A hydrographs correspond to a situation where the reservoir feature is prevalent. Conversely, Type C hydrographs are more akin to an open-channel system. Type B lies in-between.

Figure 5 summarizes the partitioning of hydrograph types according to the initial Froude number and the non-dimensional main channel width, w_{mc}/h_d. Eleven experiments led by Rifai et al. (2017), two by Michelazzo et al. (2018) and one by Elalfy et al. (2018) are presented along with tests performed within the present study. It appears that an increase in the Froude number leads to the disappearance of the hydrograph maximum, i.e., transition from Type A to Type C, which agrees with observations by Schmitz et al. (2021). The same trend appears when the non-dimensional channel width is enlarged, but to a much smaller extent. Nonetheless, no clear-cut regions related to specific hydrograph types can be identified, especially for small non-dimensional channel width. It suggests that other parameters affect the hydrograph topology, as for example the drainage system (as highlighted by Michelazzo et al. (2018)) or the water volume stored in the main channel. The channel length, l_{mc}, used in this study is about twice the one used by Rifai et al. (2017). For a given channel width, it corresponds to an increase of about 50% in the volume stored in the main channel.

After breaching, the water level in the main channel drops more slowly when the stored volume is larger, leading to a greater breach discharge during Stage 1. When F ≈ 0.15 and w_{mc}/h_d = 3.33 (Test A2), the hydrograph exhibits an absolute peak discharge during Stage 1 while only a local maximum appears in the corresponding test led by Rifai et al. (2017). In full-scale dike breaching events, the ratio between the main channel width and length is much larger than in laboratory experiments, suggesting that all hydrograph types might not be encountered in practice. However, experimental tests are still extremely useful to validate physically based numerical models, which aim at reproducing physical behaviour of the breaching process regardless of test scale. It is the goal of the following section.

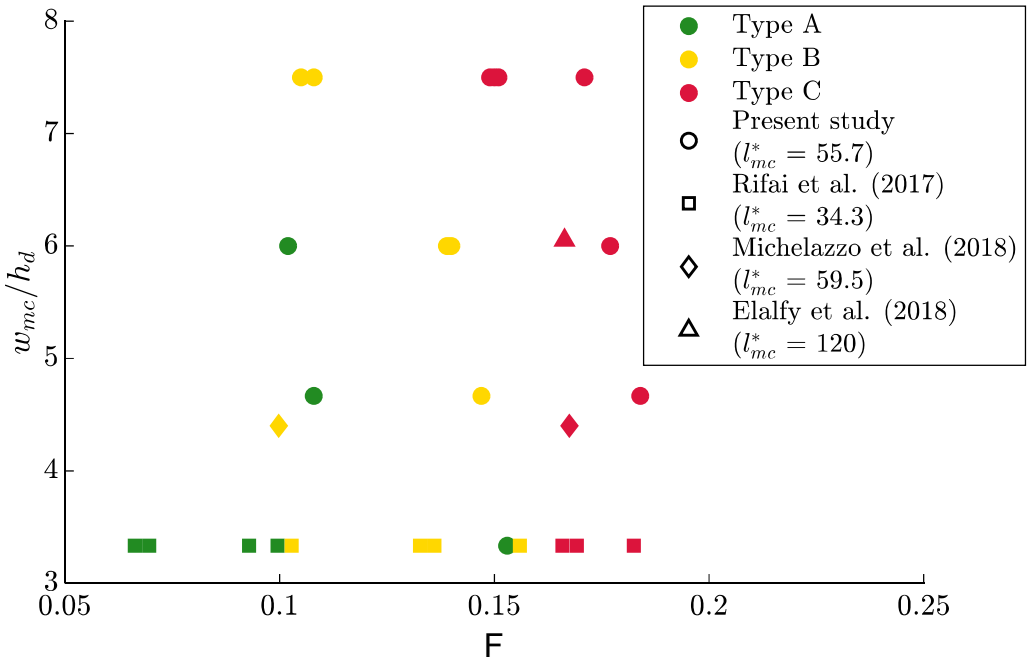


Figure 5. Partitioning of hydrograph types according to initial Froude number in main channel and non-dimensional main channel width. The non-dimensional main channel length $l^*_{mc} = l_{mc} / h_d$ is provided for each test campaign.

4.2. Numerical modelling

The influence of the main channel width on the breach discharge and expansion results from a complex combination of physical phenomena. Here, we adapted an existing semi-analytical physically based dam breaching model to the case of fluvial dike breaching. Our aim is to explore whether the model can capture the influence of the main channel width on the dike breaching dynamics.

The original model is DLBreach developed by Wu (2013), which is composed of three coupled modules: a hydrodynamic module, a sediment transport module, and a dike morphodynamic module. A flow chart of the numerical model is provided in Figs S4, S5 and S6 in Supplement, while Table S1 describes each involved parameter. The flow is assumed uniformly distributed through the entire trapezoidal breach and the breach expansion is symmetric. However, in fluvial configurations, the flow inertia in main channel direction plays an important role in the breaching process. Laboratory experiments (Fig. S3 in Supplement) and detailed 2D computation (Charrier, 2015) show that the flow through the breach is not uniformly distributed along the breach width, but rather concentrated in the most downstream part of the breach. This phenomenon induces a greater water velocity near the downstream breach extremity (Charrier, 2015), which leads to more intense erosion at this location. Conversely, limited erosion appears on the upstream extremity, leading to largely non-symmetrical breach expansion in the case of fluvial dikes.

In our modified implementation of the model of Wu (2013), we introduced the concept of "effective breach width" to describe the fraction of the total breach width that is effectively used to convey water. During Stage 1, flow velocity fields along both breach extremities tend to be similar due to the small breach width, which leads to a rather symmetrical breach expansion. Consequently, we kept the initial breach expansion description proposed by Wu (2013) until the breach bottom reached the channel bottom. Then, asymmetrical breach widening was enforced by considering an effective breach width equal to $b_{eff\%}$ = 50% of the total breach width, which seemed in line with experimental observations (Fig. S3 in Supplement). This leads to higher flow velocity for a similar breach discharge, so that:

$$U_b = \frac{Q_b}{A_{h,eff}}, \tag{0.2}$$

with $A_{h,eff} = f(b_{eff\%})$ the effective breach flow section. Erosion was computed based on this increased velocity and the related effective breach area. Erosion only appeared on the downstream breach extremity and breach expansion speeded up in that direction. No backwater effect is considered in our model.

Figure 6a compares hydrographs obtained experimentally and numerically using the modified model. Fig. S7a in Supplement provides a similar comparison using the original dam breaching model (Wu, 2013). Comparisons of the position of the breach extremities are provided in Figure 6b for the dike breaching model and in Fig. S7b in Supplement for the original dam breaching model, while Fig. S8 in Supplement compares the water levels predicted by both numerical models. Overall, results are greatly improved with the dike

breaching model, especially for the breach hydrograph. The breach expands toward downstream slightly more quickly with the Froude number, which is consistent with experimental data. However, the new model fails to qualitatively reproduce the non-monotonic influence of the channel width on the breach expansion, i.e., enlarging the channel width always leads to a reduction of the breach widening rate during Stage 2. As mentioned in Section 3.3, this non-monotonic behaviour results from the competition between two processes, namely the water volume stored in the main channel and the impinging jet on the breach downstream extremity. The adapted numerical model captures the former aspect while being unable to represent the latter. The intensity of the impinging jet on the breach downstream extremity is expected to rise with a decreasing channel width. Our adapted model currently considers the flow through the breach to be parallel to the breach extremities, which is not consistent with this observation. The insufficient breach expansion rate toward downstream during Stage 1 may also be attributed to this phenomenon.

Besides, the experimentally observed breach width is affected by substantial uncertainties because the actual breach geometry is complex and, for the sake of clarity, only the experimentally breach width at the crest level is reflected in Fig. 2b and Fig. 6b. This cannot reflect all the breach geometric features which influence the flow.

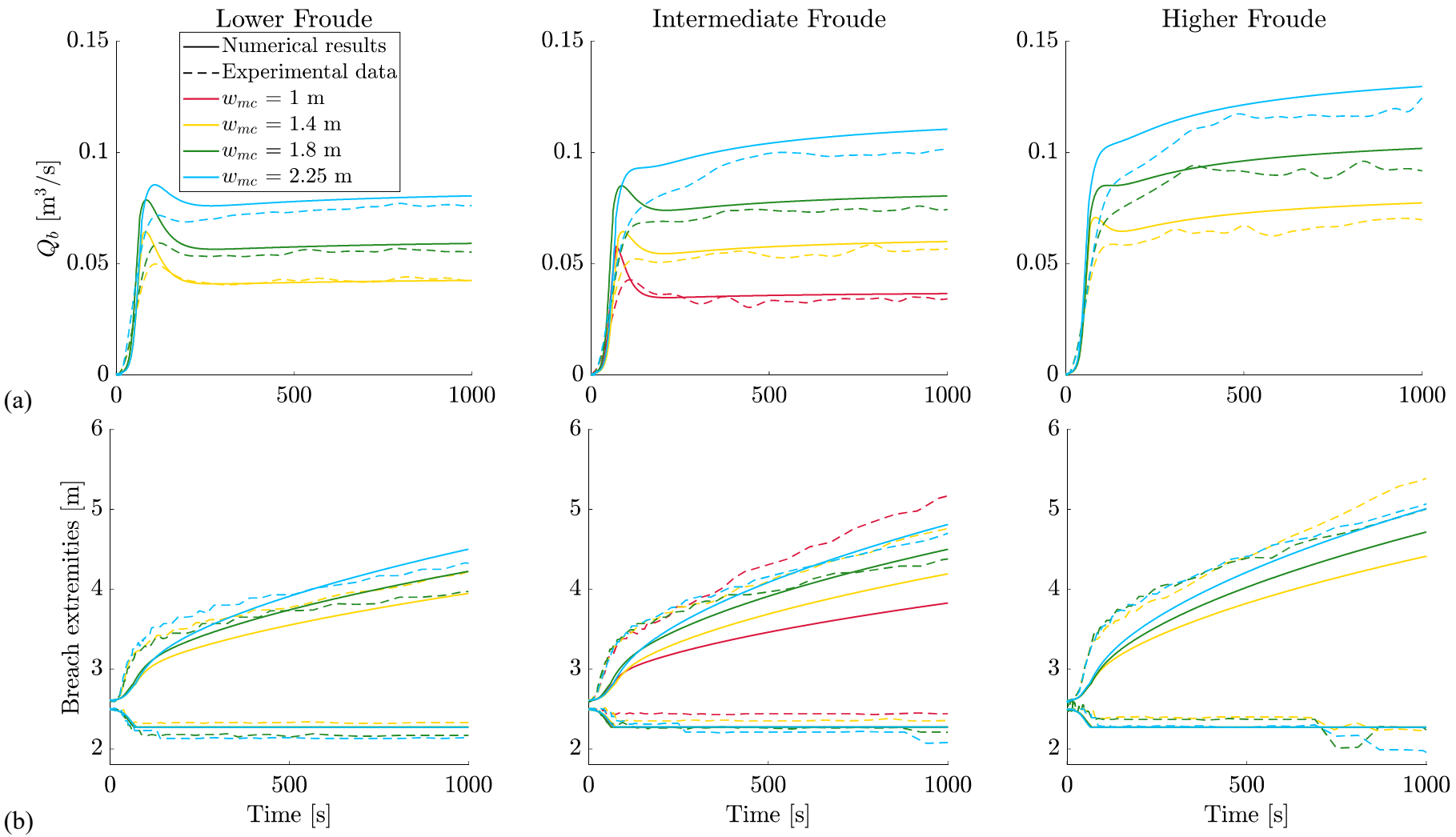


Figure 6. Comparison between results obtained experimentally and numerically with the dike breaching model, i.e., adapted model: (a) breach hydrographs; (b) position of the breach extremities.

5. Conclusions

This paper highlighted the influence of the main channel width on the breach discharge and expansion in a homogeneous and non-cohesive fluvial dike. Laboratory experiments were conducted, while varying systematically the main channel width and inlet Froude number prior to overtopping. Results showed that the breach hydrograph shape and the breach expansion dynamics highly depend on the main channel inlet Froude number and width. During Stage 1, the initial increase in the breach discharge became more gradual and no maximum was reached when the Froude number was increased. Enlarging the main channel has a similar effect as increasing the Froude number, i.e., smoother initial increase in breach discharge and disappearance of the maximum at the end of Stage 1. This influence got particularly pronounced for low Froude numbers. Though, neither the Froude number nor the channel width significantly impacted the quasi-stabilized breach discharge value during Stage 2. Note that other test features may still influence this value, such as the downstream boundary condition (Rifai et al., 2017) or the main channel bed erodibility (Michelazzo et al., 2018).

During Stage 1, the breach quickly deepened and widened, mainly in the downstream direction. Increasing the Froude number and reducing the main channel width reduced the upstream erosion while having almost no influence on the breach downstream expansion. During Stage 2, both breach deepening and upstream widening stopped. The breach expanded toward downstream at a relatively constant pace, which increased with rising Froude number. The relation between the channel width and the breach erosion rate appeared to be non-monotonic, though of limited amplitude. It is expected to be driven by the combination of two physical phenomena related to the amount of water stored in the main channel on one hand, and the velocity of water impacting the downstream extremity of the breach on the other hand. Their relative weight determined the breach expansion rate.

Hydrographs were categorized in three types, according to the classification proposed by Schmitz et al. (2021). Their partitioning according to the Froude number in the main channel and the non-dimensional channel width highlighted that Type A hydrographs mostly appear with low Froude number and narrow channels whilst Type C hydrographs emerged with larger Froude numbers and channel widths. Type B hydrographs were obtained for intermediate configurations. Though, no clear-cut limits between type zones could be identified. Additional parameters should thus be considered to fully describe this classification.

A semi-analytical physically based fluvial dike breaching model corresponding to an adapted version of the model proposed by Wu (2013) was introduced and its capabilities were assessed using the presented experimental data. The main model improvement consisted in the introduction of the concept of “effective breach width”, which lumps the effect of non-uniform flow distribution through the breach. The accuracy of numerical results was significantly improved. Though, the new model failed to capture the non-monotonic relation between the channel width and the breach downstream erosion rate during Stage 2.

A wider range of main channel widths should be tested to better understand the impact on the breaching dynamics. Also, the influence of other parameters should be investigated systematically, e.g., dike material properties, drainage system, channel bottom erodibility, but also scale effect. Indeed, larger water volumes are expected to be stored in real world river channel, leading to much slower water level decrease. This would particularly affect flow dynamics and breach expansion during Stage 2. Finally, a way of improving the predictive capabilities of the model presented in this paper would be to include the impact of the impinging jet on the breach downstream extremity, which is expected to boost erosion rate at this location.

Funding

This work was partially funded by the *Association Nationale de Recherche et de la Technologie* (ANRT) [CIFRE 2015/0015 and CIFRE 2018/1235], the European Regional Development Fund (*Programme Opérationnel Interrégional Rhône-Saône* 2014-2020) and EDF.

Supplemental data

Supplemental data can be accessed from the online version of the paper. All experimental data are available from the following Zenodo depository: https://doi.org/10.5281/zenodo.1477843.

References

ASCE/EWRI Task Committee on Dam/Levee Breaching. (2011). Earthen embankment breaching. *Journal of Hydraulic Engineering*, *137*(12), 1549-1564.

Charrier, G. (2015). Etude expérimentale des ruptures de digues fluviales par surverse (Doctoral dissertation). University of Aix-Marseille. (In French)

Di Baldassare, G., Viglione, A., Carr, G., Kuil, L, Yan, K., Brandimarte, L., & Blöschl, G. (2015). Debates-perspectives on socio-hydrology: Capturing feedbacks between physical and social processes. *Water Resources Research*, *51*(6), 4770–4781.

Elalfy, E., Tabrizi, A. A., & Chaudhry, M. H. (2018). Numerical and experimental modeling of levee breach including slumping failure of breach sides. *Journal of Hydraulic Engineering, 144*(2), 04017066.

Frank, P.-J. (2016). Hydraulics of spatial dike breach (Doctoral dissertation). ETH Zürich, Switzerland.

Froehlich, D. C. (2008). Embankment Dam Breach Parameters and their Uncertainties. *Journal of Hydraulic Engineering,* 134(12):1708-1721

Islam, S. (2012). Study on levee breach and successive disasters in low-land through numerical and experimental approaches (Doctoral dissertation). Nagoya University, Japan.

Kakinuma, T., Tobita, D., Yokoyama, H., & Takeda, A. (2013). Levee breach observation at Chiyoda experimental flume. In S. Fukuoaka, H. Nakagawa, T. Sumi, & H. Zhang (Eds.), *ISRS* 2013. *Proceedings of the 12th International Symposium on River Sedimentation* (pp. 1013–1020). Kyoto, Japan: CRS press.

LaRocque, L. A., Elkholy, M., Hanif Chaudhry, M., & Imran, J. (2013). Experiments on urban flooding caused by a levee breach. *Journal of Hydraulic Engineering*, *139*(9), 960-973.

Michelazzo, G., Oumeraci, H., & Paris, E. (2018). New hypothesis for the final equilibrium stage of a river levee breach due to overflow. *Water Resources Research*, *54*(7), 4277–4293.

Onda, S., Hosoda, T., Jaćimović, N. M., & Kimura, I. (2019). Numerical modelling of simultaneous overtopping and seepage flows with application to dike breaching. *Journal of Hydraulic Research*, *57*(1), 13–25.

Powledge, G. R., Ralston, D. C., Miller, P., Chen, Y. H., Clopper, P. E., & Temple, D. M. (1989). Mechanics of Overflow Erosion on Embankments. II: Hydraulic and Design Considerations. *Journal of Hydraulic Engineering, 115*(8), 1056–1075.

Rifai I., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., El Kadi Abderrezzak, K., & Dewals, B. (2017). Overtopping induced failure of noncohesive, homogeneous fluvial dikes. *Water Resources Research*, *53*(4), 3373–3386.

Rifai, I., El kadi Abderrezzak, K., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., & Dewals, B. (2018). Floodplain backwater effect on overtopping induced fluvial dike failure. *Water Resources Research*, *54*(11), 9060–9073.

Rifai, I., El kadi Abderrezzak, K., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., & Dewals, B. (2019). Flow and detailed 3D morphodynamic data from laboratory experiments of fluvial dike breaching. *Scientific Data*, *6*(1), 53.

Rifai, I., Schmitz, V. Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., Dewals, B. & El Kadi Abderrezzak, K. (2020). Continuous Monitoring of Fluvial Dike Breaching by a Laser Profilometry Technique. *Water Resources Research*, 10.1029/2019WR026941.

Rifai, I., El Kadi Abderrezzak, K., Hager, W. H., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., & Dewals, B. (2021). Apparent cohesion effects on overtopping induced fluvial dike breaching. *Journal of Hydraulic Research*, *59*(1), 75-87.

Schmitz, V., Erpicum, S., El Kadi Abderrezzak, K., Rifai, I., Archambeau, P., Pirotton, M., & Dewals, B. (2021). Overtopping-Induced Failure of Non–Cohesive Homogeneous Fluvial Dikes: Effect of Dike Geometry on Breach Discharge and Widening. *Water Resources Research*, *57*(7), e2021WR029660.

Schmitz, V., Arnst, M., Abderrezzak, K. El kadi, Pirotton, M., Erpicum, S., Archambeau, P., & Dewals, B. (2023). Global sensitivity analysis of a dam breaching model: To which extent is parameter sensitivity case-dependent? Water Resources Research, e2022WR033894.

Schmocker, L., Frank, P.-J., & Hager, W. H. (2014). Overtopping dike-breach: Effect of grain size distribution. *Journal of Hydraulic Research*, *52*(4), 559–564.

Wallner, S. Influence of reservoir Shape and Size on The Flood Wave Caused by Progressive Overtopping Dam Failure (Doctoral dissertation). TU Wien, 2014

Ward, P. J., Jongman, B., Aerts, J. C., Bates, P. D., Botzen, W. J., Loaiza, A. D., ... & Winsemius, H. C. (2017). A global framework for future costs and benefits of river-flood protection in urban areas. *Nature climate change, 7*(9), 642-646.

Wu, W. (2013). Simplified physically based model of earthen embankment breaching. *Journal of Hydraulic Engineering, 139*(8), 837-851.

Supplement to "Main Channel Width Effects on Overtopping-Induced Non-Cohesive Fluvial Dike Breaching"

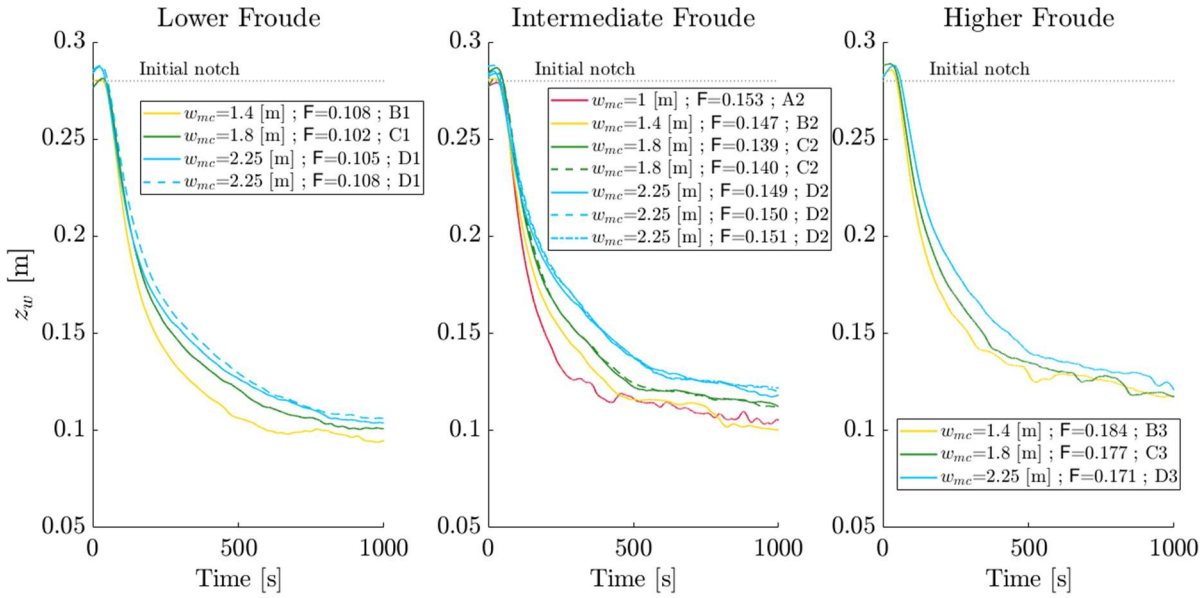


Figure S1. Water level evolution in main channel. Origin of time corresponds to start of Stage 1.

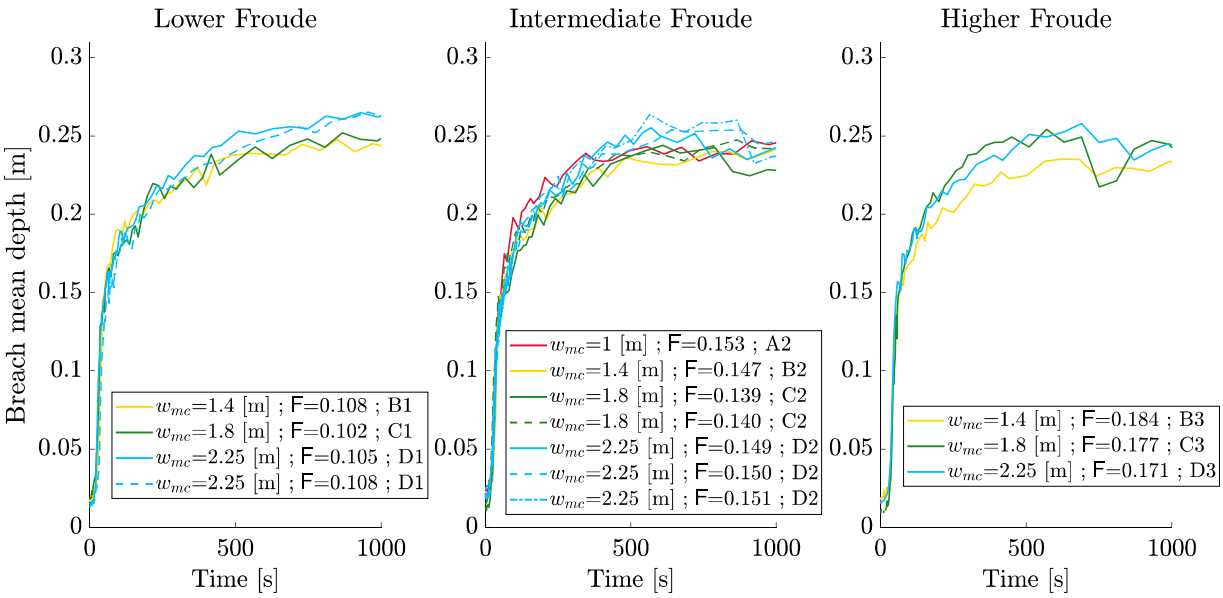
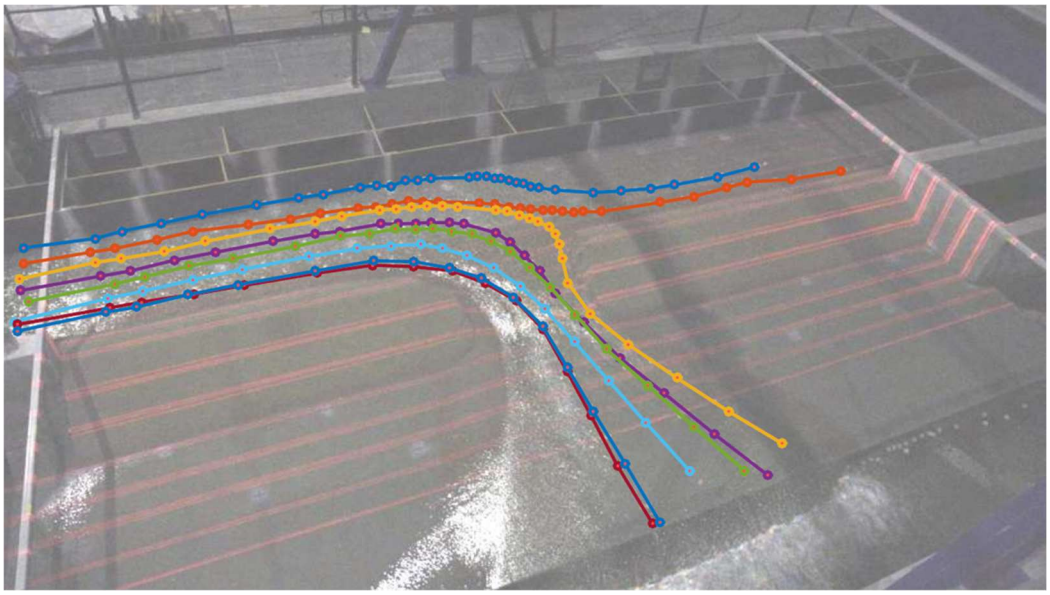


Figure S2. Breach mean depth evolution. The breach mean depth is defined as the ratio between the breach area and the maximum breach width, both computed along the dike crest center line.

(a)



(b)

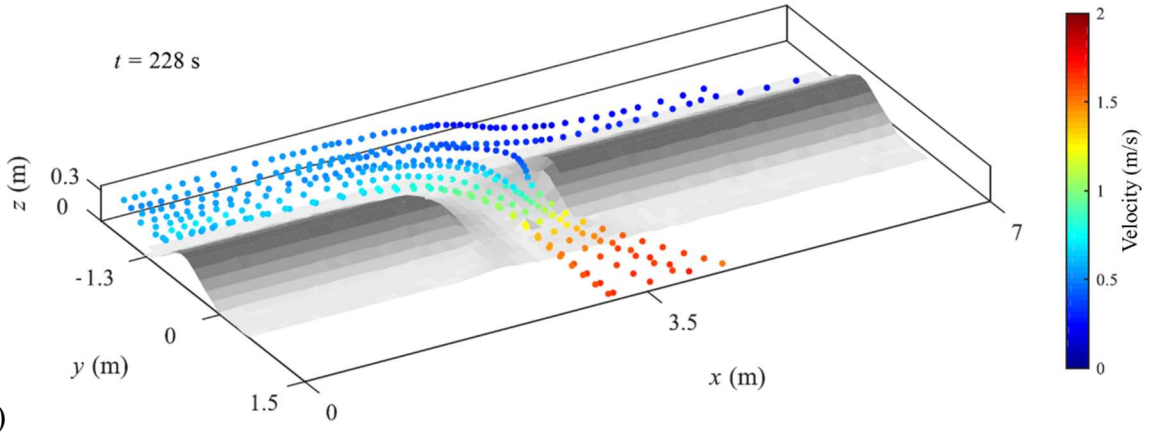


Figure S3. Laboratory Test B2. (a) Float positions and (b) associated velocity 228 s after beginning of breaching obtained using LPT.

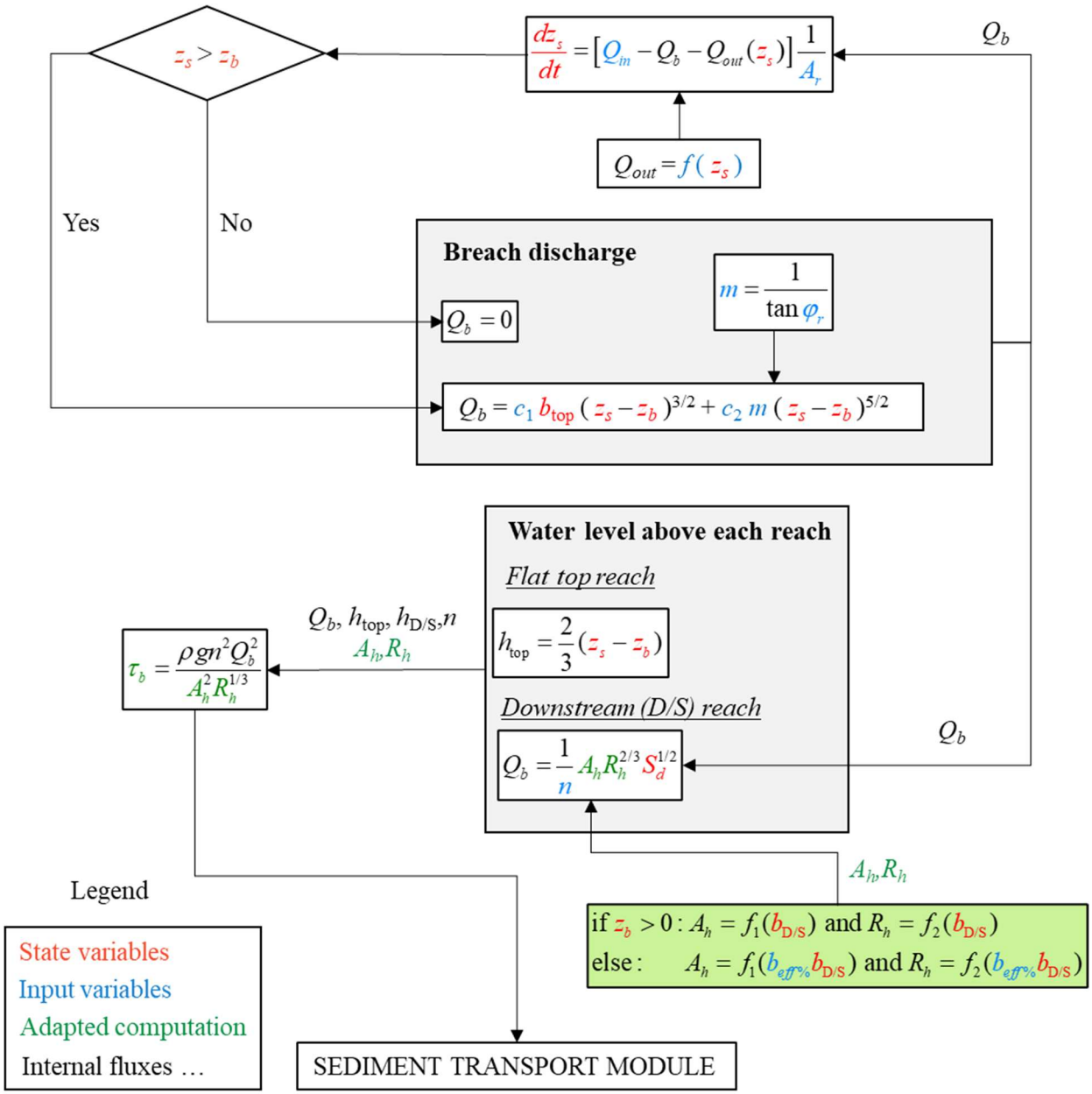


Figure S4. Flow chart of the hydrodynamic module of the numerical model. Green variables are computed using a different approach as the one used in DLBreach (Wu, 2013).

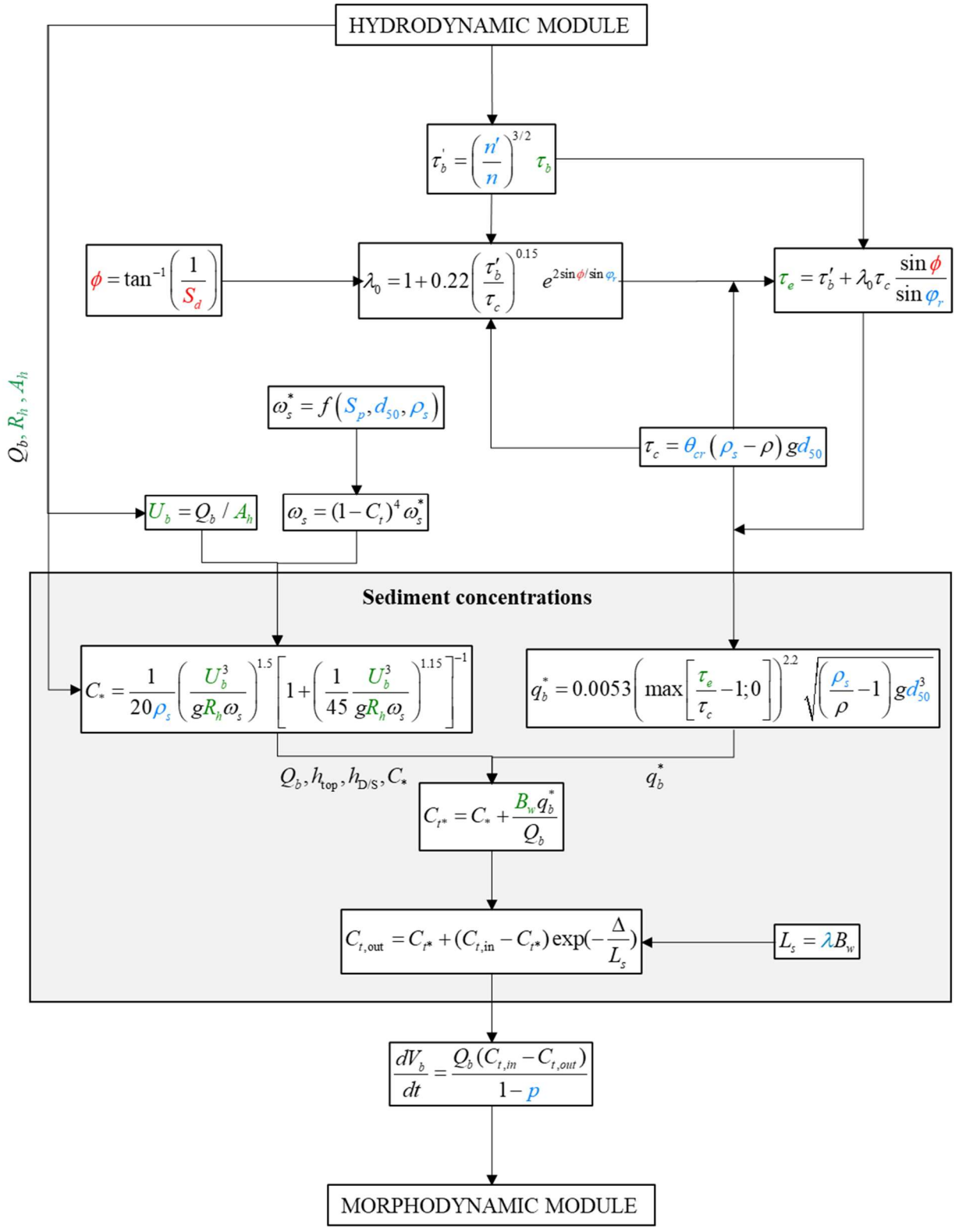


Figure S5. Flow chart of the sediment transport module.

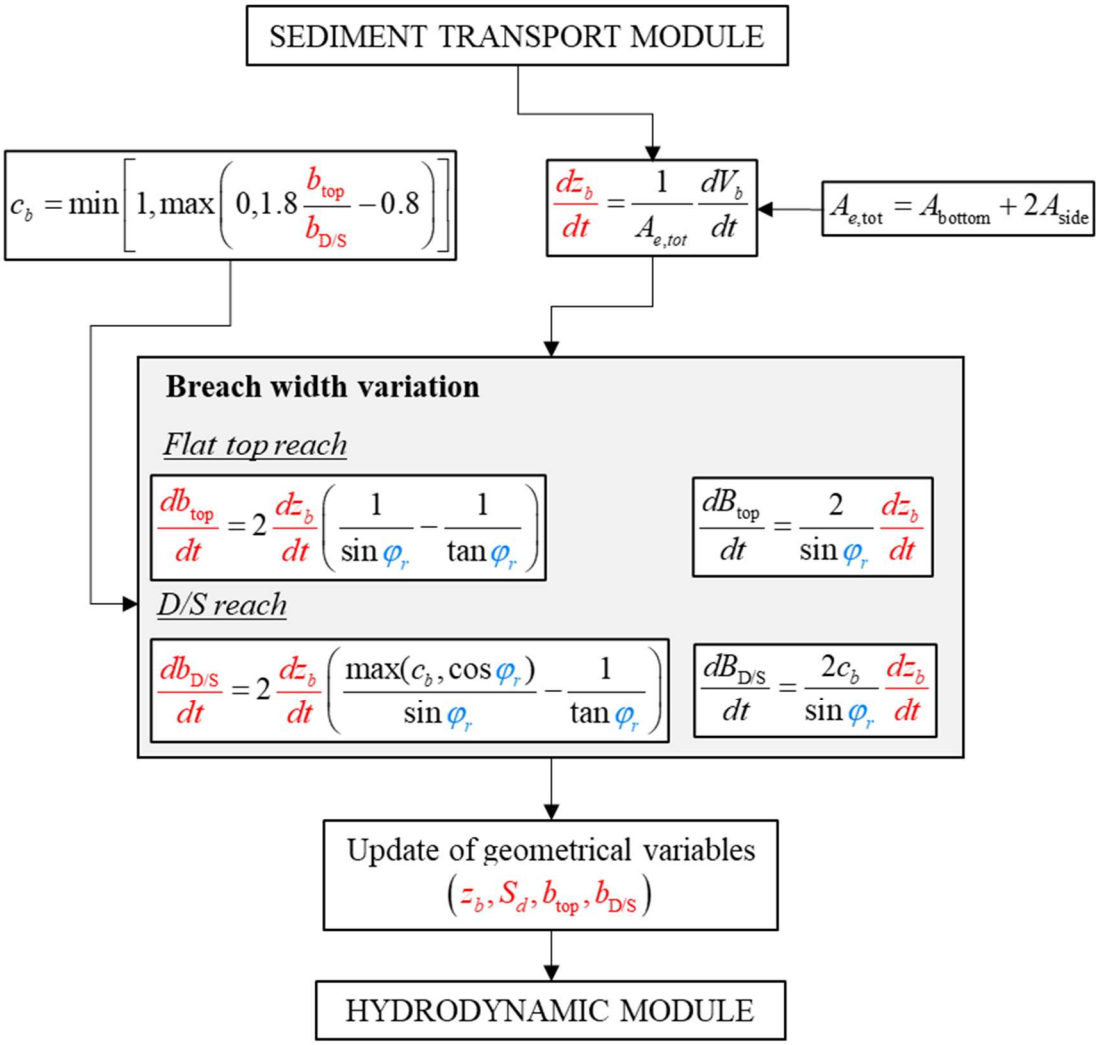


Figure S6. Flow chart of the morphodynamic module.

Supplement to Chapter 3

Table S1. Summary of parameters involved in the numerical model.

	Symbol	Description
State variables	z_s, z_b	Level of the main channel free surface and breach bottom, respectively
	S_d	Downstream slopes of the dam
	$b_{\text{top}}, b_{\text{D/S}}$	Breach bottom width on the flat top reach and downstream reach, respectively
Input variables	Q_{in}, A_r	Inflow discharge and free surface area of the main channel
	$\rho_s, d_{50}, \varphi_r$	Sediment density, median grain size and repose angle
	c_1, c_2	Weir efficiency coefficients
	n, n'	n and n' total and effective Manning's coefficients, respectively
	$b_{eff\%}$	Effective fraction of the total breach width
	S_p	Corey shape factor
	θ_{cr}	Critical Shields parameter
	λ	Empirical coefficient involved in mixing length computation
	p	Dam material porosity
Output variables	Q_b, Q_{out}	Breach and outflow discharges, respectively
	A_h, R_h	Breach wetted area and hydraulic radius, respectively
	τ_b, τ_e, τ_c	Breach shear stress, effective breach shear stress and critical shear stress
	U_b, ω_s^*	Water velocity through the breach and sediment settling velocity
	C_*, q_b^*	Suspended load and bed load sediment transport capacities
	B_w	Width of the water free surface over the effective breach section
	$C_{t^*}, C_{t,in}, C_{t,out}$	Sediment concentration at equilibrium, at the reach inlet and at the reach outlet
	Δ, L_s	Reach length and mixing length
	$V_b, A_{e,tot}$	Eroded sediment volume and erodible area of the breach

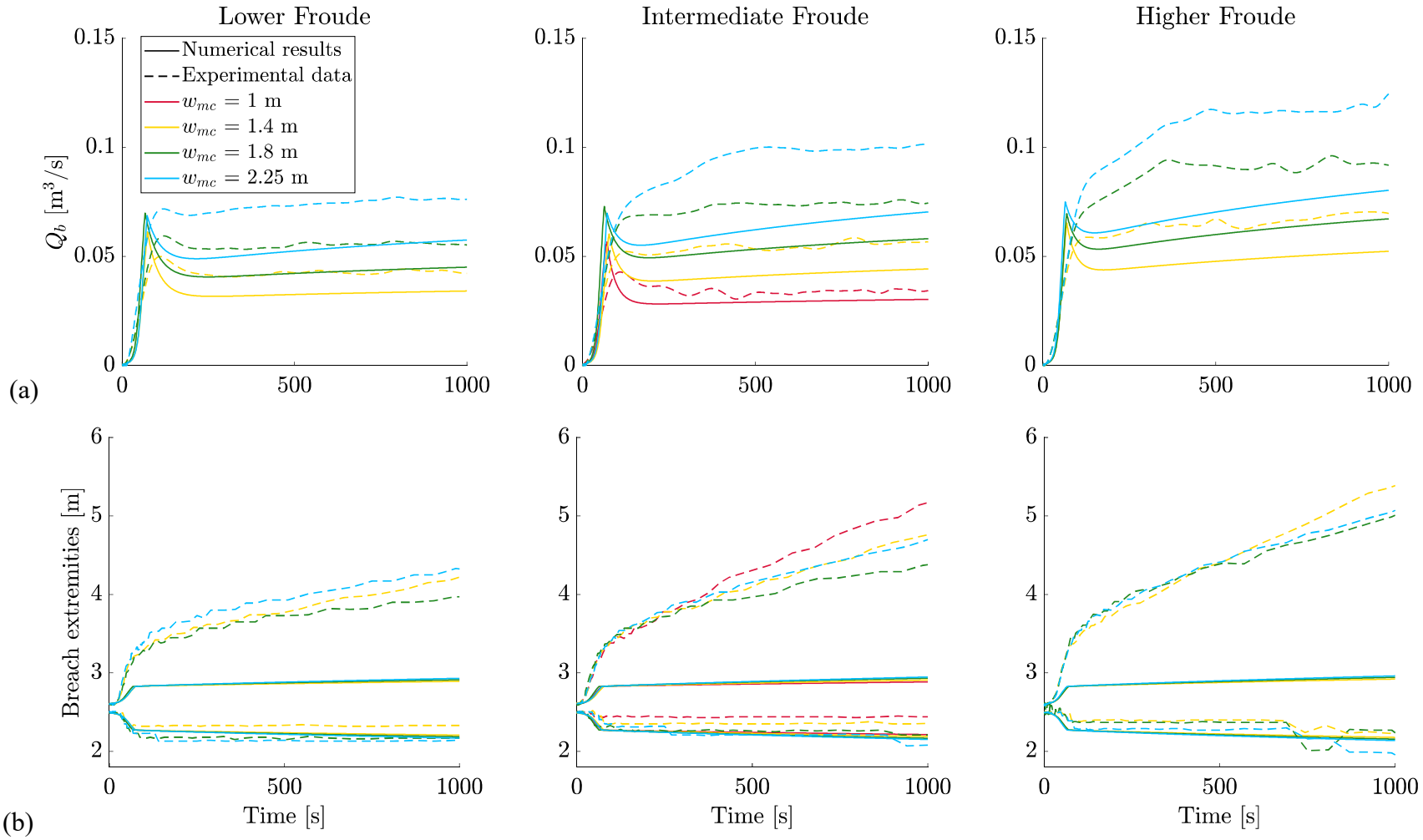


Figure S7. Comparison between results obtained experimentally and numerically with the dam breaching model (Wu, 2013), i.e., original model: (a) breach hydrographs; (b) position of the breach extremities.

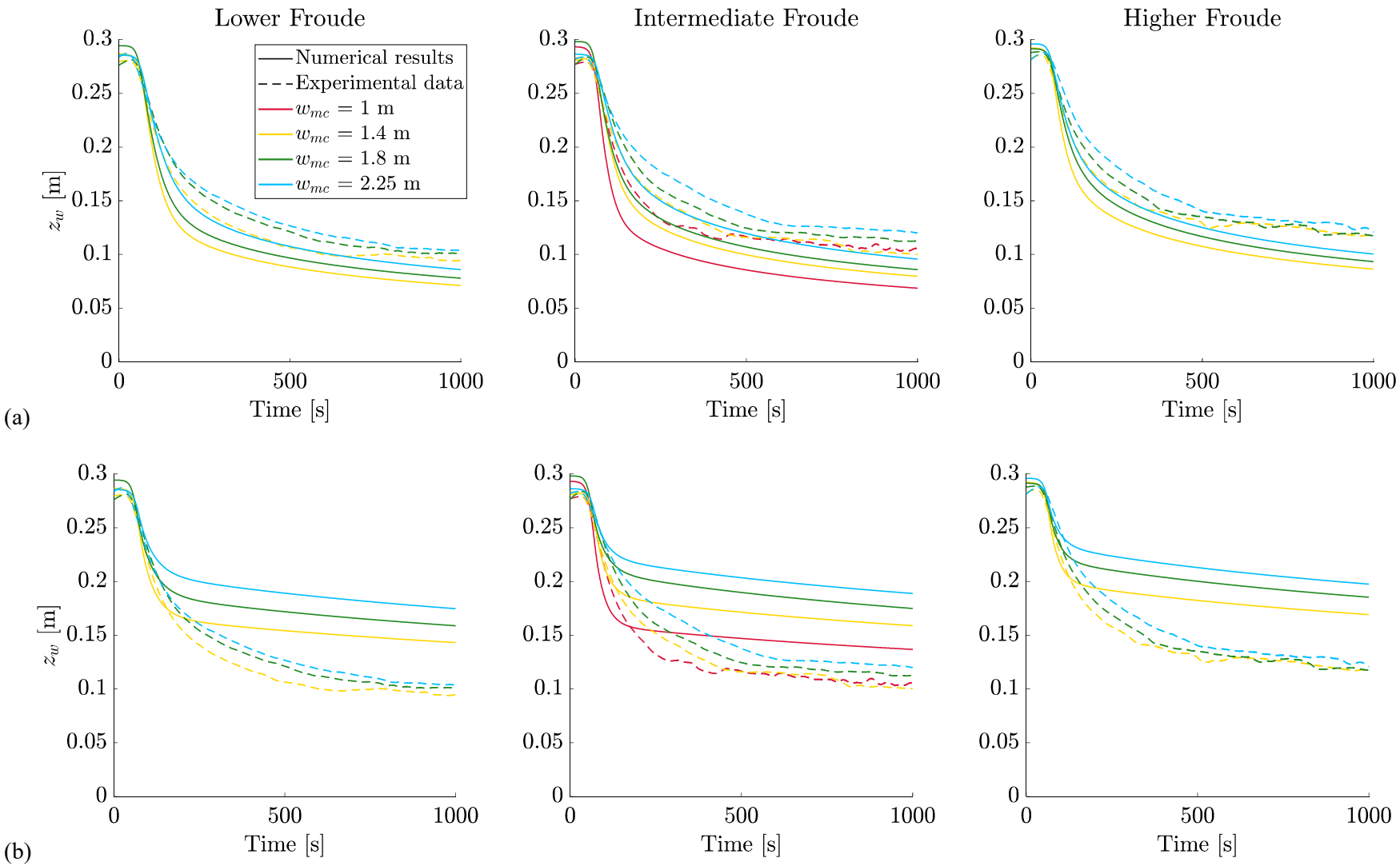


Figure S8. Evolution of the mean water level in the main channel obtained experimentally and numerically: (a) with the dike model; (b) with the dam model.

Chapter 4

Efficient Modelling of Lateral Discharge Through a Dike Breach

This chapter corresponds to the journal paper "Efficient modelling of lateral discharge through a dike breach" by V. Schmitz, V. Kitsikoudis, G. Wylock, S. Erpicum, M. Pirotton, P. Archambeau, and B. Dewals, published in 2024 in Journal of Hydrology. The PhD candidate contributed to the literature review, methodology, and numerical models, generated figures, and contributed to writing and revising the manuscript.

Efficient Modelling of Lateral Discharge Through a Dike Breach

Vincent Schmitz[1*], Vasileios Kitsikoudis[2], Gregoire Wylock[1], Sebastien Erpicum[1], Michel Pirotton[1], Pierre Archambeau[1], and Benjamin Dewals[1]

[1]*Research Group of Hydraulics in Environmental and Civil Engineering (HECE), University of Liège, Liège, Belgium.*

[2]*Water Engineering and Management, Faculty of Engineering Technology, University of Twente, 7500 AE Enschede, The Netherlands.*

Corresponding author: Vincent Schmitz (V.Schmitz@uliege.be)

Key Points:

- Two numerical models were developed to predict side discharge through a dike breach.
- The flow through a dike breach was considered similar to the flow through a side weir.
- Both models used empirical formulas to evaluate the side weir discharge coefficient.

ABSTRACT

Breaches in fluvial dikes can lead to major flooding in the hinterland with severe societal and economic consequences. The discharge partitioning at the location of a dike breach is a complex flow phenomenon with 2D and 3D flow features that needs to be predicted accurately for the estimation of flood hazard. Determining the exact location of a potential breach is highly uncertain and so are the circumstances in which it could appear. Therefore, many scenarios should be investigated. Fast and accurate modelling of the discharge partitioning with appropriate simplifications and parameterizations are required to allow for a large number of simulations within reasonable computational time. To achieve this, spatially lumped or one-dimensional flow models have been used in combination with side weir equations. For the first time, the present study systematically assesses the performance of eleven side weir equations for the determination of the lateral discharge through a breach in a dike that is parallel to the flow direction along a straight river reach. These side weir equations were implemented in a zero-dimensional spatially lumped flow model and in a one-dimensional spatially distributed flow model. Both models were evaluated against experimental data from laboratory tests with a side opening that was either fixed, or dynamically evolving. The performance of the side weir equations varied with the experimental data, highlighting the empirical nature of most of these equations. The coupling of the side weir equations with the spatially distributed flow model did not always generate better results than the coupling with the lumped model, which implies that increasing the model complexity does not systematically lead to better predictions of the dike breach discharge.

1. Introduction

Fluvial dikes or levees are engineered structures that confine the river flow at high discharges and provide flood protection. Several mechanisms can lead to a dike failure, such as overtopping (Rifai et al., 2017), slumping failure (Elalfy et al., 2018), piping (Vorogushyn et al., 2009), seepage (Onda et al., 2019), and the weakening of some parts of the dike caused by the actions of burrowing animals and poor maintenance (Orlandini et al., 2015). The failure of a fluvial dike can induce major flooding in the hinterland with detrimental consequences to local economies and potentially life losses, especially in urbanized areas. Thus, the accurate prediction of the dike breach hydrograph is of paramount importance for the estimation of flood hazard and for the determination of safe evacuation routes, especially when considering the fact that people tend to settle within flood sheltered areas (Haer et al., 2020). This kind of river flood risk is expected to increase in the future, as more people are exposed to floods (Tellman et al., 2021) and extreme precipitation is projected to increase (Madsen et al., 2014).

The discharge partitioning in a channel with a side opening is inherently a 2D and 3D flow phenomenon (Stilmant et al., 2013; Li et al., 2021; Dewals et al., 2023) with flow separation, flow recirculation, and helicoidal flow near the side opening (Michelazzo et al., 2015). The lateral flow discharge through a side opening has been modelled successfully using 2D models (Roger et al., 2009; Yu et al., 2013; Echeverribar et al., 2019; Shustikova et al., 2020). Recent studies have coupled 2D hydraulic models with physically-based erosion models (Kakinuma and Shimizu, 2014; Elalfy et al., 2018; Dazzi et al., 2019). Also, 1D-2D hydraulic models were extended using additional computational elements, such as a physically-based description of

the breach formation and evolution (Viero et al., 2013). Another example is the use of probabilistic frameworks for possible locations and timing of breaches along a dike with the aid of fragility curves for various failure mechanisms (Vorogushyn et al., 2010; Bomers et al., 2019; D'Oria et al., 2019; Maranzoni et al., 2022). Despite recent advancements (Dazzi et al., 2019; Ferrari et al., 2020), such models have either long computational times or large uncertainties regarding their parameterizations. This can become prohibitive for systematic analyses of the flood hazard from dike breaches since breaches at different dike locations can lead to different flood patterns. There is a need for simpler and faster modelling procedures that approximate the complex hydrodynamic processes near and through a dike breach. This is particularly relevant for sensitivity analysis (Schmitz et al., 2023a) and for inclusion in river models, which are also subject to further simplifications (Kitsikoudis et al., 2020), for systematic flood risk assessment.

Side weir equations provide an easy parameterization of the lateral discharge through a side opening. While side weir equations have been applied to fluvial dike breaches (Kamrath et al., 2006; Yu et al., 2013; Wei et al., 2016), the number of such studies is small compared to those modelling frontal dike breaches or earthen embankment breach as a frontal weir (ASCE/EWRI Task Committee, 2011; Schmocker and Hager, 2012; Wu, 2013). As a result, the accuracy of the side weir equations and the parameterization of the side weir discharge coefficient, C_d, for breaches in fluvial dikes have not been thoroughly evaluated yet. Mignot et al. (2020) assessed the predictive capability of eleven semi-empirical formulations for C_d. To this end, they used laboratory experiments of flow intrusion into buildings and considered the building openings (doors, windows, and gates) as rectangular side weirs. The computed side discharge was highly dependent on the formulation of C_d. The applicability of such semi-empirical equations to fluvial dike breaches is not straightforward. Dike breaches exhibit a time-dependent complex geometry, i.e., a dynamically evolving and uneven breach profile (Rifai et al., 2017), that depends on the discharge in the main channel, and the dike geometry and erodibility (Schmitz et al., 2021).

Rifai et al. (2017, 2018) identified three stages during the development of a non-cohesive dike breach caused by overtopping (Figure 1). Initially, right after overtopping occurs, the erosion rate of the dike is low because the flow depth and the velocity above the dike are also small. As the flow depth and the velocity above the dike breach increase, erosion intensifies rapidly with large water level variations near the breach (Al-Hafidh et al., 2022) and a downstream shift of the breach center. Finally, in the last stage, the upstream side of the breach remains almost fixed, and the breach grows more slowly in the downstream direction.

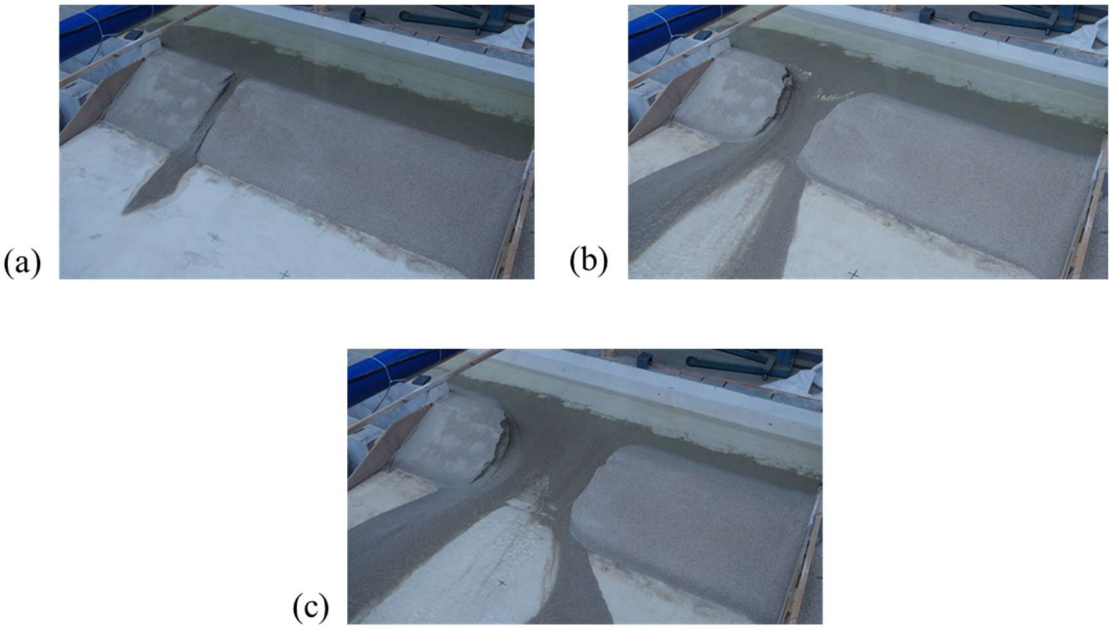


Figure 1. Three stages observed during a non-cohesive homogeneous dike breaching event due to overtopping: (a) initial low erosion phase, (b) erosion intensification period, and (c) erosion stabilization phase. The corresponding experimental test was performed at the "Laboratoire d'hydraulique des constructions" of the University of Liège, Belgium.

Similar observations were reported by Michelazzo et al. (2018) and Kakinuma and Shimizu (2014) from laboratory-scale and field-scale experiments, respectively. Overall, the dike breaching process is accelerated and the final shape of the breach gets larger with increasing water level and flow discharge in the river (Yu et al., 2013; Rifai et al., 2017; Wu et al., 2018). The dike composition also affects the breach expansion (Yu et al., 2013; Wu et al., 2018; Rifai et al., 2021), so does the channel width (Schmitz et al., 2023b) and the dike geometry, with dikes with larger volume per unit width inducing a more gradual enlargement of the breach during the rapid expansion phase (Schmitz et al., 2021). The breach location and its final shape govern to a large extent the flood inundation (Tadesse and Fröhle, 2020). The water level in the hinterland (Rifai et al., 2018) and the presence or not of riprap also alter the breach development (Ahadiyan et al., 2022).

The present study investigates the suitability of different parameterizations for lateral flow through dike breaches for the incorporation of 3D flow processes into simpler and faster models. Specifically, this study assesses the performance of eleven semi-empirical side weir equations applied in two different modelling frameworks to determine the lateral outflow discharge through a breach in a dike that is parallel to the flow main direction along a straight river reach. The assessment is performed in two frameworks: by combining the side weir equations with either a spatially lumped flow model based on mass conservation or a one-dimensional flow model based on the shallow water equations. The semi-empirical side weir equations, coupled with the two models, are firstly tested with data from channels with fixed side openings and secondly with data from a channel with a dynamic evolution of a dike breach.

Section 2 presents the numerical models used in this work. Section 3 briefly presents the experimental data from the literature that are used for the validation of the developed models

(three experiments with a lateral outflow through a fixed geometry and one experiment with an evolving dike breach geometry). Section 4 presents how the breach discharge predicted by the semi-empirical equations compares with the experimental measurements. The performance and limitations of the models are discussed in Section 5. Finally, conclusions are drawn in Section 6.

2. Modelling of the lateral flow discharge through a dike breach

2.1. Discharge coefficient for lateral side weirs

The lateral flow discharge through a breach in a dike that runs parallel to the flow direction in a straight river reach with width W (Figure 2) can be simplified as a flow over a broad-crested rectangular side weir. In such a case, the incoming flow discharge from the upstream of the channel, Q_{in}, is divided at the location of the dike breach into the lateral flow discharge, Q_b, towards the hinterland and the outgoing flow discharge, Q_{out}, towards the downstream of the channel. In a straight channel, the lateral flow discharge over a rectangular side weir is expressed as:

$$Q_b = \frac{2}{3} C_d \sqrt{2g(h-p)^3} L_s \,, \tag{1}$$

where C_d is the side weir discharge coefficient, h is the flow depth in the main channel upstream of the side opening, p is the crest height of the side weir, L_s is the length of the side weir, and g is the acceleration of gravity. The discharge coefficient, C_d, is typically estimated from analytical and semi-empirical relationships based on the weir and flow characteristics. Similarly to Mignot et al. (2020), eleven relationships from the literature for the calculation of C_d for sharp-crested weirs are used in this study (Table 1).

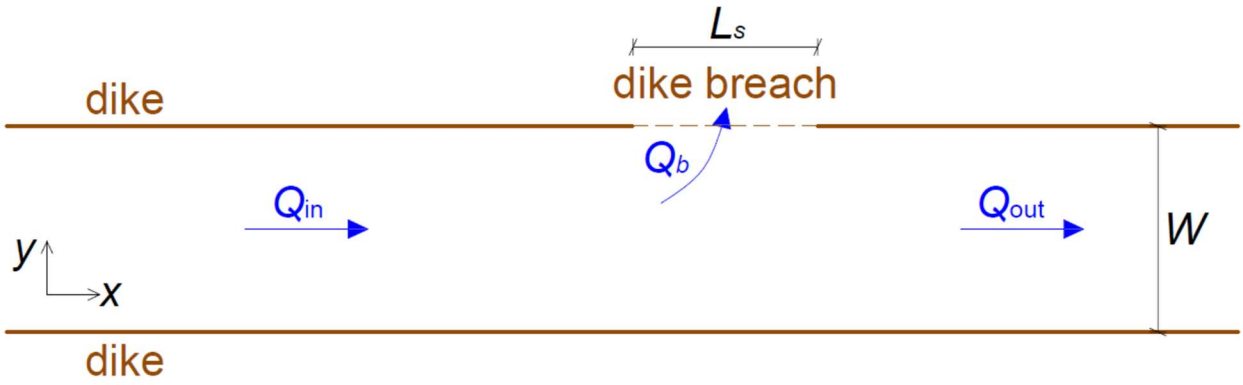


Figure 2. Flow within a straight channel with a dike breach on the side.

Table 1. Formulas for the estimation of the side weir discharge coefficient, C_d, and the associated ranges of Fr $(= U/\sqrt{gh})$, p/h, and L_s/W for which these formulas were developed. Fr and h were measured at the channel centerline, just upstream from the side breach. The table is adapted from Mignot et al. (2020).

No	Source	C_d (-)	Fr (-)	p/h (-)	L_s/W (-)
1	Nadesamoorthy and Thomson (1972)	$0.432\left(\frac{2+\mathrm{Fr}^2}{1+2\mathrm{Fr}^2}\right)^{0.5}$	0.02 – 4.3	0 – 0.96	0.2 – 1
2	Subramanya and Awasthy (1972)	$0.611\left(1-\frac{3\mathrm{Fr}^2}{2+\mathrm{Fr}^2}\right)^{0.5}$	0.02 – 0.9	0.2 – 0.96	0.2 – 1
3	Yu-Tek (1972)	$0.622-0.222\mathrm{Fr}$	0.02 – 4.3	0 – 0.96	0.2 – 1
4	Ranga Raju et al. (1979)	$0.81-0.6\mathrm{Fr}$	0.1 – 0.5	n/a	0.33 – 0.5
5	Hager (1987)	$0.636\left(1+\frac{(H-p)^3}{7H^3}\right)\left(\frac{H-p}{3H-2h-p}\right)^{0.5}$	Analytical approach		
6	Singh et al. (1994)	$0.33-0.18\mathrm{Fr}+0.49\frac{p}{h}$	0.22 – 0.42	0.45 – 0.85	0.4 – 0.8
7	Swamee et al. (1994)	$0.447\left[\left(\frac{44.7p}{49p+h}\right)^{6.67}+\left(\frac{h-p}{h}\right)^{6.67}\right]^{-0.15}$	0.1 – 0.93	0 – 0.31	0.4 – 1
8	Jalili and Borghei (1996)	$0.71-0.41\mathrm{Fr}-0.22\frac{p}{h}$	0.1 – 2	0.05 – 0.87	0.67 – 2.5
9	Borghei et al. (1999)	$0.7-0.48\mathrm{Fr}-0.3\frac{p}{h}+0.06\frac{L_s}{W}$	0.1 – 0.9	0.02 – 0.87	0.33 – 2.33
10	Emiroglu et al. (2011)	$\left[0.836+\left(-0.035+0.39\left(\frac{p}{h}\right)^{12.69}+0.158\left(\frac{L_s}{W}\right)^{0.59}+0.049\left(\frac{L_s}{h}\right)^{0.42}+0.244\mathrm{Fr}^{2.125}\right)^{3.018}\right]^{5.36}$	0.08 – 0.92	0.34 – 0.91	0.3 – 3
11	Bagheri et al. (2014)	$-1.423\mathrm{Fr}^{0.138}+0.744\left(\frac{h-p}{L_s}\right)^{-0.083}+0.723\left(\frac{h-p}{p}\right)^{0.088}+0.182\left(\frac{L_s}{W}\right)^{-0.241}$	0.08 – 0.91	0.22 – 0.9	0.5 – 1.5

2.2. Numerical models for flow through a dike breach

Two simplified models were developed using *MATLAB* software to predict the lateral flow discharge, Q_b, through a side opening representing a dike breach. Section 2.2.1 presents a lumped model and Section 2.2.2 presents a one-dimensional finite volume model. Both models are coupled with Eq. (1) and with an equation for the determination of C_d (Table 1).

2.2.1. Lumped model (zero-dimensional)

The lumped model is based on a side weir discharge equation for the determination of Q_b (Eq. (1)) and on the mass balance equation in a control volume, which writes:

$$\frac{dV}{dt} = A_{xy}(h_{xy})\frac{dh_{xy}}{dt} = Q_{\text{in}} - Q_b(h_{xy}) - Q_{\text{out}}(h_{xy}) \,, \tag{2}$$

where V is the volume of water in the control volume, t is the time, A_{xy} is the horizontal surface area of the control volume, and h_{xy} is the flow depth spatially averaged across the area A_{xy}.

This numerical model estimates the spatially averaged flow depth h_{xy} for every timestep dt, based on the value of h_{xy} in the previous timestep, the measured value of Q_{in}, an experimental rating curve for Q_{out}, and the value of Q_b computed at the previous timestep with Eq. (1). The surface area A_{xy} is considered as constant in each case.

2.2.2. Spatially distributed model (one-dimensional)

The spatially distributed model solves numerically the one-dimensional Saint-Venant equations, i.e., the conservation of mass (Eq. (3)) and the conservation of momentum (Eq. (4)):

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = -q_b \tag{3}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x}(QU + gAy) + gA\left(S_f - S_0\right) = -Uq_b \tag{4}$$

where A is the flow cross-sectional area, Q is the discharge in the channel, q_b is the specific lateral discharge through the side opening, U is the longitudinal velocity averaged over the cross-section, y is the cross-section average of the depth, S_f is the friction slope, and S_0 is the bed slope. By considering $U = Q/A$, expressing y as a function of the cross-sectional area, $y = f(A)$, and estimating the friction slope with the Manning formula, the only remaining unknowns in Eqs. (3) and (4) are Q and A.

Eqs. (3) and (4) are discretized spatially with a finite volume numerical scheme (Kerger et al., 2011) and in time with a two-step Runge-Kutta algorithm. q_b is computed on each spatial step, Δx, as a function of the local water depth and the local height of the side weir crest (h and p are considered constant over each cell). Δx was 1 cm in the experiments of Michelazzo et al. (2015) and Mignot et al. (2020), and 10 cm in the experiments of Roger et al. (2009) and Rifai et al. (2017) (see Section 3).

2.3. Initial and boundary conditions

The initial water level in the main channel for each simulation was the same as the water level under steady flow conditions with the prescribed incoming discharge, Q_{in}, of the corresponding experiment but without a dike breach. In practice, this water level was computed using our numerical models without considering any breach discharge. The equilibrium breach discharge was obtained when the flow reached a new steady state in the presence of a dike breach. The downstream boundary condition was a rating curve at the weir (Roger et al., 2009), sluice gate (Michelazzo et al., 2015), crested tailgate (Mignot et al., 2020), or perforated plate (Rifai et al., 2017) of each experiment.

3. Experimental data

The accuracy of the eleven semi-empirical side discharge equations (Section 2.1) is evaluated by comparing the modelling results with experimental data from the literature. The considered experimental data comprise data from (a) experiments with a fixed side opening (Roger et al., 2009; Michelazzo et al., 2015; Mignot et al., 2020) and (b) experiments with a dynamically evolving side opening (Rifai et al., 2017). In all experiments, the side opening was parallel to the main flow direction and in a straight channel (Figure 2), which had uniform and steady flow characteristics. The experimental setups and methods used to generate these data are briefly described in this section and summarized in Table 2. The setup used for each test campaign is displayed in Figure 3. More details can be found in the referenced studies.

3.1. Experiments with fixed geometry of the side opening

3.1.1. Experiments of Roger et al. (2009)

Roger et al. (2009) carried out laboratory experiments of a dike breach in a 1 m-wide rectangular flume. The rectangular side opening was 0.7 m long and its crest had zero height, simulating the complete failure of a dike. The lateral discharge through the side opening was propagated into a 3.5 × 4.0 m^2 basin that was at the same level as the flume bottom and was made of glass. The water flowed freely off the edges of the basin, while at the downstream end of the flume the water flowed over a weir. The tested inflow discharges were equal to 0.2 m^3/s and 0.3 m^3/s and the tested undisturbed flow depths were 0.4 and 0.5 m, resulting in four experimental combinations in total.

3.1.2. Experiments of Michelazzo et al. (2015)

Michelazzo et al. (2015) conducted laboratory experiments in a 30 cm wide recirculating flume with a slope of 0.1% and rectangular cross section. Fine gravel was glued on the bed of the flume to provide roughness. A vertical sluice gate was placed at the downstream extremity of the flume to adjust the flow depth. A rectangular side weir with zero height diverted a portion of the flow to a lateral channel at a lower level, which conveyed water toward a storage basin through its downstream end. Ten different side weir lengths were tested and varied from 3 to 47 cm. In all cases, the incoming discharge was about 0.01 m^3/s while the sluice gate at the downstream end of the flume was always set in the same way.

3.1.3. Experiments of Mignot et al. (2020)

Mignot et al. (2020) investigated the flow intrusion from a flooded street into buildings through openings, such as doors, gates, and windows, during urban floods. The flow through building openings, without nearby obstacles, can be considered as flow through a rectangular side weir with zero and non-zero crest height for doors/gates and windows, respectively. Alternatively, these cases can represent the total or partial collapse of a dike section. The experiments were carried out in a 0.79 m-wide flume with a smooth bed and a rectangular cross section. The slope of the flume was 0.18% and the flow depth was regulated by a sharp crested tailgate at the downstream end of the flume. The side opening was 7.9 cm long for the cases with zero weir crest height and 15.4 cm long for the cases with non-zero weir crest height. Four and two inflow discharges were tested with zero and non-zero weir crest height, respectively. Mignot et al. (2020) also investigated the impact of urban obstacles, such as parked cars, on the flow intrusion but these cases are not considered in this study.

3.2. Experiments of Rifai et al. (2017) with dynamic evolution of the geometry of the side opening

Rifai et al. (2017) investigated the evolution of a breach in a sandy homogeneous trapezoidal dike due to overtopping. The experiments were conducted in a 10 m-long and 1 m-wide flume with a trapezoidal cross section. The erodible dike stretched over 3 m along the right side of the flume and separated the flume from a 4.3 m x 2.5 m area behind the dike. The bottom of this area and of the main flume was coated with impermeable whitewash to ensure a uniform bed roughness over the entire setup. A small notch was carved at the crest of the dike to initiate erosion at this specific spot. Once the dike breach began forming, its evolution was monitored with a nonintrusive laser profilometry technique (Rifai et al., 2019). Rifai et al. (2017) conducted experiments for many inflow discharges and downstream boundary conditions. In this work, we focus on four experiments based on four different inflow discharges ranging from 0.02 m^3/s to 0.05 m^3/s. In all cases, a perforated plate was used at the flume downstream extremity to regulate the water level.

In the experiments of Rifai et al. (2017) the side opening evolved and grew bigger in time since the dike was made of sand that is eroded by the flow. In the cases with fixed geometry of the side opening, the lateral flow could be considered steady; however, in the experiments of Rifai et al. (2017) the flow that goes through the side opening depends on time. For each time step, the crest height of the side weir, i.e., the breach in the dike, was determined from a scanned longitudinal line in the middle of the dike in the transverse direction. The profile of the dike breach is typically irregular; however, the models require a single crest height, either for the whole dike breach for the lumped model or for each spatial step $\boldsymbol{\Delta x}$ for the spatially distributed model. For the spatially distributed model, the crest height of each spatial step was chosen equal to the breach elevation at the center of the step, which was obtained using a linear interpolation between scanning measurements of the breach (spatial resolution of 1 cm). The side weir equations were then applied on each one-dimensional cell individually. For the lumped model, a representative crest height was chosen equal to the 15^{th} percentile of the elevation of the scanned points along the dike crest center line that were within the dike breach.

Table 2. Main characteristics of the experimental tests considered in this study.

	Fixed breach geometry				Dynamic breach geometry
	Zero lateral crest height			Non-zero lateral crest height	
Source	Roger et al. (2009)	Michelazzo et al. (2015)	Mignot et al. (2020)	Mignot et al. (2020)	Rifai et al. (2017)
Number of tests	4	10	4	2	4
Channel width (m)	1	0.3	0.79	0.79	1
Channel length (m)	9	5.1	8.35	8.35	10
Slope (%)	-	0.3	0.18	0.18	-
Breach length (m)	0.7	0.03 – 0.47	0.079	0.154	dynamic
Manning coefficient ($s/m^{1/3}$)	0.015	0.024	0.01	0.01	0.018
Q_{in} (m^3/s)	0.2 and 0.3	~ 0.01	$5.78 \cdot 10^{-4}$ – $4.19 \cdot 10^{-3}$	$1.142 \cdot 10^{-3}$ / $1.273 \cdot 10^{-3}$	0.02 – 0.05
Downstream boundary condition	weir	sluice gate	crested tailgate	crested tailgate	perforated plate

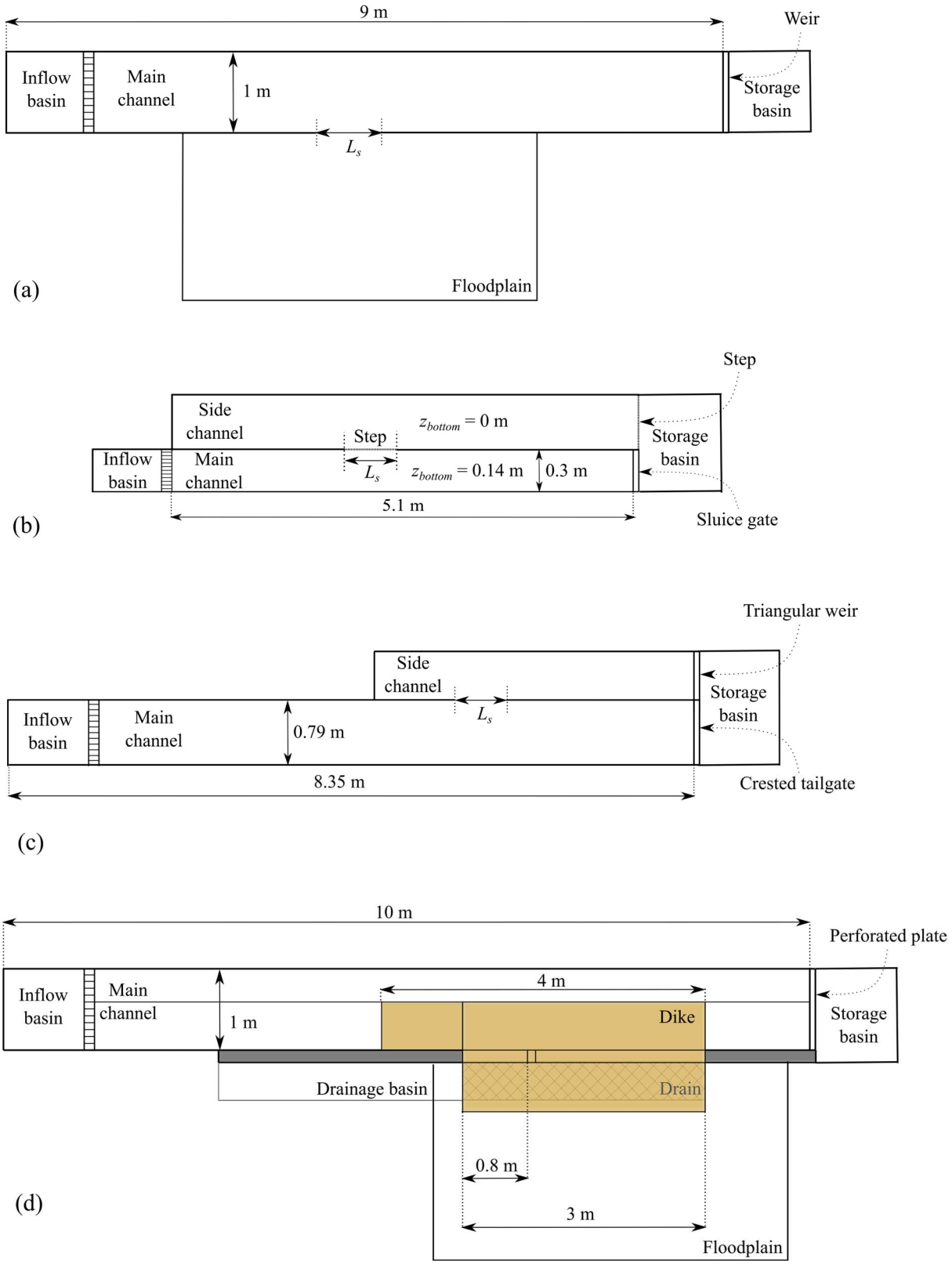


Figure 3. Experimental setups used by (a) Roger et al. (2009), (b) Michelazzo et al. (2015), (c) Mignot et al. (2020), and (d) Rifai et al. (2017). The subfigures are adapted from the cited studies.

4. Results

Sections 4.1 and 4.2 present a comparison between the results obtained from the two numerical models of Section 1.1 coupled with the side weir discharge coefficients from Table 1 and the corresponding measurements from the experimental configurations with fixed and dynamic geometry of the side opening, respectively.

4.1. Side opening with fixed geometry

As shown in Figure 4, for the four experimental cases of Roger et al. (2009), the lumped model performed better than the spatially distributed model in estimating the lateral discharge $\boldsymbol{Q_b}$ through the side opening for almost every $\boldsymbol{C_d}$ formula that was tested. A notable difference is that the spatially distributed model always underestimated $\boldsymbol{Q_b}$, while the lumped model generated more variable results. This difference may be understood by considering that, in practice, the water level significantly drops in the vicinity of the side breach due to water acceleration. This trend is captured by the one-dimensional model. Conversely, the lumped model overestimates the water level close to the breach as it only considers an averaged value over the entire control volume. As a result, using Eq. (1) with the average water height in the lumped model leads to larger estimations of the breach discharge.

Both models, and particularly the spatially distributed model, performed better for the smaller discharge and the larger flow depth, i.e., for smaller inflow Froude number. The height of the side weir crest in the experiments of Roger et al. (2009) was zero and as a result the $\boldsymbol{C_d}$ formula of Bagheri et al. (2014) was not tested because the weir crest height, $\boldsymbol{p}$, is used in the denominator. From the remaining discharge coefficient equations in Table 1, the equations of Subramanya and Awasthy (1972) and Borghei et al. (1999) with the lumped model performed best on average , while the spatially distributed model was most successful when coupled with the equation of Nadesamoorthy and Thomson (1972), followed by Yu-Tek (1972). The equation of Singh et al. (1994) performed worst with both models. However, it needs to be noted that while the discharge coefficient equations were compared with the same data, the semi-empirical equations, such as the equation of Singh et al. (1994), were developed in different data ranges (Table 1). As a result, some of them were used out of their calibration range, which is the case with the Singh et al. (1994) formulation. Roger et al. (2009) also modelled their experiments with a 2D finite volume model and a 2D finite element model. The models underestimated the $\boldsymbol{Q_b}$ measurements by approximately 6% to 12% (Figure 4). The absolute deviation of the 2D models from the measurements was larger than that of the lumped model, regardless of the $\boldsymbol{C_d}$ formula with the exception of Singh et al. (1994) and Emiroglu et al. (2011), and larger than most combinations of the spatially distributed model with a $\boldsymbol{C_d}$ formula (Figure 4).

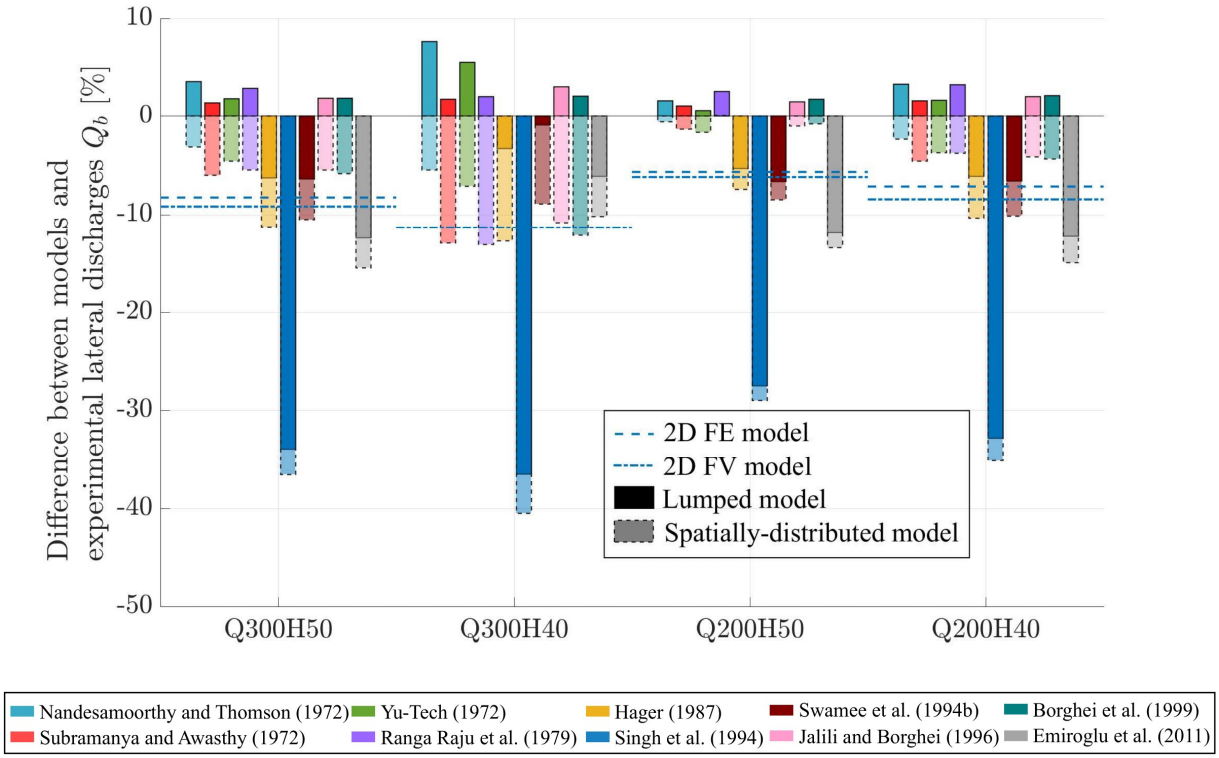


Figure 4. Comparison between modelled and measured dike breach discharges, Q_b, for the experiments of Roger et al. (2009). The results from the 2D finite element (FE) and 2D finite volume (FV) models used by Roger et al. (2009) are also included in the comparison.

The experiments of Michelazzo et al. (2015) are similar to those of Roger et al. (2009) when considering the height of the side weir crest, which is also zero. However, in this case the length of the side weir is shorter as it varies from 3 cm to 47 cm. Contrary to the cases of Roger et al. (2009), in the experiments of Michelazzo et al. (2015) both the lumped model and the spatially distributed model mostly overpredicted the lateral discharge $\boldsymbol{Q_b}$, particularly as the length of the side weir increased (Figure 5). The only consistent exception is the case where the models were coupled with the discharge coefficient equation of Singh et al. (1994), presumably because it was used outside of the data range for which it was developed. Overall, the accuracy of both models decreased as the side weir length increased. Another difference with the experiments of Roger et al. (2009) is that in the data of Michelazzo et al. (2015) the spatially distributed model performed better than the lumped model, for almost every coupled discharge coefficient equation. The discharge coefficient equation of Subramanya and Awasthy (1972) generated the best results when it was coupled with either of the two models. The rest of the discharge coefficient equations exhibited a rather erratic behavior when coupled either with the lumped model or the spatially distributed model. The formula of Bagheri et al. (2014) was again not tested because it is not applicable to a side weir with zero crest height.

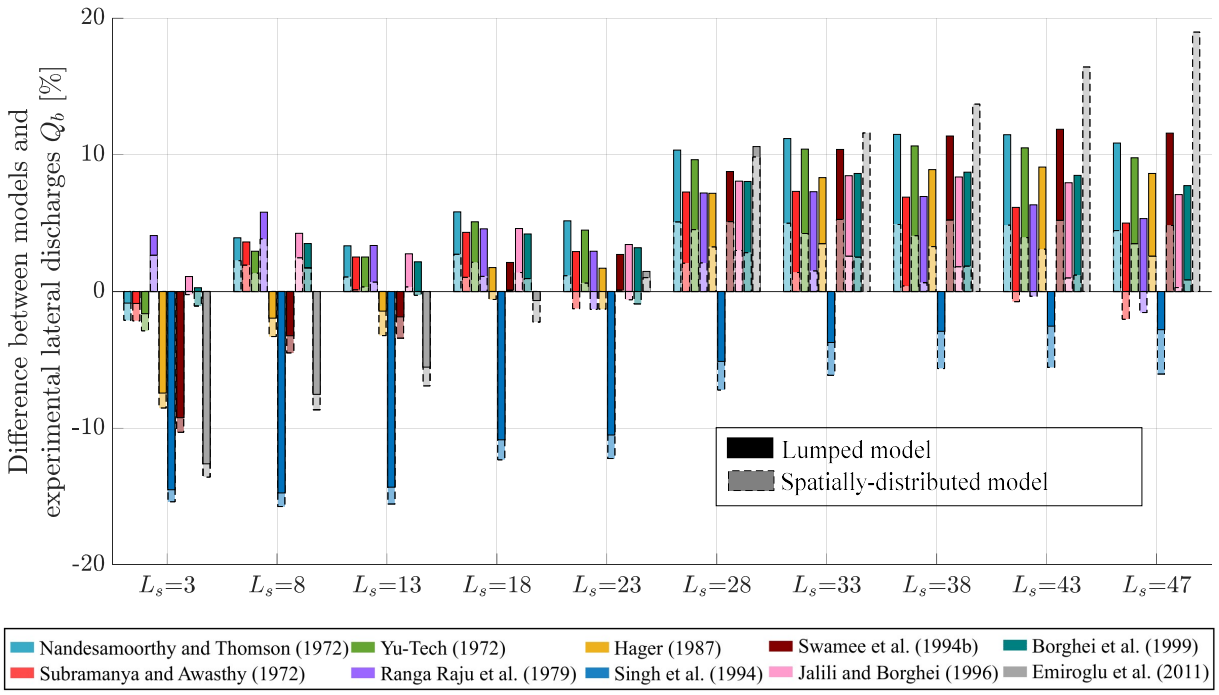


Figure 5. Comparison between modelled and measured dike breach discharges, Q_b, for experiments of Michelazzo et al. (2015). The side opening widths, L_s, are in cm. No converged value could be obtained for $L_s \geq 33$ cm when using the formula of Emiroglu et al. (2011) in the lumped model.

With regard to the experiments of Mignot et al. (2020), for the cases with a fixed side opening with zero crest height there was considerable underestimation of the lateral discharge through the side opening on several occasions. The C_d equations that performed best were the same as those in the data of Roger et al. (2009), i.e., the equation of Subramanya and Awasthy (1972) for the lumped model and the equation of Nadesamoorthy and Thomson (1972) for the spatially distributed model. When the height of the crest of the side weir became non-zero, both models generated less accurate results compared to the zero crest height and underpredicted the experimental measurements by more than 25% for most couplings with a discharge coefficient equation (Figure 6). A notable exception was the coupling with the discharge coefficient equation of Singh et al. (1994), where both models performed much better. Overall, the lumped model performed better than the spatially distributed model for most cases of Mignot et al. (2020).

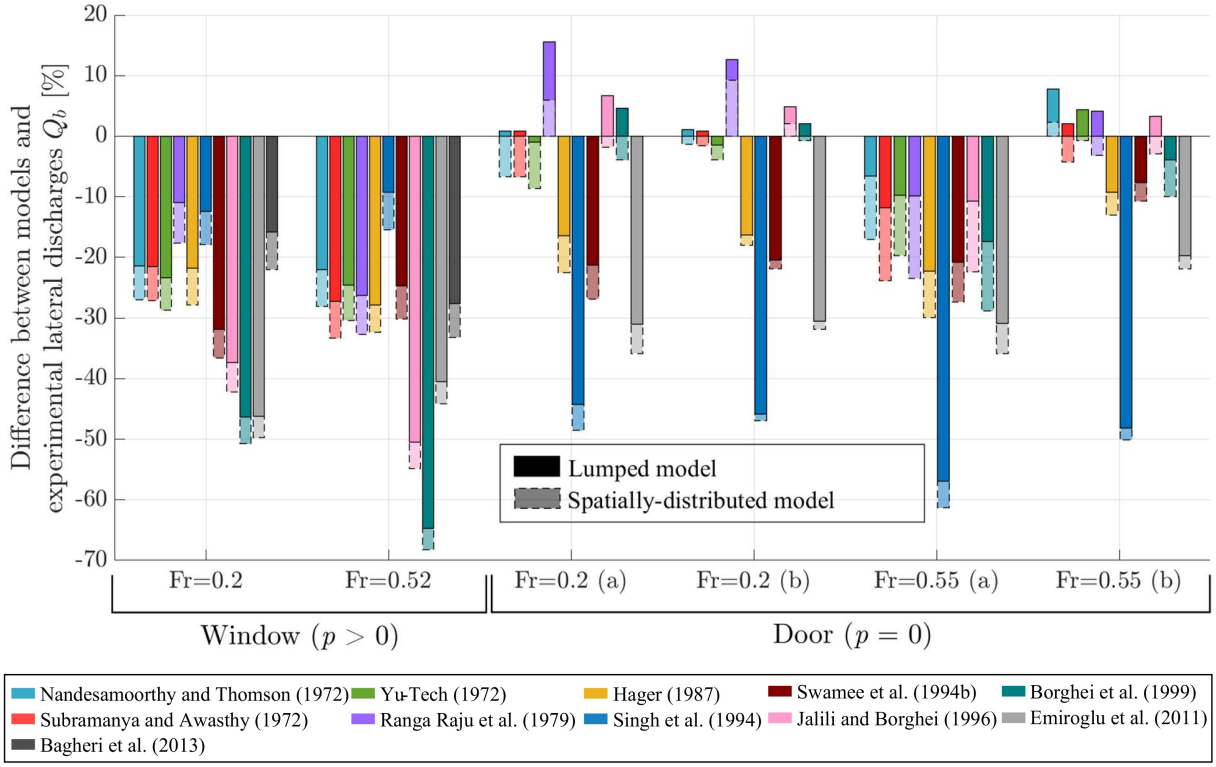


Figure 6. Comparison between modelled and measured dike breach discharges, Q_b, for experiments of Mignot et al. (2020). Labels (a) and (b) refer to different combinations of inflow discharge and water height in the main channel but leading to the same Froude number.

4.2. Side opening with dynamic evolution of its geometry

As shown in Figure 7, for the four representative experimental cases of Rifai et al. (2017), the lumped model followed quite well the overall evolution trend of the discharge through the gradually augmenting dike breach, although in a less smooth way. The model did not capture accurately the discharge lowering after the initial peak in Test 1. For the lower upstream flow discharge, the discharge coefficient formulations did not have a significant impact on $\boldsymbol{Q_b}$; however, as the flow discharge increased the modelling results of $\boldsymbol{Q_b}$ exhibited a considerable scatter, particularly towards the end of each experiment when the dike breach was rather stabilized. Compared to the lumped model, the spatially distributed model predicted more accurately the initial peak of $\boldsymbol{Q_b}$ and was also able to reproduce the trend of the lateral discharge through the variable dike breach (Figure 8). However, the modelled $\boldsymbol{Q_b}$ at the stabilization phase, after approximately 100 s, mostly underpredicted the measured values and gave worst results than the lumped model for the two lowest values of $\boldsymbol{Q_{in}}$, i.e., 20 and 30 l/s. The different discharge coefficient formulations that were tested in the spatially distributed model did not affect $\boldsymbol{Q_b}$ as much as they did in the lumped model, regardless of the upstream discharge that was prescribed in each experiment. The formulas of Emiroglu et al. (2011) and Bagheri et al. (2014) were not evaluated in this experiment because for the former there was no convergence in the numerical models and the latter is not applicable to a side weir with zero crest height.

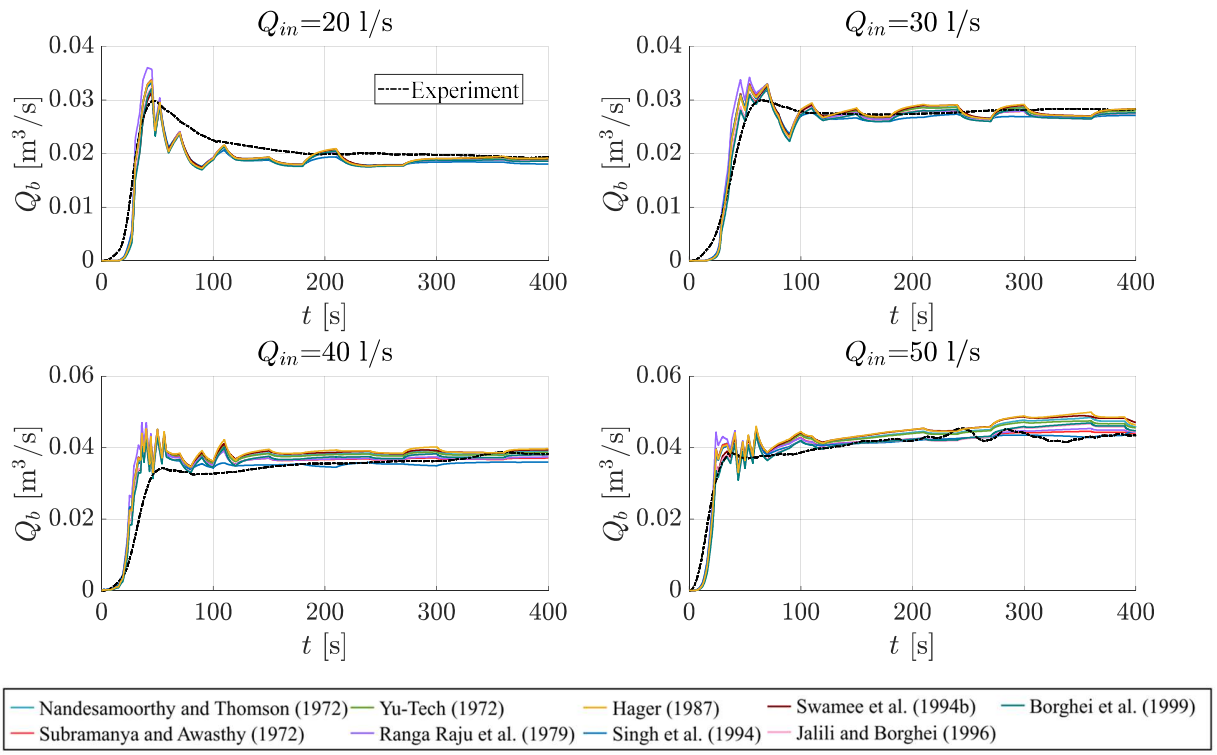


Figure 7. Evolution of the dike breach discharge, Q_b, with the lumped model for experiments of Rifai et al. (2017). The formula of Emiroglu et al. (2011) did not converge and the formula of Bagheri et al. (2014) cannot be applied when $p = 0$.

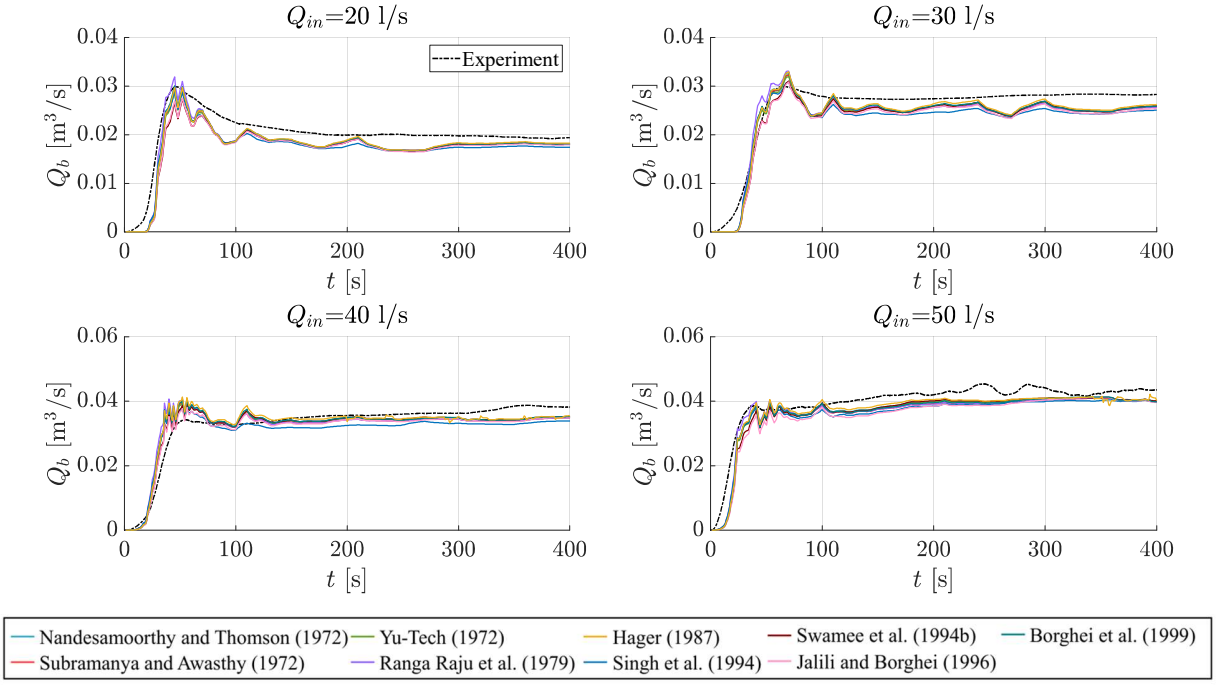


Figure 8. Evolution of the dike breach discharge, Q_b, with the spatially distributed model for experiments of Rifai et al. (2017).

The peak breach discharge, $Q_{b,peak}$, was more accurately predicted by the spatially distributed model (Figure 9). Indeed, the lumped model led to an average absolute error that is more than two times higher than the absolute error of the spatially distributed model. The lumped model systematically overestimated the peak breach discharge, while the spatially distributed model sometimes underpredicted the peak lateral discharge. The discharge coefficient formulations that were tested exhibited a rather erratic behavior with respect to the different tests. Overall, the results of the different formulations exhibited greater variability compared to the previous cases with fixed openings and no formulation outperformed consistently the others when coupled with either of the two models.

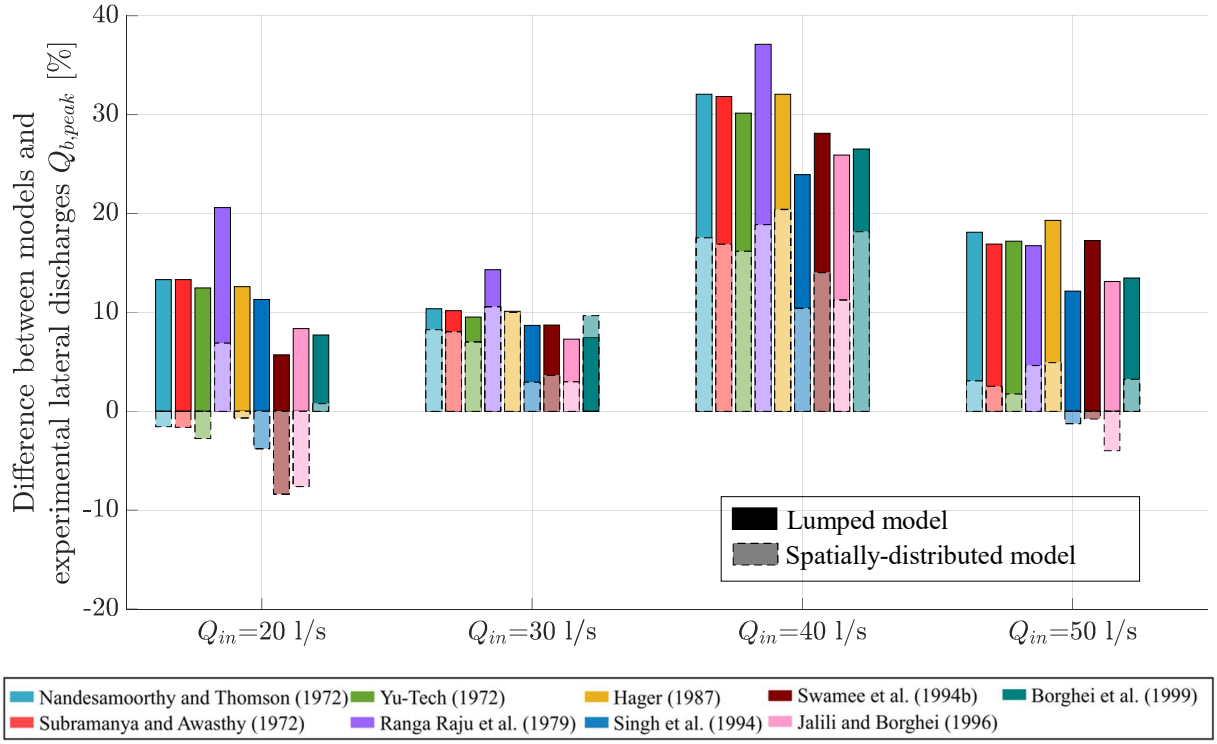


Figure 9. Comparison between modelled and measured breach peak discharges, $Q_{b,peak}$, for experiments of Rifai et al. (2017).

5. Discussion

The estimation of the discharge coefficient in weir equations has typically a large empirical component that depends on the flow conditions of the laboratory experiments from which it was developed. As such, weir equations can be confidently applied in data ranges for which they have been calibrated, but their performance outside these calibration ranges is uncertain. Figure 10 shows how the range of the experimental data used in this study relates to the data ranges in which the semi-empirical equations for the estimation of C_d were developed. It is evident that in some cases, some formulations were used outside of their development range, which may be a reason for their deteriorating performance.

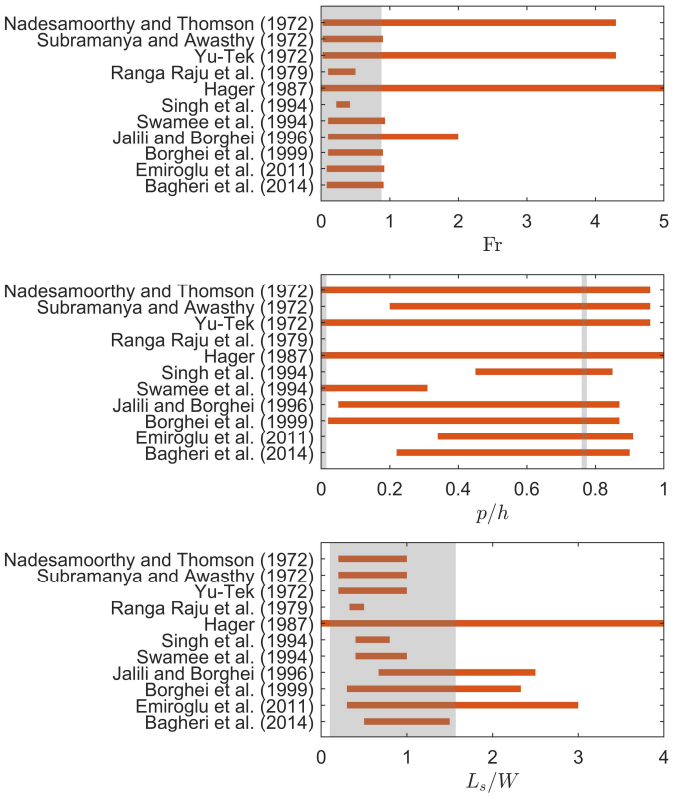


Figure 10. Data ranges in which the semi-empirical discharge coefficient equations were developed. The grey shaded areas denote the ranges of the experimental data from the static side openings used in this study, i.e., the experiments of Rifai et al. (2017) are excluded. The formulation of Hager (1987) covers the whole range because it was developed based on theoretical arguments.

Ranga Raju et al. (1979) and Singh et al. (1994) used a 90° side channel as the side weir. The hinterland was thus confined as it corresponded to a relatively narrow perpendicular channel. In these cases, the predicted value of C_d should be underestimated compared to the case when the hinterland is not confined. This trend is very obvious for Singh et al. (1994) when considering a fixed breach geometry and a sharp crest. This formula performs better for Rifai et al. (2017), probably because there is a broad-crested weir in this case, which confines a bit the flow when it goes through the breach.

The implementation of the different models using the data of Roger et al. (2009) and Mignot et al. (2020) showed that increasing the model complexity, i.e., switch from a lumped to a discretized hydraulic description, does not always improve the accuracy of the side weir equation. This is the case not only when comparing our zero-dimensional lumped model to our one-dimensional model, but also when comparing to more detailed 2D models that Roger et al. (2009) used to simulate their experiments. A possible explanation for this result is that the flow near dike breaches has 3D features (Michelazzo et al., 2015) that can be more easily parameterized when modelled with a lower dimensionality model. Despite their simplicity, lumped models have exhibited a versatile and reliable behavior in a broad range of hydraulic applications, e.g., in interactions between

surface flows and drainage systems (Kitsikoudis et al., 2021), and as such it is not surprising that such modelling tools perform relatively well when properly calibrated.

The importance of proper calibration of the discharge coefficient in lumped models becomes evident from the data from the experiments of Rifai et al. (2017), where the lumped model was outperformed by the spatially distributed model for the majority of discharge coefficient equations. The different discharge coefficient equations have been calibrated with data from laboratory experiments where the crest of the side weir was horizontal. In the experiments of Rifai et al. (2017) the dike breach, which was considered as a broad-crested side weir, exhibited an irregular cross-section at the different time-steps due to the spatial distribution of the erosion process induced by turbulent flow. In such cases with variability of the bed level across the dike breach, the variations in flow depth will affect the discharge distribution over the dike breach (Michelazzo et al., 2018). In the experiments of Rifai et al. (2017), the lumped model was used by considering a straight weir crest at the 15th percentile of the elevation of the scanned points within the dike breach. This simplification led to a considerable overestimation of the breach peak discharge by the lumped model. This inaccuracy could also be attributed to sharp variations of the breach dimensions owed to sudden breach side slope failures, which instantaneously impact the value of C_d when using the empirical formulas. These sharp variations of C_d may also be the reason why the evolution of Q_b is smoother in the experiments compared to the modelling. The spatially distributed model predicted much better the peak discharge in the dynamic dike breach (Figure 9), highlighting the importance of spatial discretization in cases with large spatial variability of the side opening.

For accurate hazard predictions related to fluvial dike breaching, a discretized model should be preferred. Nevertheless, the zero-dimensional model is more conservative as it tends to systematically overestimate the breach discharge (Figure 7 and Figure 9). To be on the safe side while optimizing the prediction accuracy, a trade-off consists in combining the one-dimensional model with a formulation that overestimates the breach discharge coefficient, e.g., the one proposed by Ranga Raju et al. (1979) or Borghei et al. (1999). It is also advised to test different formulations for the discharge coefficient to obtain an envelope curve, allowing for better uncertainty assessment.

6. Conclusions

Breaches in river dikes can pose a hazard of intense flooding to nearby areas. In such cases, flood risk assessment requires not only accurate modelling of the dike breach discharge towards the hinterland, but also fast modelling tools that can be used efficiently. To satisfy these requirements, in this study the flow through a breach in a dike with its axis parallel to a straight river reach was considered similar to the flow over a side weir. Eleven semi-empirical equations from the literature were tested for the parameterization of the discharge coefficient of the side weir equation, which was coupled to either a lumped model (zero-dimensional) or a spatially distributed model (one-dimensional). These models were tested in cases with increasing complexity: from cases with a fixed side opening with zero and non-zero crest height to a dynamic dike breach that evolves over time.

The performance of the different equations for the parametrization of the discharge coefficient of the side weir equation varied with the different cases, highlighting the empirical nature of most of them. The equations of Subramanya and Awasthy (1972) and Nadesamoorthy and Thomson (1972) performed best for a fixed opening with zero crest height while the equation of Singh et al. (1994) performed best for a fixed opening with non-zero crest height. For the cases with the dynamic dike breach, several formulations derived good results and there was no formula that was consistently superior. Overall, the modelling results were less accurate when transitioning from the fixed opening cases of zero to non-zero crest height and to the cases with a dynamic breach. Despite its simplicity, the lumped model generated better results than the spatially distributed model for many cases with fixed openings, but the spatially distributed model was more accurate in the dynamic breach cases. This highlights the importance of spatial discretization in cases where the dike breach exhibits irregular geometry. This model type should be selected when assessing hazard related to real-world dike breaching. For the sake of safety, it should be coupled with a formulation that overestimates the breach discharge coefficient, such as the one of Ranga Raju et al. (1979) or Borghei et al. (1999). Uncertainty may also be considered by creating an envelope curve for the breach discharge by testing different formulations for the breach discharge.

References

Ahadiyan, J., Bahmanpouri, F., Adeli, A., Gualtieri, C., Khoshkonesh, A., 2022. Riprap Effect on Hydraulic Fracturing Process of Cohesive and Non-cohesive Protective Levees. Water Resources Management 36, 625–639.

Al-Hafidh, I.A.I., Calamak, M., LaRocque, L.A., Chaudhry, M.H., Imran, J., 2022. Experimental Investigation of Flood Management by an Instantaneous Levee Breach. Journal of Hydraulic Engineering 148, 04021056.

Bagheri, S., Kabiri-Samani, A.R., Heidarpour, M., 2014. Discharge coefficient of rectangular sharp-crested side weirs, Part I: Traditional weir equation. Flow Measurement and Instrumentation 35, 109–115.

Bomers, A., Schielen, R.M.J., Hulscher, S.J.M.H., 2019. Consequences of dike breaches and dike overflow in a bifurcating river system. Natural Hazards 97, 309–334.

Borghei, S.M., Jalili, M.R., Ghodsian, M., 1999. Discharge Coefficient for Sharp-Crested Side Weir in Subcritical Flow. Journal of Hydraulic Engineering 125, 1051–1056.

D'Oria, M., Maranzoni, A., Mazzoleni, M., 2019. Probabilistic Assessment of Flood Hazard due to Levee Breaches Using Fragility Functions. Water Resources Research 55, 8740–8764.

ASCE/EWRI Task Committee on Dam/Levee Breaching, 2011. Earthen Embankment Breaching. Journal of Hydraulic Engineering 137, 1549–1564.

Dazzi, S., Vacondio, R., Mignosa, P., 2019. Integration of a Levee Breach Erosion Model in a GPU-Accelerated 2D Shallow Water Equations Code. Water Resources Research 55, 682–702.

Dewals, B., Kitsikoudis, V., Mejía-Morales, M.A., Archambeau, P., Mignot, E., Proust, S., Erpicum, S., Pirotton, M., Paquier, A., 2023. Can the 2D shallow water equations model flow intrusion into buildings during urban floods? Journal of Hydrology 619, 129231.

Echeverribar, I., Morales-Hernández, M., Brufau, P., Garcıa-Navarro, P., 2019. Use of internal boundary conditions for levees representation: application to river flood management. Environmental Fluid Mechanics 19, 1253–1271.

Elalfy, E., Tabrizi, A.A., Chaudhry, M.H., 2018. Numerical and Experimental Modeling of Levee Breach Including Slumping Failure of Breach Sides. Journal of Hydraulic Engineering 144, 04017066.

Emiroglu, M.E., Agaccioglu, H., Kaya, N., 2011. Discharging capacity of rectangular side weirs in straight open channels. Flow Measurement and Instrumentation 22, 319–330.

Ferrari, A., Dazzi, S., Vacondio, R., Mignosa, P., 2020. Enhancing the resilience to flooding induced by levee breaches in lowland areas: a methodology based on numerical modelling. Natural Hazards and Earth System Sciences 20, 59–72.

Haer, T., Husby, T.G., Botzen, W.J.W., Aerts, J.C.J.H., 2020. The safe development paradox: An agent-based model for flood risk under climate change in the European Union. Global Environmental Change 60, 102009.

Hager, W.H., 1987. Lateral Outflow Over Side Weirs. Journal of Hydraulic Engineering 113, 491–504.

Jalili, M.R., Borghei, S.M., 1996. Discussion: Discharge Coefficient of Rectangular Side Weirs. Journal of Irrigation and Drainage Engineering 122, 132–132.

Kakinuma, T., Shimizu, Y., 2014. Large-Scale Experiment and Numerical Modeling of a Riverine Levee Breach. Journal of Hydraulic Engineering 140, 04014039.

Kamrath, P., Disse, M., Hammer, M., Köngeter, J., 2006. Assessment of Discharge through a Dike Breach and Simulation of Flood Wave Propagation. Natural Hazards 38, 63–78.

Kerger, F., Archambeau, P., Erpicum, S., Dewals, B.J., Pirotton, M., 2011. A fast universal solver for 1D continuous and discontinuous steady flows in rivers and pipes. International Journal for Numerical Methods in Fluids 66, 38–48.

Kitsikoudis, V., Becker, B.P.J., Huismans, Y., Archambeau, P., Erpicum, S., Pirotton, M., Dewals, B., 2020. Discrepancies in Flood Modelling Approaches in Transboundary River Systems: Legacy of the Past or Well-grounded Choices? Water Resources Management 34, 3465–3478.

Kitsikoudis, V., Erpicum, S., Rubinato, M., Shucksmith, J.D., Archambeau, P., Pirotton, M., Dewals, B., 2021. Exchange between drainage systems and surface flows during urban flooding: Quasi-steady and dynamic modelling in unsteady flow conditions. Journal of Hydrology 602, 126628.

Li, X., Kitsikoudis, V., Mignot, E., Archambeau, P., Pirotton, M., Dewals, B., Erpicum, S., 2021. Experimental and Numerical Study of the Effect of Model Geometric Distortion on Laboratory Modeling of Urban Flooding. Water Resources Research 57, e2021WR029666.

Madsen, H., Lawrence, D., Lang, M., Martinkova, M., Kjeldsen, T.R., 2014. Review of trend analysis and climate change projections of extreme precipitation and floods in Europe. Journal of Hydrology 519, 3634–3650.

Maranzoni, A., D'Oria, M., Mazzoleni, M., 2022. Probabilistic Flood Hazard Mapping Considering Multiple Levee Breaches. Water Resources Research 58, e2021WR030874.

Michelazzo, G., Oumeraci, H., Paris, E., 2015. Laboratory Study on 3D Flow Structures Induced by Zero-Height Side Weir and Implications for 1D Modeling. Journal of Hydraulic Engineering 141, 04015023.

Michelazzo, G., Oumeraci, H., Paris, E., 2018. New Hypothesis for the Final Equilibrium Stage of a River Levee Breach due to Overflow. Water Resources Research 54, 4277–4293.

Mignot, E., Camusson, L., Riviere, N., 2020. Measuring the flow intrusion towards building areas during urban floods: Impact of the obstacles located in the streets and on the facade. Journal of Hydrology 583, 124607.

Nadesamoorthy, T., Thomson, A., 1972. Discussion of "Spatially Varied Flow over Side-Weirs. Journal of the Hydraulics Division 98, 2234–2235.

Onda, S., Hosoda, T., Jacimovic, N.M., Kimura, I., 2019. Numerical modelling of simultaneous overtopping and seepage flows with application to dike breaching. Journal of Hydraulic Research 57, 13–25.

Orlandini, S., Moretti, G., Albertson, J.D., 2015. Evidence of an emerging levee failure mechanism causing disastrous floods in Italy. Water Resources Research 51, 7995–8011.

Ranga Raju, K.G., Gupta, S.K., Prasad, B., 1979. Side Weir in Rectangular Channel. Journal of the Hydraulics Division 105, 547–554.

Rifai, I., El Kadi Abderrezzak, K., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., Dewals, B., 2018. Floodplain Backwater Effect on Overtopping Induced Fluvial Dike Failure. Water Resources Research 54, 9060–9073.

Rifai, I., El Kadi Abderrezzak, K., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., Dewals, B., 2019. Flow and detailed 3D morphodynamic data from laboratory experiments of fluvial dike breaching. Scientific Data 6.

Rifai, I., El Kadi Abderrezzak, K., Hager, W.H., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., Dewals, B., 2021. Apparent cohesion effects on overtopping-induced fluvial dike breaching. Journal of Hydraulic Research 59, 75–87.

Rifai, I., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., El Kadi Abderrezzak, K., Dewals, B., 2017. Overtopping induced failure of noncohesive, homogeneous fluvial dikes. Water Resources Research 53, 3373–3386.

Roger, S., Dewals, B.J., Erpicum, S., Schwanenberg, D., Schüttrumpf, H., Köngeter, J., Pirotton, M., 2009. Experimental and numerical investigations of dike-break induced flows. Journal of Hydraulic Research 47, 349–359.

Schmitz, V., Arnst, M., El Kadi Abderrezzak, K., Pirotton, M., Erpicum, S., Archambeau, P., Dewals, B., 2023a. Global Sensitivity Analysis of a Dam Breaching Model: To Which Extent Is Parameter Sensitivity Case-Dependent? Water Resources Research 59, e2022WR033894.

Schmitz, V., Erpicum, S., Abderrezzak, K. El kadi, Rifai, I., Archambeau, P., Pirotton, M., Dewals, B., 2021. Overtopping-Induced Failure of Non–Cohesive Homogeneous Fluvial Dikes: Effect of Dike Geometry on Breach Discharge and Widening. Water Resources Research 57.

Schmitz, V., Rifai, I., Kheloui, L., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., El Kadi Abderrezzak, K., Dewals, B., 2023b. Main channel width effects on overtopping-induced non-cohesive fluvial dike breaching. Journal of Hydraulic Research 61, 601–610.

Schmocker, L., Hager, W.H., 2012. Plane dike-breach due to overtopping: effects of sediment, dike height and discharge. Journal of Hydraulic Research 50, 576–586.

Shustikova, I., Neal, J.C., Domeneghetti, A., Bates, P.D., Vorogushyn, S., Castellarin, A., 2020. Levee Breaching: A New Extension to the LISFLOOD-FP Model. Water 12, 942.

Singh, R., Manivannan, D., Satyanarayana, T., 1994. Discharge Coefficient of Rectangular Side Weirs. Journal of Irrigation and Drainage Engineering 120, 814–819.

Stilmant, F., Pirotton, M., Archambeau, P., Roger, S., Erpicum, S., Dewals, B., 2013. Dike-break induced flows: a simplified model. Environmental Fluid Mechanics 13, 89–100.

Subramanya, K., Awasthy, S.C., 1972. Spatially Varied Flow over Side-Weirs. Journal of the Hydraulics Division 98, 1–10.

Swamee, P.K., Pathak, S.K., Mohan, M., Agrawal, S.K., Ali, M.S., 1994. Subcritical Flow over Rectangular Side Weir. Journal of Irrigation and Drainage Engineering 120, 212–217.

Tadesse, Y.B., Fröhle, P., 2020. Modelling of Flood Inundation due to Levee Breaches: Sensitivity of Flood Inundation against Breach Process Parameters. Water 12, 3566.

Tellman, B., Sullivan, J.A., Kuhn, C., Kettner, A.J., Doyle, C.S., Brakenridge, G.R., Erickson, T.A., Slayback, D.A., 2021. Satellite imaging reveals increased proportion of population exposed to floods. Nature 596, 80–86.

Viero, D.P., D'Alpaos, A., Carniello, L., Defina, A., 2013. Mathematical modeling of flooding due to river bank failure. Advances in Water Resources 59, 82–94.

Vorogushyn, S., Merz, B., Apel, H., 2009. Development of dike fragility curves for piping and micro-instability breach mechanisms. Natural Hazards and Earth System Sciences 9, 1383–1401.

Vorogushyn, S., Merz, B., Lindenschmidt, K.-E., Apel, H., 2010. A new methodology for flood hazard assessment considering dike breaches. Water Resources Research 46, W08541.

Wei, H., Yu, M., Wang, D., Li, Y., 2016. Overtopping breaching of river levees constructed with cohesive sediments. Natural Hazards and Earth System Sciences 16, 1541–1551.

Wu, S., Yu, M., Wei, H., Liang, Y., Zeng, J., 2018. Non-symmetrical levee breaching processes in a channel bend due to overtopping. International Journal of Sediment Research 33, 208–215.

Wu, W., 2013. Simplified Physically Based Model of Earthen Embankment Breaching. Journal of Hydraulic Engineering 139, 837–851.

Yu, M., Wei, H., Liang, Y., Zhao, Y., 2013. Investigation of non-cohesive levee breach by overtopping flow. Journal of Hydrodynamics 25, 572–579.

Yu-Tek, L., 1972. Discussion of "Spatially Varied Flow over Side-Weirs. Journal of the Hydraulics Division 98, 2047–2048.

Chapter 5

How Does Machine Learning Compare to Hybrid Approaches for Predicting Fluvial Dike Breach Discharge?

This chapter corresponds to the manuscript "How does machine learning compare to hybrid approaches for predicting fluvial dike breach discharge?" by V. Schmitz, R. Vandeghen, S. Erpicum, M. Pirotton, P. Archambeau, and B. Dewals, submitted in 2024 to Water Resources Research. The PhD candidate developed the methodology, implemented the numerical models, carried out the analytical developments, and was the primary contributor to the manuscript and figures.

How Does Machine Learning Compare to Hybrid Approaches for Predicting Fluvial Dike Breach Discharge?

V. Schmitz[1], R. Vandeghen[2], S. Erpicum[1], M. Pirotton[1], P. Archambeau[1], and B. Dewals[1]

[1]*Research Group of Hydraulics in Environmental and Civil Engineering (HECE), University of Liège, Liège, Belgium.*

[2]*Department of Electrical Engineering and Computer Science, University of Liège, Belgium*

Corresponding author: Vincent Schmitz (V.Schmitz@uliege.be)

Key Points:

- Comparison between machine learning algorithms, a new analytical approach, and empirical formulas for predicting the breach discharge.
- The extremely randomized trees algorithm leads to particularly accurate results, both when evaluated inside and outside its training space.
- The new analytical model enhanced with machine learning is a promising tool.

Chapter 5: Machine Learning versus Hybrid Approaches

ABSTRACT

The breaching of a fluvial dike can have devastating consequences for flooded areas. Accurate prediction of breach discharge is crucial to enhance preventive measures and emergency planning. So far, most studies have relied on empirical formulas developed for simplified configurations, which fail at capturing the complexity of a real dike breaching event. In this context, machine learning (ML) models offer promising predictive capabilities. This study focuses on decision-tree-based models, trained on 43 dike breaching laboratory tests, and compares three predictive approaches: (1) direct prediction using ML, (2) direct prediction using empirical formulas developed for simplified configurations, and (3) a new analytical approach with an empirical parameter computed using ML or empirical regressions. Additionally, a ML-based corrective term is applied to approaches (2) and (3) to further improve their accuracy. The extremely randomized trees algorithm demonstrates particularly high accuracy when predicting the breach discharge (approach 1), both in interpolation and extrapolation, i.e., when tested inside or outside the ML training space. The ML-based corrective term significantly improves results associated with the analytical model (approach 2) and the empirical formulas (approach 3). The definition of the dike breach invert level, i.e., one of the model inputs, was varied but it had little influence on the models' performance. Expanding the experimental dataset by leading new laboratory or field tests would enhance the accuracy and reliability of the ML models. Future studies may explore alternative ML models, including physics-guided deep learning algorithms, which, although in their early stages, hold substantial potential for future applications.

1. Introduction

With the increase in extreme meteorological events, growing urbanization in hinterlands and aging infrastructure, fluvial dikes are becoming more prone to breaches while their potential impact increases substantially (Flynn et al., 2022). As a result, developing predictive dike breach models has become critically important for ensuring safe land-use planning and effective emergency response strategies. Among numerical models, spatially non-discretized models offer a good approximation of the results in a very short time (ASCE/EWRI Task Committee, 2011).

Dike breaching models are required to accurately reproduce the breach discharge and the resulting dike erosion. However, the capability of existing experimental and analytical side weir discharge formulas in predicting the flow through a real dike breach is limited (Schmitz et al., 2024). Those formulas mostly rely on laboratory setups or analytical assumptions that do not represent breaching events faithfully, e.g., prismatic sharp-crested weir geometry (Hager, 1987; Borghei et al., 1999; Emiroglu et al., 2011; Bagheri et al., 2014; Elalfy et al., 2018; Lee, 2019) or narrow constraint hinterland (Ranga Raju et al., 1979; Cheong, 1991; Haddadi and Rahimpour, 2012; Ibrahim et al., 2022; Wang et al., 2024).

There is no straightforward solution to tackle those limitations as dike breaching events are characterized by complex features, such as 3D flow patterns close to the side breach (Neary et al., 1999; Michelazzo et al., 2015; Cheng et al., 2022; Chowdhury et al., 2022) and highly non-

uniform breach geometry (Rifai et al., 2019). Oversimplified experimental setups and strong theoretical assumptions fail at capturing those features.

Within this context, supervised machine learning techniques may be of great help as they are able to replicate complex relationships when properly trained. Specifically, regression machine learning algorithms are considered when predicting a continuous target variable, i.e., the dike breach discharge, based on a limited number of features, i.e., upstream flow characteristics, channel and dike geometry, and breach morphology (Haghbin and Sharafati, 2022). Many machine learning techniques have been used in previous works dedicated to the analysis of the flow through a side weir, including linear methods (Jamei et al., 2021), decision-tree-based techniques (Hameed et al., 2021), support vector regressions (Balahang and Ghodsian, 2023; Li et al., 2024), genetic algorithms (Roushangar et al., 2016; Azimi et al., 2017), and neural networks (Parsaie, 2016; Ebtehaj et al., 2018).

However, most works considered prismatic weirs with idealized geometries, e.g., triangular (Jamei et al., 2021; Balahang and Ghodsian, 2023), rectangular (Parsaie, 2016; Azimi et al., 2017; Ebtehaj et al., 2018; Hameed et al., 2021), trapezoidal (Roushangar et al., 2016), or semi-circular (Li et al., 2024). Also, they often focused on the determination of the discharge coefficient, whose definition may vary from one experiment to another. Di Bacco and Scorzini (2019) highlighted that this approach could lead to inconsistent datasets and biased regression models.

The choice of a specific machine learning model depends on the data on hand (e.g., type, sample size, distribution, relationships…) and the user's objective (accuracy, speed, interpretability, ease of use…). In this work, multiple linear and decision-tree-based models are considered. Linear methods are intuitive, easy to use and computationally efficient. Methods based on decision trees are also particularly appropriate due to their simple use, i.e., very few hyperparameters, and ability to handle non-linear dependencies. Data collected from experimental tests (Rifai et al., 2019; Schmitz et al., 2021) are used to feed the machine learning models.

The objective of this work is to evaluate the capability of linear and decision-tree-based machine learning methods to predict the dike breach discharge based on the hydrodynamic data in the main channel and the breach geometry, and to confront them with classical empirical formulas. The predictive capability of the machine learning models is assessed in interpolation and extrapolation, i.e., inside and outside the training set space, respectively. Additionally, a new semi-analytical approach for the determination of the breach discharge is introduced. It relies on analytical developments and involves a single fitting parameter, whose value is computed either using machine learning models or empirical formulas derived from the experimental dataset on hand. Finally, the influence of the definition of features used by machine learning models is discussed, especially the impact of those that characterize the dike breach geometry.

Section 2 presents the machine learning models used in this work. Experimental data used to train and validate them are also introduced in this section. Section 3 introduces a new semi-analytical model for the determination of the breach discharge. Then, it shows how the

breach discharge predicted by the different machine learning models compares with the experimental measurements. The performance and limitations of the machine learning models compared to the semi-analytical model and empirical formulas are discussed in Section 4. This section also investigates the influence of the definition of the breach geometry on the numerical results. Finally, conclusions are drawn in Section 5.

2. Methods and data

This section starts with a brief description of the machine learning techniques used in this work (Section 2.1). The experimental data and the way datasets are generated to feed the machine learning models are introduced in Sections 2.2 and 2.3. Finally, a new analytical model for the breach discharge is detailed in Section 3.1.

2.1. Machine learning techniques

In this work, two types of machine learning (ML) techniques were used to predict the dike breach discharge directly, or indirectly through a new analytical model (Section 3.1). In both cases, experimental data collected by Rifai et al. (2019) and Schmitz et al. (2021) were used to feed the models. ML methods providing easy interpretation and involving a few hyperparameters were considered. Specifically, we focused on multiple linear regression (MLR) and decision-tree-based regression methods readily available in *scikit-learn* library, i.e., a machine learning library developed in Python (Pedregosa et al., 2011).

The MLR techniques aim at modelling a target value as a linear combination of the features. The main advantages of this method are its simple interpretation and its limited computation cost. However, it fails to reproduce highly non-linear phenomena.

The decision tree regression technique predicts the target variable by relying on simple decision rules, i.e., boolean logic. In this case, the target variable takes a constant value as long as the boolean results of the decision rules are not modified. This leads to a piecewise constant approximation of the target variable. This algorithm mainly relies on three hyperparameters:

1. the minimum number of training points required to build a leaf, i.e., the minimum number of training points that fulfil all conditions of an entire decision chain;
2. the minimum number of training points required to split an internal node;
3. the tree maximum depth, i.e., the maximum number of decision rules contained in a decision chain.

Although this method is simple to use and to interpret, it tends to overfit the training data and returns highly discontinuous predictions for the target variable.

To tackle those limitations, more advanced algorithms were developed, e.g., random forests (Breiman, 2001) and extremely randomized trees (Geurts et al., 2006). Random forests rely on multiple decision trees, each built from a bootstrap sample of the training set, i.e., a sample drawn with replacement from the training set. The final prediction corresponds to the average prediction of all individual trees. In extremely randomized trees, the whole training set is used to build each individual tree, i.e., no bootstrapping, but randomness is added when generating the decision rules. Instead of selecting the most discriminative threshold for each individual

feature when defining a new decision rule, random thresholds are assigned to each feature. This step is repeated *k* times for the considered split, with *k* the number of features. Finally, the best of these *k* randomly generated threshold sets is selected as the decision rule. The procedure is thus used each time a decision rule is generated in each decision tree. The final prediction is obtained by averaging the prediction of each individual decision tree. Random forests and extremely randomized trees considerably reduce sensitivity to the training dataset. In Section 3, the predictive capability of the different machine learning techniques is compared.

When building ML models, experimental data are required. Those are generally divided into three datasets, called training, validation, and test sets. The training set is used to fit the model parameters so that it minimizes the predictive error on the target variable. The validation set is then used to select the best model hyperparameters, such as the tree structures in decision-tree-based models. Finally, the test set serves as a benchmark for evaluating the model's ability to predict the target variable. To allow for a relevant and fair evaluation of the ML model performance, the composition of each set should depend on the experimental data specificities, as detailed in Section 2.3.

2.2. Experimental data

Data used in this work were collected from laboratory experiments. Most of them were previously presented and discussed by Rifai et al. (2019) and Schmitz et al. (2021). In total, 43 tests are considered here (Table 1).

The experimental setup consists of a horizontal trapezoidal straight main channel (10 × 1 m) (Figure 1). A 3-m-long trapezoidal dike built with uniform sand of median diameter d_{50} = 1 mm was erected along the right side of the main channel. A horizontal floodplain was present beside the dike (4.3 × 2.5 m). Its bottom and the main channel bottom were at the same level. The main channel and floodplain were covered with an impermeable whitewash coating to ensure roughness continuity between the flume, floodplain, and sand dike (Rifai et al., 2017). A drainage system was installed at the dike bottom to control the seepage through the dike body. A perforated plate, followed by a reservoir, was placed at the downstream extremity of the main channel. The perforated plate was adjusted such that, at the beginning of each test, the water level corresponded to the dike crest elevation. To trigger breaching, a 2-cm-deep and 10-cm-wide initial notch was created in the dike crest, 0.8 m from its upstream end.

Three ultrasonic sensors measured the water level in the main channel, one in the outflow tank, and one in the drainage tank (accuracy of ±1 mm). A spatially averaged water level was computed based on the measurements from the three sensors located in the main channel. This value was considered as the main channel water level for the rest of the analysis. The inflow discharge Q_{in} was measured using an electromagnetic flowmeter (accuracy of ± 0.75 l/s). The outflow discharge was derived from the discharge passing through a V-notch weir, i.e., by means of its rating curve, and mass balance in the outflow tank. The drainage discharge Q_d was estimated from the evolution of the water level measured in the drainage basin. The breach discharge Q_b was determined from mass balance in the main channel.

The 3D breach geometry was recorded for 13 tests using the laser profilometry technique presented by Rifai et al. (2020), with a resolution of 1cm x 1cm. Alternatively, the breach widening in the other tests was monitored using a simplified version of this technique: the laser sheet was not rotated, which enabled to reliably monitor the breach widening at the crest level as the laser sheet was aligned with the dike crest, but not to produce 3D reconstructions of the breach evolving geometry.

Table 1. Experimental tests features. The target Froude number, F_{target}, is computed using the target inflow discharge, Q_{in}, and the water level equal to the dike crest. S_u and S_d stand for the dike slope on the main channel and floodplain side, respectively, while L_k is the dike crest width. The three tests corresponding to underlined values of the inlet discharge Q_{in} were repeated twice. The test corresponding to the bold underlined Q_{in} value was repeated three times. The stars depict tests for which the full 3D breach geometry was recorded.

S_u (-)	S_d (-)	L_k (m)	Q_{in} (l/s)			F_{target}		
2.0	2.0	0.15	<u>25</u>	<u>40</u>	55	<u>0.083</u>	<u>0.133</u>	0.183
1.5	1.5	0.15	25	40	55	0.071	0.114	0.157
1.5	2.0	0.15	25	40	55	0.071	0.114	0.157
2.0	1.5	0.15	25	40	<u>55</u>	0.083	0.133	<u>0.183</u>
2.0	2.5	0.15	25	40	55	0.083	0.133	0.183
2.0	3.0	0.15	25	40	55	0.083	0.133	0.183
2.0	3.0	0.30	25	-	55	0.083	-	0.183
2.0	2.0	0.00	25	40	55	0.083	<u>0.133</u>	0.183
2.0	2.0	0.30	25	40	55	0.083	0.133	0.183
2.0	2.0	0.60	-	-	55	-	-	0.183
2.0*	2.0*	0.1*	20	21	28	0.066	0.07	0.093
			30	31	**<u>40</u>**	0.100	0.103	**<u>0.133</u>**
			41	47	50	0.136	0.156	0.166
			51	55		0.169	0.182	

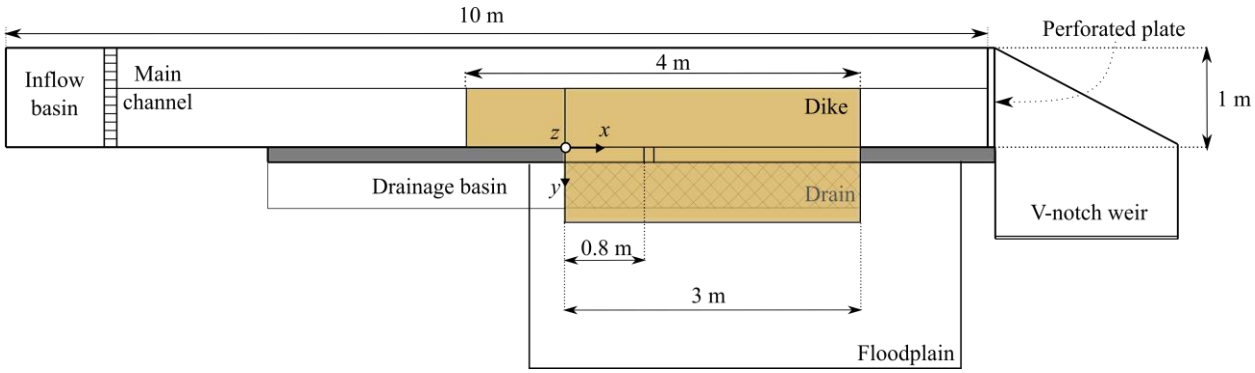


Figure 1. Laboratory setup used to collect the experimental data. Not at scale.

2.3. Datasets generation

As mentioned earlier, the target variable in this work is the dike breach discharge, Q_b. Its value is known to mainly depend on the main channel hydrodynamic state and the main channel, dike and breach shapes (Schmitz et al., 2021; Schmitz et al., 2023a). Within this context, five representative non-dimensional features were considered:

$$X = \left[z_{r,adim} = \frac{z_r}{w_{r,FS}}; \mathsf{F}; S_d; L_{k,adim} = \frac{L_k}{w_{r,FS}}; B_{top,adim} = \frac{B_{top}}{w_{r,FS}} \right], \quad (1)$$

With z_r the water level in the main channel, $w_{r,FS}$ the main channel width at the free surface, F the Froude number upstream from the breach, S_d the dike slope on the floodplain side, L_k the dike crest width, and B_{top} the breach top width. The breach top width was selected to represent the breach geometry as it is the only breach feature recorded during all experimental campaigns used in this work. During a breaching event, all parameters in Eq. (1) are time-dependent, except S_d and L_k.

In all tests, the breach top width was extracted at intervals that range between 1 and 60 seconds, depending on the test phase. In contrast, the hydrodynamic variables, i.e., the water level and the breach discharge, were obtained every 0.02 to 0.1 second, depending on the experiment. A linear variation was assumed between two successive geometry reconstructions to obtain a breach width associated to each recorded value of the hydrodynamic variables. Additionally, each experimental data point is assumed to be independent of the value of the parameters measured in the past, i.e., each data point is considered as an independent instance of a relation between the output and the input parameters. That way, several thousand data points were obtained for each laboratory experiment, leading to about 1.5 10^6 experimental data points overall.

Once the global dataset has been generated, it must be divided into three subsets: training, validation, and test sets. In the present case, no hyperparameters need to be fixed in the MLR algorithm, and only a few are required in the decision-tree-based models. For the latter, the hyperparameters values selected in this work are listed in Table S1 in Supplement. This set leads to satisfactory results while limiting overfitting (Pedregosa et al., 2011). Consequently, no validation set was generated in this work. Still, dispatching experimental data points between

the training and the test sets should be done carefully to avoid biased evaluation due to too similar training and test data. Data clustering helps gathering alike data subsets to avoid training and testing to be performed on data from the same subset. In this work, this should be done by avoiding training and testing ML models on data points emanating from identical or too similar laboratory experiments, i.e., data points encompassed in alike feature spaces. The laboratory tests could be differentiated by considering three main features, namely the main channel Froude number at overtopping initiation, F_{OT}, the non-dimensional dike crest width at overtopping initiation, $L_{k,adim}$, and the dike slope on the floodplain side, S_d. Figure 2 shows that 21 clusters were identified using the OPTICS algorithm (Ankerst et al., 1999) with a minimum cluster membership of two. Each cluster is labeled with a number between -9 and 11. Degenerated clusters containing only one experimental test were labeled with a red negative number. As a result, the test set corresponds to one cluster, or degenerated cluster, while the rest of the data forms the training set. Once the training set is defined, the features are standardized based on this dataset, i.e., mean removal and division by the standard deviation (Shanker et al., 1996).

To obtain a global performance score for each ML technique, the mean absolute relative error (MARE) on the target value is computed when successively considering each cluster as the test set, so that

$$MARE = \frac{1}{n}\sum_{i=1}^{n}\left|X_{exp}^{i} - X_{num}^{i}\right|, \tag{2}$$

with n the number of data points in the considered test set, X_{num}^{i} the value of the target variable obtained from the ML model fed with features associated to data point i, and X_{exp}^{i} the associated experimental value. The averaged value of MARE obtained for all the different test sets illustrates the global performance of the considered ML model.

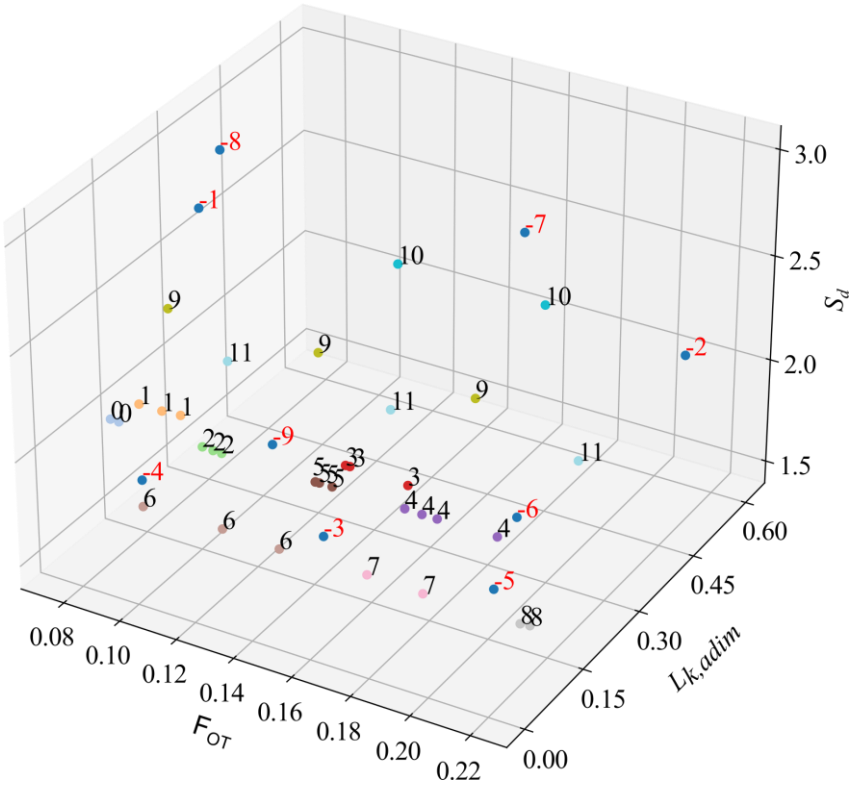


Figure 2. Clustering of the laboratory experiments. Tests contained in each cluster are listed in Table S2 in Supplement.

3. Results

3.1. New analytical model for the breach discharge

Instead of predicting the breach discharge by machine learning directly from the dike and flow parameters, an analytical expression can be derived based on sound assumptions and physics reasoning. Here, we consider shallow water flows and hydrostatic pressure distribution everywhere, and a subcritical flow in the main channel. If the backwater effect in the hinterland is negligible, a critical section must occur somewhere in the breach, so that

$$\mathrm{F}_b = \frac{U_c}{\sqrt{g\dfrac{A_c}{L_c}}} = 1, \tag{3}$$

where F_b and U_c stand for the Froude number and the mean flow velocity across the critical section in the breach, and A_c and L_c, the area and free surface width of this section.

The breach discharge Q_b can then be derived as

$$Q_b = U_c A_C = \sqrt{g\frac{A_c^3}{L_c}}\,. \tag{4}$$

To simplify A_c and L_c expressions while being consistent with experimental observations (Rifai et al., 2018), the breach cross section is approximated by a trapezoid with flat bottom and side slopes equal to the dike material repose angle.

Neglecting head losses between the main channel and the breach critical cross section, the energy conservation equation writes

$$(z_r - z_b) + \frac{U_r^2}{2g} = h_c + \frac{U_c^2}{2g}, \tag{5}$$

with z_b the breach invert level, U_r the mean flow velocity in the main channel upstream from the breach, h_c the critical water depth above the breach bottom, and g the gravity acceleration (Figure ***3***a). Injecting Eq. (4) into Eq. (5) yields

$$(z_r - z_b) + \frac{U_r^2}{2g} = h_c + \frac{A_c}{2L_c}. \tag{6}$$

Using a trapezoidal breach shape allows for explicitly linking A_c and L_c with the critical water depth h_c, leading to one equation with one single unknown. Nonetheless, laboratory experiments (Rifai et al., 2017), field tests (Kakinuma et al., 2013) and detailed 2D computations (Charrier, 2015) show that the flow velocity is highly non-uniform through the breach section, especially when the dike breach is large. Kakinuma and Shimizu (2014) observed that the flow concentrates close to the breach downstream extremity whilst recirculations take place at the upstream extremity. This suggests that the whole wetted breach area should not be considered when computing the critical breach section as only a reduced section of the breach conveys most of the flow. Similarly to Schmitz et al. (2023a), we introduce a parameter α that represents the fraction of the breach width that conveys most of the breach discharge and that should be considered as the breach critical section top width :

$$L_c = \alpha L_b = \alpha\left(b_{bot} + 2mh_c\right) \tag{7}$$

with L_b the total breach width at free surface level, b_{bot} the breach bottom width and m the breach side slopes (Figure 3b).

Note that shortly after the breaching onset, the breach is narrow, and the entire breach section is actively used by the flow. Stilmant et al. (2013) showed that the critical section at the breach could be curved, leading to a critical section width that is larger than the one of the breach straight cross-section. In this case, α might be larger than unity.

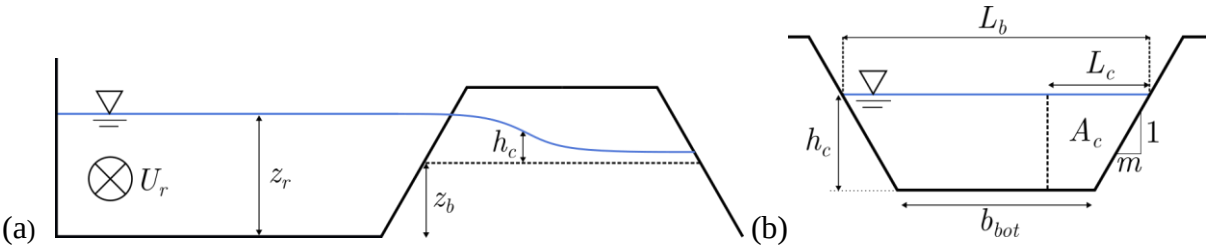


Figure 3. (a) Channel cross-section and (b) breach cross-section.

The critical height h_c is then derived from Eq. (6) as a function of α. For detailed computations, see the Appendix. From then, A_c and L_c are determined and injected in Eq. (4) to obtain the breach discharge.

Similarly, α can be obtained if the breach discharge and geometry are known. The breach bottom width, b_{bot}, and bottom elevation, z_b, can be derived from the experimental breach top width, i.e., the only breach feature recorded in all experiments. This is done by assuming a uniform erosion over the entire breach surface and constant side slopes equal to the repose angle of the dike wet material. That way, an "experimental" set of α values can be generated from the experimental data in hand (Figure 4). Machine learning models presented in Section 2.1 can then be applied to predict α, following the same procedure as the one presented for the breach discharge (Section 2.3). In this work, we limited the maximum value of α to 2. Without any limit, inconsistent extreme values are encountered at the beginning of the breaching event. During this period, the breach discharge is very small, and the relative error induced by experimental uncertainties becomes substantial, which largely impacts the value of α computed through the analytical model.

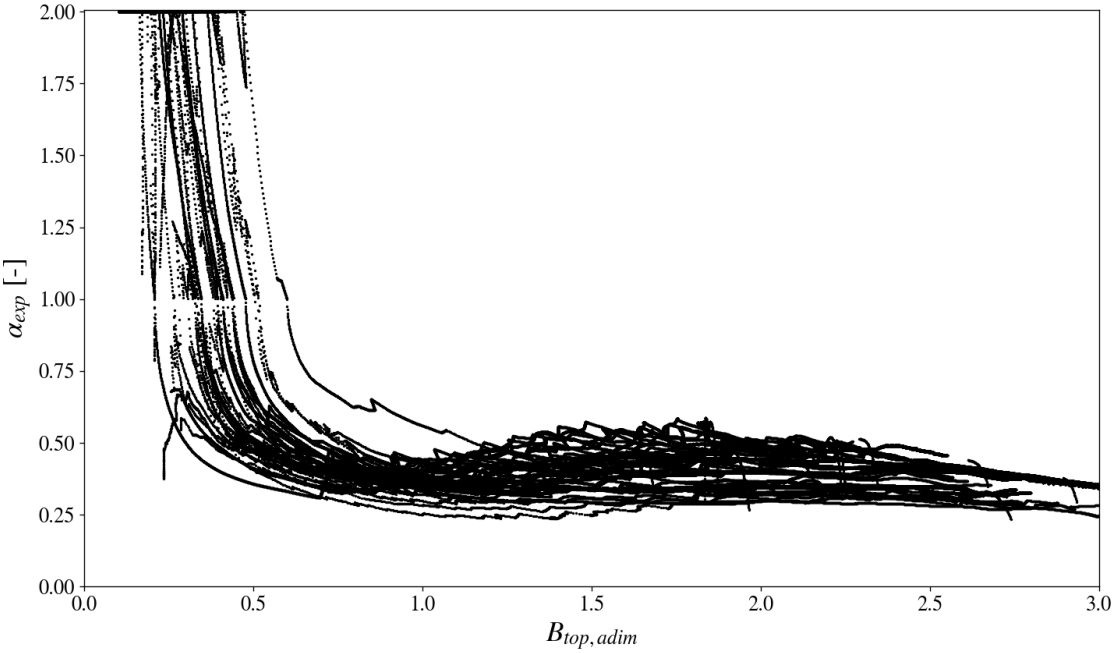


Figure 4. α values derived from experimental data as a function of the non-dimensional breach top width.

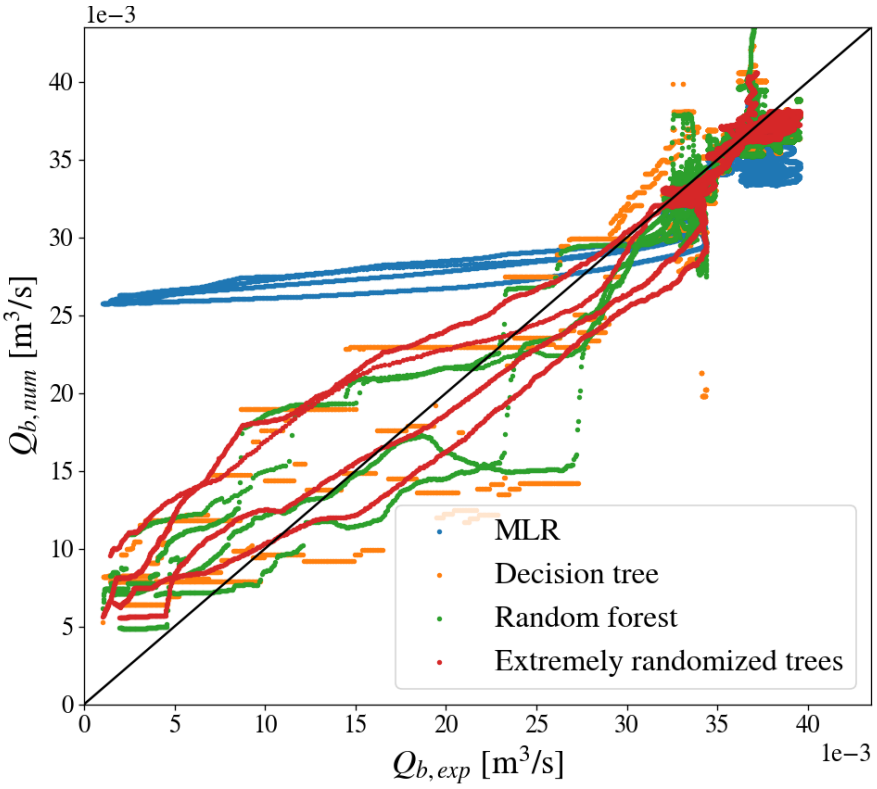
3.2. Comparison of Machine Learning techniques

In this section, the performance of the four ML techniques presented in Section 2.1 is compared for the direct evaluation of the breach discharge, or its indirect evaluation through the new analytical model presented in Section 3.1. As an illustration, Figure 5 presents the results generated when using cluster #5 as the test set (Figure 2) and the target variable being Q_b (Figure 5a), or α (Figure 5b). The breach discharge computed from the predicted values of α displayed in Figure 5b, $Q_b(\alpha)$, is presented in Figure 5c. Table 2 provides the averaged value of MARE on Q_b, α, and $Q_b(\alpha)$, as defined in Section 2.3.

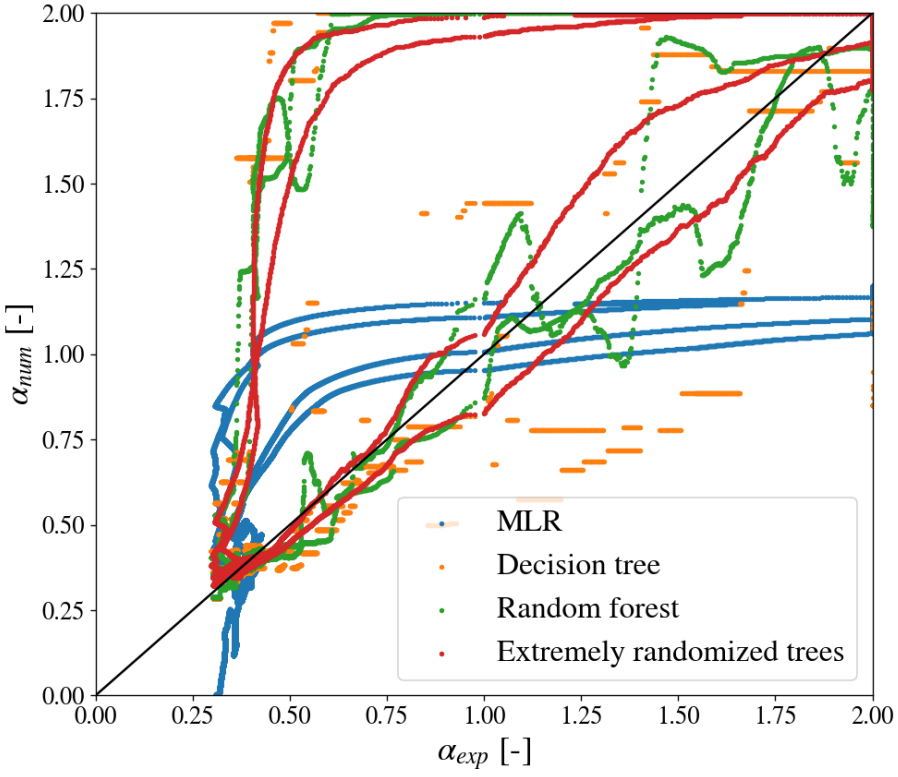
Overall, all ML techniques lead to more accurate results when predicting Q_b (Figure 5a) rather than α (Figure 5b), i.e., the averaged MARE on Q_b varies between 5% and 14%, whilst it ranges from 7% to 34% for α, depending on the ML method (Table 2). The error on the breach discharge derived from α, $Q_b(\alpha)$, grows further (Figure 5c and 5d). Depending on the ML model, the error on $Q_b(\alpha)$ increases by 12% to 27% compared to the one obtained for α (Table ***2***). This might be induced by an error amplification when converting α to $Q_b(\alpha)$ through the analytical model. In all cases, the discrepancies with the experimental data are particularly large when the breach discharge is small and α is large, i.e., shortly after overtopping initiation when the breach is still narrow. During this period, the breach geometry and the flow conditions evolve quickly (Figure 4). As experimental features evolve quickly, a slight temporal mismatch between hydraulic and breach morphologic variables may have a significant impact on the models' predictive capabilities (Figure 5b). Also, much fewer data points can be collected during this highly transient period, i.e., less training points available, which induces weaker model performance. Although large discrepancies seem to appear in Figure 5a, 5b and 5c, the overall models performances (Table 2) are mostly influenced by the accuracy on larger values of the breach discharge and smaller values of α, i.e., when the latter varies between 0.25 and 0.5 (Figure 4), which are mostly represented in the dataset.

The MLR model appears to perform poorly when predicting Q_b and α over the whole tested range. Table 2 confirms this trend, with an averaged MARE about 13% when predicting Q_b, and above 30% for both α and $Q_b(\alpha)$. Figure 6 highlights the particularly high variability of the model performance when varying the test set. The dike breaching process being highly non-linear, a linear model was indeed expected to provide low fidelity results.

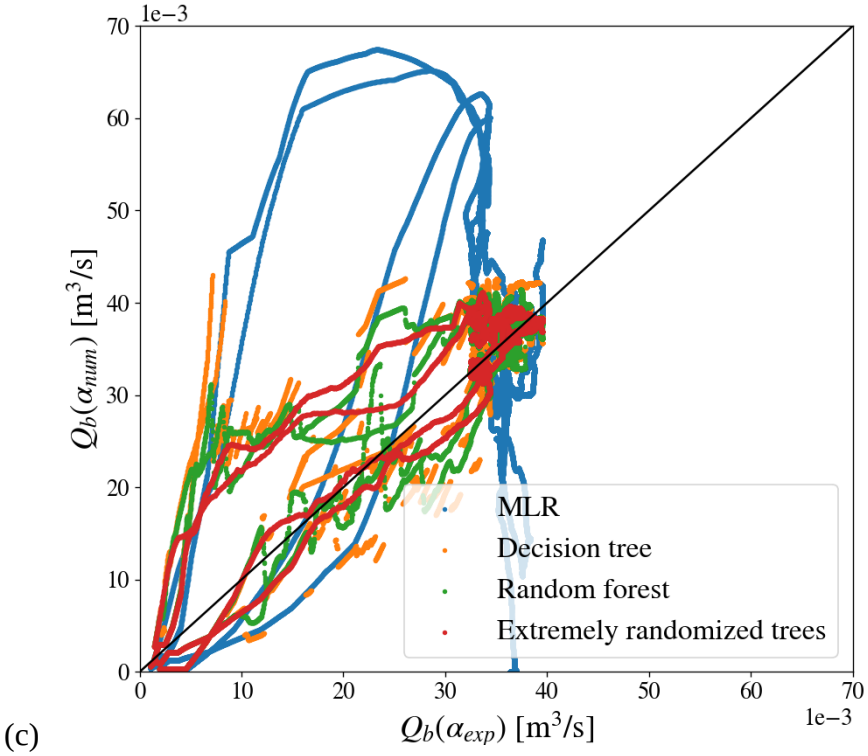
On the contrary, decision-tree-based ML techniques provide much better results. Among them, the extremely randomized trees model performs the best for all predicted variables, while having the smallest performance variability (Figure 6). Noteworthily, sharp discontinuities are observed with the basic decision-tree model. These discontinuities are still present, but to a smaller extent, when using the random forest model. Finally, they almost completely fade when the extremely randomized trees model is selected. The latter model presents the most physical behavior as the breach discharge should vary continuously. As a conclusion, the extremely randomized trees model should be preferred in the present case as it reduces the error on the target values while avoiding non-physical discontinuities in the results. This model will be the only one considered when discussing results in Section 4.


1e−3
40
35
30
25
20
15
10
5
0
$Q_{b,num}$ [m^3/s]
0
5
10
15
20
25
30
35
40
1e−3
$Q_{b,exp}$ [m^3/s]
MLR
Decision tree
Random forest
Extremely randomized trees

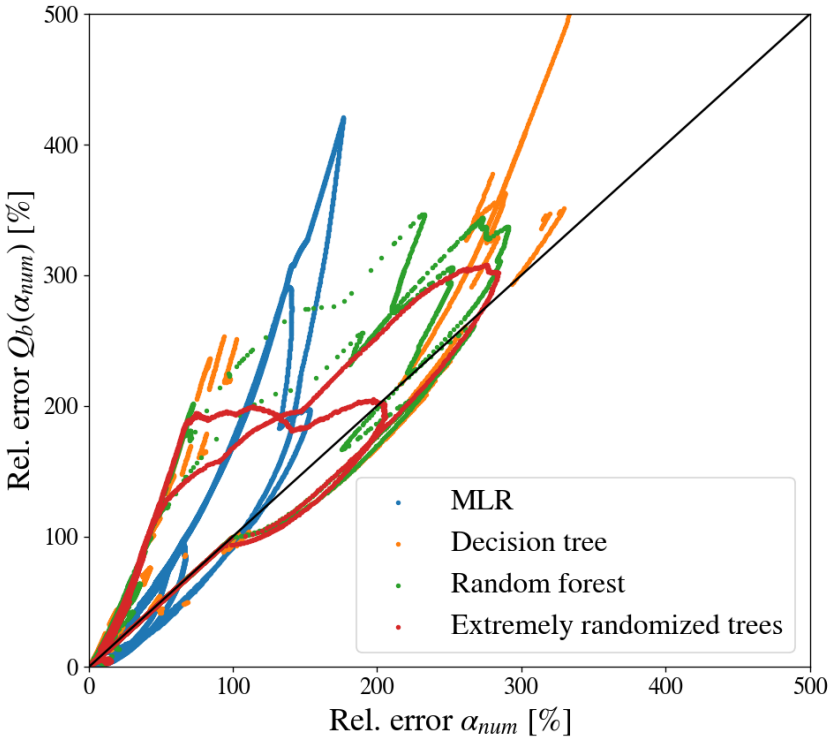
(a)


2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
α_{num} [-]
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
α_{exp} [-]
MLR
Decision tree
Random forest
Extremely randomized trees

(b)



(c)



(d)

Figure 5. Results predicted by the ML models for (a) the breach discharge, Q_b, (b) the efficient breach width parameter, α, and (c) the breach discharge computed from α, $Q_b(\alpha)$. In each case, the test set is cluster #5, which contains data from four laboratory experiments.

Table 2. Mean absolute relative error (MARE) [%] related to each ML model with Q_b or α as the target variable, and when computing $Q_b(\alpha)$ using the predicted values of α. The values presented here are averaged MARE obtained when all clusters are successively considered as test sets.

ML models	Q_b	α	$Q_b(\alpha)$
MLR	13.4	34.1	38.3
Decision tree	7.6	7.9	10
Random forest	6.7	7.3	9.3
Extremely randomized trees	5.8	7.1	8.8

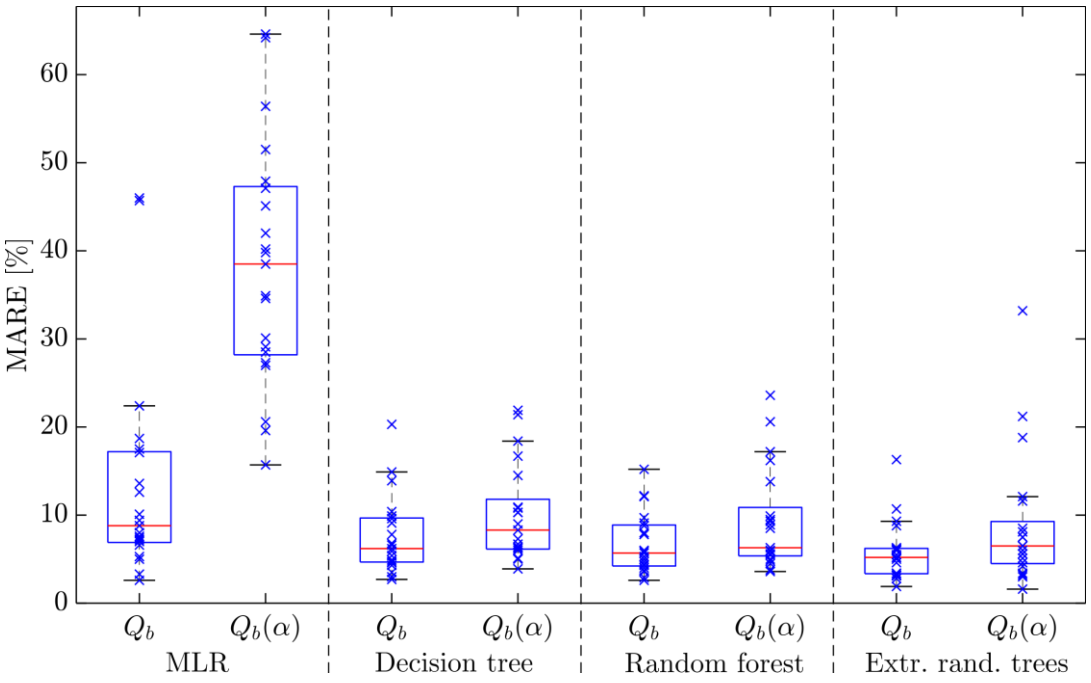


Figure 6. Boxplot of the MARE on the breach discharge associated to each ML technique and each target value when varying the test set.

4. Discussion

In this section, the performance of the extremely randomized trees model is compared against empirical formulas for Q_b and α. Interpolation and extrapolation capabilities of the ML model are assessed individually, i.e., by considering the test set to be inside or outside of the training set space. Finally, the impact of the breach geometry definition on the results is evaluated.

4.1. Comparison between machine learning and empirical formulas

Many empirical or semi-empirical formulas have been proposed in literature for the definition of the lateral breach discharge coefficient, C_D, (e.g., Subramanya and Awasthy (1972), Hager (1987), Jalili and Borghei (1996), Bagheri et al. (2014),…) which is directly involved in the expression of the breach discharge as

$$Q_b = \frac{2}{3} C_D \sqrt{2g\left(h - z_b\right)^3} L_s, \tag{8}$$

with g the gravitational acceleration [m/s^2], h the flow depth in the main channel upstream of the side opening [m], z_b the crest height of the side weir [m], and L_s the length of the side weir [m].

A thorough evaluation of the predictive capabilities of eleven empirical and semi-empirical formulas was conducted by Schmitz et al. (2024) for various side breach configurations. Formulas proposed by Jalili and Borghei (1996) and Singh et al. (1994) turned to be the most accurate when considering realistic fluvial dike breaches. These formulas respectively write

$$C_d = 0.33 - 0.18\ \mathrm{F} + 0.49\frac{z_b}{h} \qquad \text{(Singh et al., 1994),} \tag{9}$$

$$C_d = 0.71 - 0.41\ \mathrm{F} - 0.22\frac{z_b}{h} \qquad \text{(Jalili and Borghei, 1996).} \tag{10}$$

As stated earlier, the breach top width is the only information about the breach geometry that was measured in all experimental cases. To obtain the breach invert, z_b, the breach is assumed to be perfectly trapezoidal with its side slopes equal to the repose angle of the dike wet material. Additionally, erosion is assumed to be uniform over the entire breach area, which allows for the direct computation of z_b based on B_{top}. The length of the side weir, L_s, is chosen equal to the breach bottom width.

Alternatively, two new empirical relationships for the calculation of the effective breach width coefficient, α, are proposed in this work. These equations were derived by inspection of the experimental data presented in Section 2.2 (Figure S1 in Supplement). The first regression relies solely on the non-dimensional breach top width as its impact on the value of α appeared to be predominant (Schmitz et al., 2023a):

$$\alpha_{R1} = \frac{1}{3} + 0.035 B_{top,adim}^{-2.5}. \tag{11}$$

The second regression uses three input parameters, namely the non-dimensional breach top width, $B_{top,adim}$, a non-dimensional velocity, $U_{adim} = \frac{U_r}{\sqrt{gB_{top}}}$, and a non-dimensional dike width, $w_{d,adim} = \frac{w_{d,FS}}{B_{top}}$, with $w_{d,FS}$ the dike cross-section width at the water free surface level:

$$\alpha_{R2} = 1.789 w_{d,adim}^{0.242} \left(\frac{t_{adim}}{B_{top,adim}} \right)^{0.225} - 0.5 . \quad (12)$$

Finally, a corrective term computed through ML (extremely randomized tree algorithm) may be applied to these experimental formulas, leading to an improved evaluation of the breach discharge, denoted Q_b^+ (Figure 7). The target variable of this ML model is the difference between the experimental breach discharge and the breach discharge predicted by an empirical formula. The same features as the ones considered for the direct evaluation of Q_b and α are considered (Eq. (1)). In addition, the empirical value of α is added to the features when evaluating the corrective term related to the breach discharge computed by resorting Eq. (11) or (12), i.e., $Q_b^+(\alpha)$ in Figure 7.

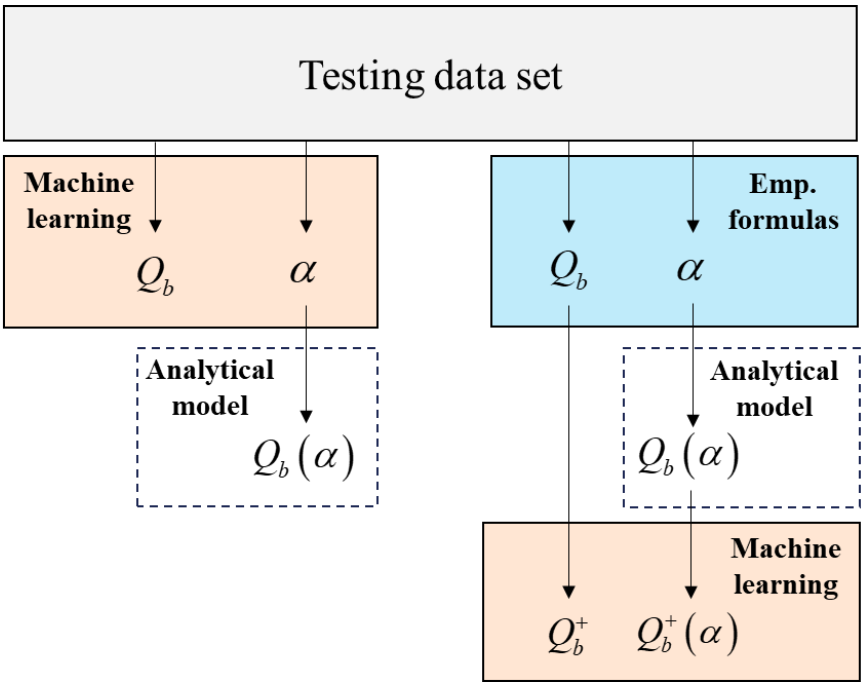


Figure 7. Flow chart of the different approaches adopted in this work to compute the breach discharge.

Table 3 summarizes the mean absolute relative error obtained with the extremely randomized trees model used for the direct evaluation of Q_b or α, with both empirical formulas for Q_b, with the new empirical regressions for α, and with each empirical formula combined to a ML-based corrective term, i.e., Q_b^+ and $Q_b^+(\alpha)$.

In all cases, machine-learning-based approaches perform better than the empirical formulas alone. In particular, the models proposed by Singh et al. (1994) and Jalili and Borghei (1996) lead to significant errors. These are much larger than those obtained by Schmitz et al. (2024)

due to the different procedure they adopted in their work. In Schmitz et al. (2024) the entire breaching event was modeled as a dynamic process, in which the water level was considered as an unknown and computed at each time step based on mass conservation in the main channel. Presently, the breach discharge is computed based on the associated instantaneous value of the features, with no dynamic considerations. For this reason, the experimental water level is used as an input variable. To a smaller extent, the definition of the breach geometry may also have an impact on the results. In the present work, the breach width at the free surface and the breach invert are derived from the breach top width, whilst Schmitz et al. (2024) used the experimental breach geometry as they only considered tests for which full breach reconstructions were available.

Both empirical regressions for α (Eq. (11) and (12)) lead to a similar accuracy. Also, they perform much better than the empirical formulas derived for the direct prediction of Q_b (Eq. (9) and (10)). This observation might be biased as α regressions were fit on this specific dataset. The use of a machine-learning-based corrective term systematically reduces the error on the results predicted by empirical formulas.

Table 3. Mean absolute relative error (MARE) [%] for each model and target variable.

	Q_b	$Q_b(\alpha)$	Q_b^+	$Q_b^+(\alpha)$
Extremely randomized trees	5.8	8.8	-	-
Singh et al. (1994) (Eq. (9))	37.4	-	7.8	
Jalili and Borghei (1996) (Eq. (10))	72	-	12.8	
α regression 1 (Eq. (11))	-	18.8	-	12.1
α regression 2 (Eq. (12))	-	18.5	-	8.8

4.2. Model performance in interpolation and extrapolation

Decision-tree-based ML models provide constant predictions when evaluated outside their training space. In this case, their performance is expected to degrade. Features associated with real dike configurations are diverse, but covering the entire feature space in the training set is hardly feasible due to the huge amount of experimental data that should be collected. This limitation highlights the need for good predictive capabilities in both interpolation and extrapolation, i.e., inside and outside the training space, respectively. Within this context, empirical formulas or analytical models are interesting tools as they usually capture the general physical trends, and their results keep evolving outside their fitting space.

In this section, the predictive capability of the ML and empirical formulas evaluated in Section 4.1 are assessed in interpolation and extrapolation separately. To this end, Figure 8 highlights clusters that were considered in extrapolation (red), i.e., located on the edge of the dataset space, and in interpolation (green). In total, 10 clusters (25 experimental tests) are considered in interpolation when individually selected as the test set, while 11 clusters are in extrapolation (18 experimental tests).

As expected, Table 4 and Figure 9 show that results are systematically more accurate in interpolation than in extrapolation when considering the extremely randomized trees algorithm alone. In interpolation, the averaged MARE on Q_b and $Q_b(\alpha)$ is limited to around 5% in both cases. While the accuracy on both variables drops in extrapolation, the error variability associated to Q_b is more limited than the one on $Q_b(\alpha)$ (Figure 10), making the direct prediction of Q_b more reliable.

In the present work, the concept of extrapolation and interpolation is meaningless for Q_b empirical formulas as they were fitted on a different dataset than the one considered in this work. Their poor and highly variable performance, especially for the formula of Jalili and Borghei (1996) (Figure ***10***), suggests that they generalize badly outside their fitting space. The performance of both α regression formulas is largely better in interpolation and extrapolation compared to the empirical formulas for Q_b. However, it decreases by 25% (α regression 1) to 140% (α regression 2) in extrapolation compared to the scores in interpolation.

Results are greatly improved when adding a ML-based corrective term to the empirical formulas, i.e., Q_b^+ and $Q_b^+(\alpha)$. In this case, the averaged MARE grows to a smaller extent when evaluated in extrapolation, especially for α regressions, although the performance variability still substantially increases (Figure 10). Overall, very satisfactory results are obtained in both interpolation and extrapolation when using the extremely randomized trees model for the prediction of Q_b and $Q_b(\alpha)$, and when adding a ML-based corrective term to the empirical formula for Q_b derived by Singh et al. (1994) and for α regression 2 (Table 4). Nonetheless, the predictive capability of ML-based models decreases when evaluated further from their training space. In this case, α regression 2 associated with a ML-based corrective should be favored as the associated analytical model will continue evolving based on physical considerations even outside the training space, limiting the impact of a poorly performing ML model.

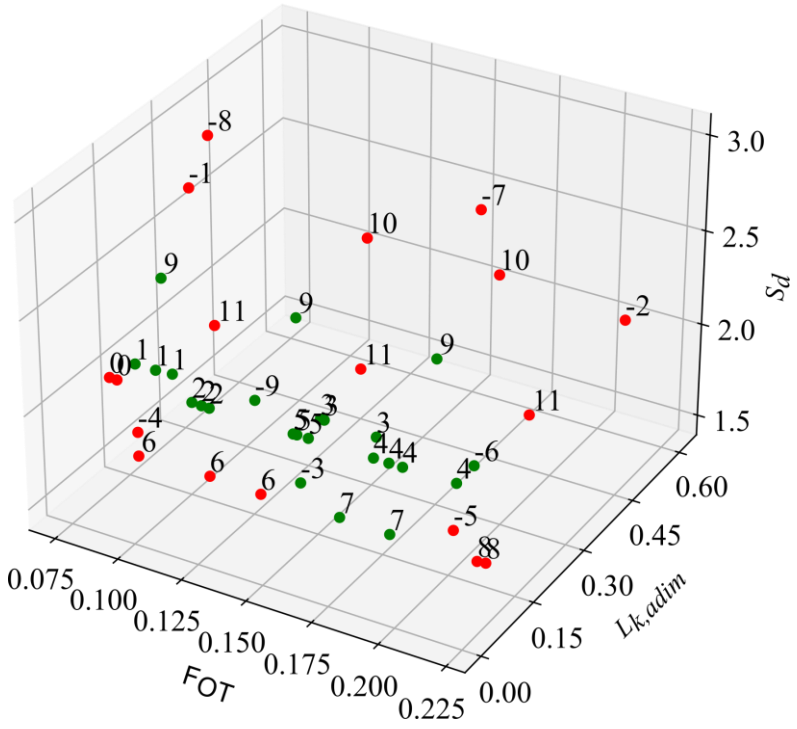


Figure 8. Clusters in extrapolation (red dots), and in interpolation (green dots) when used as test sets. Numbers refer to the clusters ID of each laboratory experiment.

Table 4. Averaged mean absolute relative error (MARE) [%] for each model and target variable when evaluated in interpolation and extrapolation.

	Q_b		$Q_b(\alpha)$		Q_b^+		$Q_b^+(\alpha)$	
	Interp.	Extrap.	Interp.	Extrap.	Interp.	Extrap.	Interp.	Extrap.
Extremely randomized trees	4.5	7	5.8	11.4	-	-	-	-
Singh et al. (1994) (Eq. (9))	36.7	38.1	-	-	5.5	9.9	-	-
Jalili and Borghei (1996) (Eq. (10))	75.9	68.5	-	-	9.9	15.4	-	-
α regression 1 (Eq. (11))	-	-	16.5	20.9	-	-	12.2	13.9
α regression 2 (Eq. (12))	-	-	10.2	24.2	-	-	6.9	10.5

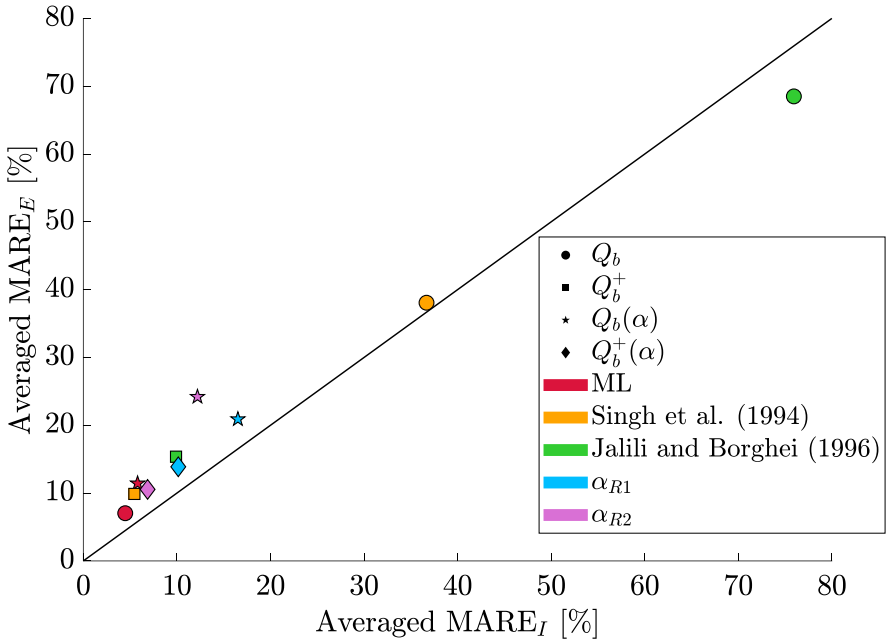


Figure 9. Averaged MARE obtained in interpolation (MARE$_I$) and extrapolation (MARE$_E$).

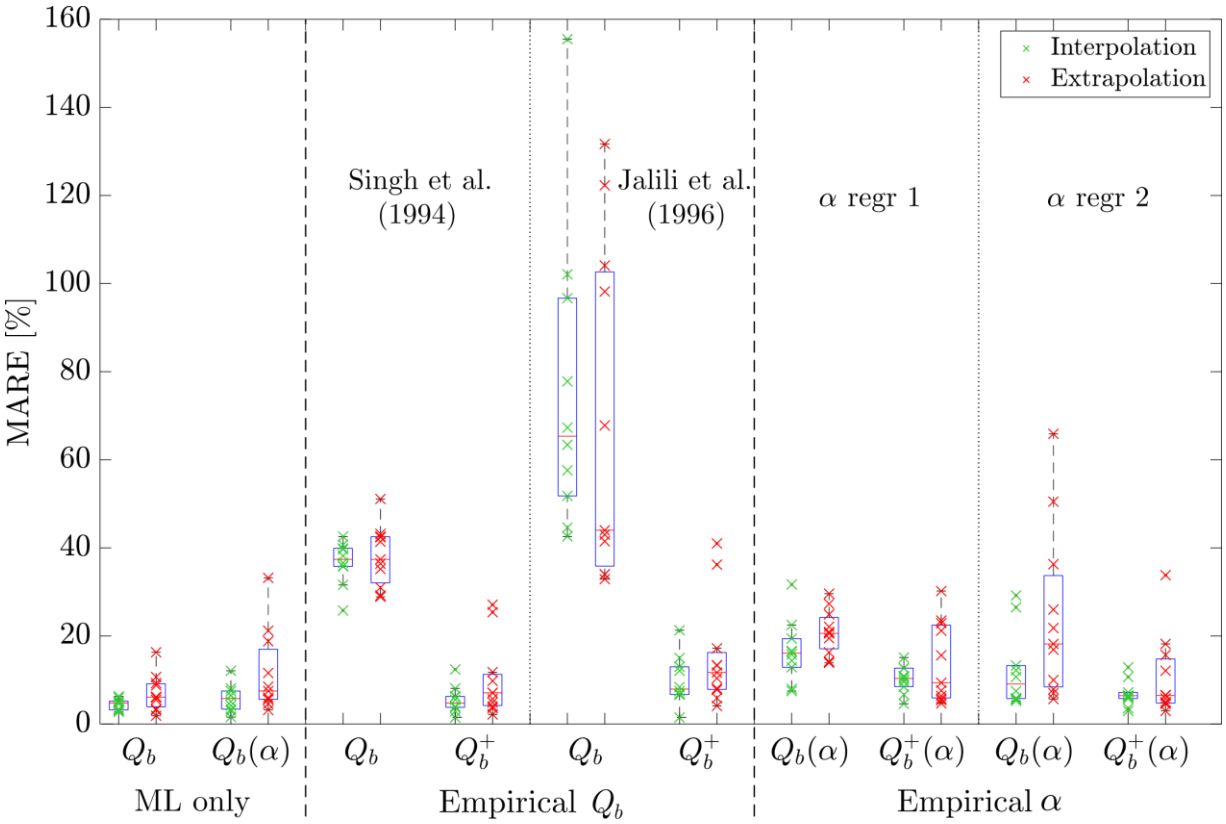


Figure 10. Boxplot of the MARE values associated with each approach in interpolation and extrapolation.

4.3. Impact of the breach geometry definition

Considering different features or adding new ones to the ML models may have a significant impact on the results. So far, the breach invert level, z_b, has been discarded as it was not recorded in all experiments (Section 2.2). In this section, only tests data for which the 3D breach geometry was recorded are considered (Rifai et al., 2019). The performance of the different models is assessed when the experimental breach invert is added to the features of the ML models and considered in the empirical formulas and the analytical model. In this case, the breach is still assumed trapezoidal with its side slopes equal to the repose angle of the dike wet material, but assuming a uniform breach erosion is no longer required. Figure 11 compares the performance of this new approach with the one obtained when z_b is not part of the features of the ML models and approximated based on B_{top} in the analytical model (Section 3.1) and the empirical formulas (Section 4.1). Defining a unique value for z_b is not trivial as the breach bottom is highly non-uniform (Rifai et al., 2018). Also, previous works showed that results might be largely influenced by the associated experimental uncertainties (Schmitz et al., 2023b). The sensitivity of the different models to z_b was assessed by defining it as (a) the minimum breach invert level, (b) the 10th percentile, (c) and the 20th percentile. In all cases, only the breach cross-section parallel to the dike main axis and passing through the center of the crest was considered.

Figure 11 shows that results are systematically improved with the experimental z_b when ML is used at last in the procedure, i.e., when computing Q_b with ML only, Q_b^+, and $Q_b^+(\alpha)$. This improvement is, however, limited in most cases, while the definition of z_b has almost no impact, i.e., a relative difference of less than 2% on the averaged MARE values is observed in all cases.

Considering the experimental z_b reduces the predictive capability of the new analytical model when ML is not used at the end of the procedure, i.e., when $Q_b(\alpha)$ is computed. In contrast, the predictions from empirical formulas proposed by Singh et al. (1994) and Jalili and Borghei (1996) with no ML-based corrective term are greatly improved when using the experimental breach invert. Here again, its definition has a negligible impact on the results (maximum relative variation of the averaged MARE of 12%).

It can be concluded that considering the experimental z_b is only recommended when using empirical formulas for Q_b on their own. In the other cases, the performance improvement is limited, and might even significantly decrease in some cases, e.g., when predicting $Q_b(\alpha)$. Also, it should be borne in mind that increasing the number of ML features increases the model training time. Finally, the definition of the experimental z_b has a negligible impact on the results.

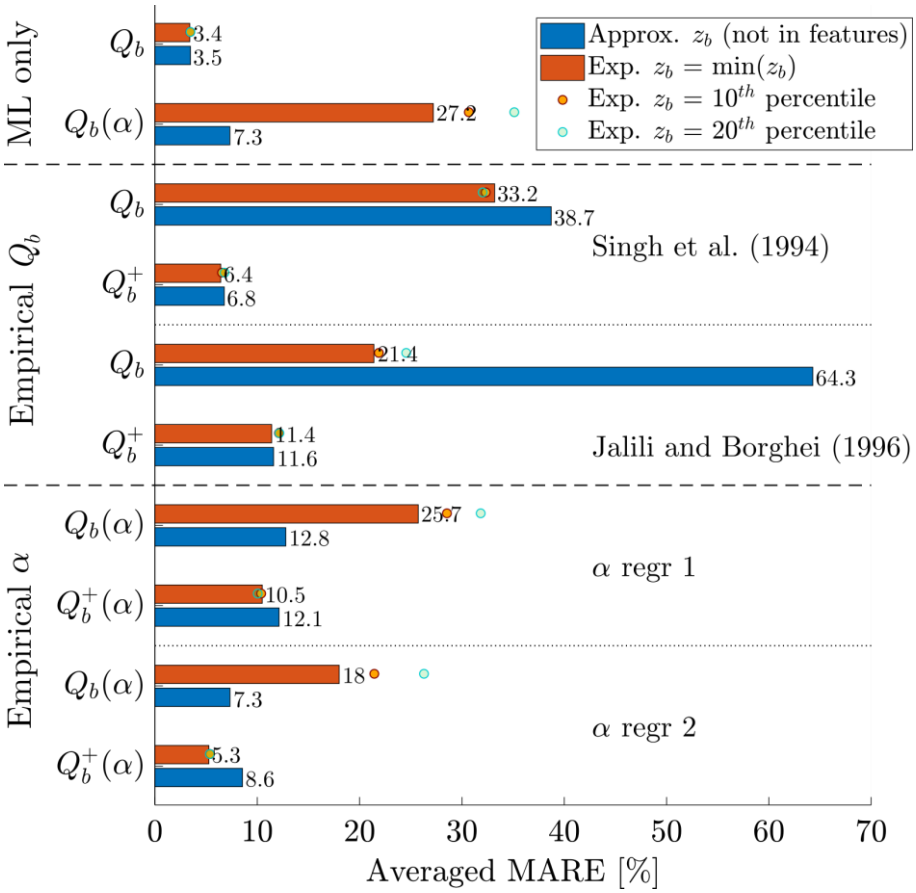


Figure 11. Impact of the definition of the breach invert on the models' accuracy.

5. Conclusions

In this work, two different approaches for the prediction of the discharge through a fluvial dike breach were presented, namely the direct prediction of the discharge Q_b, or its indirect prediction through the parameter α using a new analytical model (Section 3.1). Among the linear and decision-tree-based machine learning techniques presented in this work for the prediction of those parameters (Section 3.2), the extremely randomized trees algorithm provides almost continuous outputs while being very accurate (mean absolute relative error between 5% and 10% on average).

The ML model used to predict Q_b or $Q_b(\alpha)$ was compared to empirical formulas for the prediction of Q_b and α, respectively (Section 4.1). Empirical formulas were less accurate, especially when predicting Q_b. A ML-based corrective term was then added to the different empirical formulas results. It significantly improved the results accuracy, especially when using Singh et al. (1994) formula and the second alpha regression (Eq. (12)).

In all cases, the model performance decreased slightly when evaluated in extrapolation, i.e., outside the space of the training set, whilst the variability of the model accuracy on the tested set rose (Figure 10). It is expected that the further the ML models are evaluated from their training space, the lower the model performance. In the case of decision-tree-based ML techniques, the predictions become constant outside the training space. In this context, the use of the second empirical regression for α coupled to a ML-based corrective term is

recommended. While offering accurate results, this method keeps evolving outside the training set space as it relies on the new analytical model, which uses continuously evolving parameters.

The performance sensitivity to the definition and the addition of the breach invert level, z_b, in the features of the ML models was also investigated (Section 4.3). The use of the experimental value of z_b is only recommended when considering empirical formulas for the prediction of Q_b. In the other cases, its use degrades the results or slightly improves them in the best case, to the expense of a longer training time of the ML model.

To further improve the performance of the models, more experimental data is needed to increase the dataset space. Notably, some parameters were considered as constant over the entire tested space, e.g., dike height, material median diameter, or cohesion. These parameters should be incorporated in the ML features when expanding the training set with configurations involving different values for these parameters. Most importantly, new data and experimental setups should be comprehensively documented to ensure the consistency and inter-comparability of the database (Di Bacco and Scorzini, 2019).

The use of other ML methods for the prediction of the fluvial dike breach discharge could also be investigated, e.g. neural networks. However, purely data-driven approaches usually do not obey the governing laws of physical systems (Yin et al., 2021). Within this context, physics-guided deep learning methods have become increasingly popular when analyzing dynamic systems. In these methods, the loss function is reformulated to take into account physical constraints imposed by the user (Dona et al., 2022). Although still at its early stages, this research field is evolving quickly and may hold substantial potential for future applications (Wang and Yu, 2023).

Acknowledgements

The authors gratefully acknowledge S. Pierard from the Department of Electrical Engineering and Computer Science of the University of Liège for its valuable contribution to the discussions on the implementation of machine learning techniques.

Appendix

In Section 3.1, the energy conservation between the main channel and the breach cross-section was simplified as

$$H_r = (z_r - z_b) + \frac{U_r^2}{2g} = h_c + \frac{A_c}{2L_c},$$

with $L_c = \alpha L_b = \alpha(b_{bot} + 2mh_c)$ and $m > 0$. The effective critical section, A_c, should then be expressed as a function of the critical height, h_c, i.e. the only unknown. Three cases are encountered, which depend on the value of h_c. For this reason, an iterative solving scheme should be adopted to solve the equation.

$\underline{L_c < mh_c}$

In this case, A_c corresponds to a triangle and writes

$$A_c = \frac{L_c^2}{2m} = \frac{\alpha^2}{2m}\left(b_{bot} + 2mh_c\right)^2,$$

which leads to

$$h_c = \frac{-\frac{\alpha}{4m}b_{bot} + H_r}{1 + \frac{\alpha}{2}}.$$

$mh_c \le L_c \le b_{bot} + mh_c$

In this case, A_c corresponds to a trapezoid and writes

$$A_c = \left(L_c - mh_c\right)h_c + m\frac{h_c^2}{2} = \left[\alpha b_{bot} + mh_c\left(2\alpha - 0.5\right)\right]h_c .$$

The resulting expression of the critical height is

$$h_c = \frac{4\alpha}{12\alpha - 1}\left[H_r - 0.75\frac{b_{bot}}{m} + \sqrt{\left(0.75\frac{b_{bot}}{m}\right)^2 + \frac{H_r\left(6\alpha - 1\right)}{4\alpha}\frac{b_{bot}}{m} + H_r^2}\right].$$

$b_{bot} + mh_c < L_c \le b_{bot} + 2mh_c$

In this configuration, A_c corresponds to a pentagon and writes

$$A_c = \left(L_c + b_{bot}\right)\frac{h_c}{2} - \frac{\left(L_b - L_c\right)^2}{2m} = b_{bot}h_c + mh_c^2 - \frac{\left(1-\alpha\right)^2}{2m}\left(b_{bot} + 2mh_c\right)^2$$

It comes

$$h_c = \frac{\left(2\alpha^2 - 6\alpha + 1\right)b_{bot} + 4\alpha mH_r + \sqrt{\Delta}}{-m\left(4\alpha^2 - 16\alpha + 2\right)},$$

with $\Delta = \left[\left(2\alpha^2 - 6\alpha + 1\right)b_{bot} + 4\alpha mH_r\right]^2 + 4\left[mb_{bot} - m\left(2\alpha^2 - 8\alpha + 1\right)\right]\left[\frac{\left(1-\alpha\right)^2}{2m}b_{bot}^2 + 2\alpha H_r b_{bot}\right]$.

$L_c > b_{bot} + 2mh_c$

This last case is encountered when the critical section, A_c, is larger than the entire breach cross-section, A_b. It yields $A_c = \alpha A_b$, and

$$h_c = \frac{2}{5}\left[H_r - \frac{3}{4}\frac{b_{bot}}{m} + \sqrt{\left(\frac{3}{4}\frac{b_{bot}}{m}\right)^2 + \frac{3}{4}\frac{b_{bot}}{m}H_r + H_r^2}\right].$$

References

Ankerst, M., Breunig, M.M., Kriegel, H.-P., Sander, J., 1999. OPTICS: ordering points to identify the clustering structure, in: Proceedings 1999 ACM SIGMOD International ConferenceManagement Data, SIGMOD'99. Association for Computing Machinery, Philadelphia, Pennsylvania, USA, pp. 49–60.

ASCE/EWRI Task Committee on Dam/Levee Breaching, 2011. Earthen Embankment Breaching. Journal of Hydraulic Engineering 137, 1549–1564.

Azimi, H., Bonakdari, H., Ebtehaj, I., 2017. A highly efficient gene expression programming model for predicting the discharge coefficient in a side weir along a trapezoidal canal. Irrigation and drainage 66, 655–666.

Bagheri, S., Kabiri-Samani, A.R., Heidarpour, M., 2014. Discharge coefficient of rectangular sharp-crested side weirs, Part I: Traditional weir equation. Flow Measurement and Instrumentation 35, 109–115.

Balahang, S., Ghodsian, M., 2023. Evaluating performance of various methods in predicting triangular sharp-crested side weir discharge. Applied Water Science 13, 171.

Borghei, S.M., Jalili, M.R., Ghodsian, M., 1999. Discharge Coefficient for Sharp-Crested Side Weir in Subcritical Flow. Journal of Hydraulic Engineering 125, 1051–1056.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Charrier, G., 2015. Etude expérimentale des ruptures de digues fluviales par surverse (Doctoral dissertation, Aix-Marseille).

Cheng, Y., Song, Y., Liu, C., Wang, W., Hu, X., 2022. Numerical Simulation Research on the Diversion Characteristics of a Trapezoidal Channel. Water 14.

Cheong, H., 1991. Discharge Coefficient of Lateral Diversion from Trapezoidal Channel. Journal of Irrigation and Drainage Engineering 117, 461–475.

Chowdhury, M.K., Konsoer, K.M., Hiatt, M., 2022. Effect of Lateral Outflow on Three-Dimensional Flow Structure in a River Delta. Water Resources Research 58, e2021WR031346.

Di Bacco, M., Scorzini, A.R., 2019. Are We Correctly Using Discharge Coefficients for Side Weirs? Insights from a Numerical Investigation. Water 11.

Dona, J., Déchelle, M., Gallinari, P., Levy, M., 2022. Constrained Physical-Statistics Models for Dynamical System Identification and Prediction, in: InternationalConference Learning Representations.

Ebtehaj, I., Bonakdari, H., Gharabaghi, B., 2018. Development of more accurate discharge coefficient prediction equations for rectangular side weirs using adaptive neuro-fuzzy inference system and generalized group method of data handling. Measurement 116, 473–482.

Elalfy, E., Tabrizi, A.A., Chaudhry, M.H., 2018. Numerical and experimental modeling of levee breach including slumping failure of breach sides. Journal of Hydraulic Engineering 144.

Emiroglu, M.E., Agaccioglu, H., Kaya, N., 2011. Discharging capacity of rectangular side weirs in straight open channels. Flow Measurement and Instrumentation 22, 319–330.

Flynn, S., Zamanian, S., Vahedifard, F., Shafieezadeh, A., Schaaf, D., 2022. Data-Driven Model for Estimating the Probability of Riverine Levee Breach Due to Overtopping. Journal of Geotechnical and Geoenvironmental Engineering 148, 04021193.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Machine learning 63, 3–42.

Haddadi, H., Rahimpour, M., 2012. A discharge coefficient for a trapezoidal broad-crested side weir in subcritical flow. Flow Measurement and Instrumentation 26, 63–67.

Hager, W.H., 1987. Lateral Outflow Over Side Weirs. Journal of Hydraulic Engineering 113, 491–504.

Haghbin, M., Sharafati, A., 2022. A review of studies on estimating the discharge coefficient of flow control structures based on the soft computing models. Flow Measurement and Instrumentation 83, 102119.

Hameed, M.M., AlOmar, M.K., Khaleel, F., Al-Ansari, N., 2021. An Extra Tree Regression Model for Discharge Coefficient Prediction: Novel, Practical Applications in the Hydraulic Sector and Future Research Directions. Mathematical Problems in Engineering 2021.

Ibrahim, I., Riviere, N., Lebouteiller, I., Mignot, E., 2022. Discharge Distribution in Open-Channel T-Shape Bifurcations: Effect of a Reduced Side Branch Width. Journal of Hydraulic Engineering 148, 04022015.

Jalili, M.R., Borghei, S.M., 1996. Discussion: Discharge Coefficient of Rectangular Side Weirs. Journal of Irrigation and Drainage Engineering 122, 132–132.

Jamei, M., Ahmadianfar, I., Chu, X., Yaseen, Z.M., 2021. Estimation of triangular side orifice discharge coefficient under a free flow condition using data-driven models. Flow Measurement and Instrumentation 77, 101878.

Kakinuma, T., Shimizu, Y., 2014. Large-scale experiment and numerical modeling of a riverine levee breach. Journal of Hydraulic Engineering 140.

Kakinuma, T., Tobita, D., Yokoyama, H., Takeda, A., 2013. Levee breach observation at Chiyoda experimental flume, in: 12thInternationalSymposium River Sedimentation (ISRS), IRTCES, Kyoto, Japan.

Lee, K., 2019. Simulation of Dam-Breach Outflow Hydrographs Using Water Level Variations. Water Resources Management 33, 3781–3797.

Li, S., Shen, G., Parsaie, A., Li, G., Cao, D., 2024. Discharge modeling and characteristic analysis of semi-circular side weir based on the soft computing method. Journal of Hydroinformatics 26, 175–188.

Michelazzo, G., Oumeraci, H., Paris, E., 2015. Laboratory Study on 3D Flow Structures Induced by Zero-Height Side Weir and Implications for 1D Modeling. Journal of Hydraulic Engineering 141, 04015023.

Neary, V.S., Sotiropoulos, F., Odgaard, A.J., 1999. Three-Dimensional Numerical Model of Lateral-Intake Inflows. Journal of Hydraulic Engineering 125, 126–140.

Parsaie, A., 2016. Predictive modeling the side weir discharge coefficient using neural network. Modeling Earth Systems and Environment 2, 1–11.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Ranga Raju, K.G., Gupta, S.K., Prasad, B., 1979. Side Weir in Rectangular Channel. Journal of the Hydraulics Division 105, 547–554.

Rifai, I., El Kadi Abderrezzak, K., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., Dewals, B., 2018. Floodplain Backwater Effect on Overtopping Induced Fluvial Dike Failure. Water Resources Research 54, 9060–9073.

Rifai, I., El Kadi Abderrezzak, K., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., Dewals, B., 2019. Flow and detailed 3D morphodynamic data from laboratory experiments of fluvial dike breaching. Scientific data 6, 53.

Rifai, I., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., El Kadi Abderrezzak, K., Dewals, B., 2017. Overtopping induced failure of noncohesive, homogeneous fluvial dikes. Water Resources Research 53, 3373–3386.

Rifai, I., Schmitz, V., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., Dewals, B., El Kadi Abderrezzak, K., 2020. Continuous Monitoring of Fluvial Dike Breaching by a Laser Profilometry Technique. Water Resources Research 56, e2019WR026941.

Roushangar, K., Khoshkanar, R., Shiri, J., 2016. Predicting trapezoidal and rectangular side weirs discharge coefficient using machine learning methods. ISH Journal of Hydraulic Engineering 22, 254–261.

Schmitz, V., Arnst, M., Abderrezzak, K. El kadi, Pirotton, M., Erpicum, S., Archambeau, P., Dewals, B., 2023b. Global sensitivity analysis of a dam breaching model: To which extent is parameter sensitivity case-dependent? Water Resources Research e2022WR033894.

Schmitz, V., Erpicum, S., Abderrezzak, K. El kadi, Rifai, I., Archambeau, P., Pirotton, M., Dewals, B., 2021. Overtopping-Induced Failure of Non-Cohesive Homogeneous

Fluvial Dikes: Effect of Dike Geometry on Breach Discharge and Widening. Water Resources Research 57, e2021WR029660.

Schmitz, V., Kitsikoudis, V., Wylock, G., Erpicum, S., Pirotton, M., Archambeau, P., Dewals, B., 2024. Efficient modelling of lateral discharge through a dike breach. Journal of Hydrology 640, 131660.

Schmitz, V., Rifai, I., Kheloui, L., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., El Kadi Abderrezzak, K., Dewals, B., 2023a. Main channel width effects on overtopping-induced non-cohesive fluvial dike breaching. Journal of Hydraulic Research 61, 601–610.

Shanker, M., Hu, M.Y., Hung, M.S., 1996. Effect of data standardization on neural network training. Omega 24, 385–397.

Singh, R., Manivannan, D., Satyanarayana, T., 1994. Discharge Coefficient of Rectangular Side Weirs. Journal of Irrigation and Drainage Engineering 120, 814–819.

Stilmant, F., Pirotton, M., Archambeau, P., Roger, S., Erpicum, S., Dewals, B., 2013. Dike-break induced flows: a simplified model. Environmental fluid mechanics 13, 89–100.

Subramanya, K., Awasthy, S.C., 1972. Spatially Varied Flow over Side-Weirs. Journal of the Hydraulics Division 98, 1–10.

Wang, R., Yu, R., 2023. Physics-Guided Deep Learning for Dynamical Systems: A Survey.

Wang, Y., Lv, M., Wang, W., Meng, M., others, 2024. Discharge Formula and Hydraulics of Rectangular Side Weirs in the Small Channel and Field Inlet. Water 16, 713.

Yin, Y., Le Guen, V., Dona, J., Bézenac, E. de, Ayed, I., Thome, N., Gallinari, P., 2021. Augmenting physical models with deep networks for complex dynamics forecasting. Journal of Statistical Mechanics: Theory and Experiment 2021, 124012.

Supplement to "How Does Machine Learning Compare to Hybrid Approaches for Predicting Fluvial Dike Breach Discharge?"

Table S1. Values of the hyperparameters used in the machine learning models. Values labelled with a star (*) are recommended in the *scikit-learn* documentation.

<table>
<tr><th></th><th>Maximum tree depth</th><th>Split criterion</th><th>Minimum samples for split</th><th>Minimum samples in a leaf</th><th>Number of estimators</th></tr>
<tr><td>Decision trees</td><td rowspan="3">15</td><td rowspan="3">Mean square error*</td><td rowspan="3">2*</td><td rowspan="3">1*</td><td>-</td></tr>
<tr><td>Random forests</td><td rowspan="2">1000</td></tr>
<tr><td>Extremely randomized trees</td></tr>
</table>

Table S2. Definition of the clusters. S_u = dike slope on the main channel side; S_d = dike slope on the floodplain side; Q_{in} = target inflow discharge; F_{OT} = main channel Froude number at overtopping initiation.

Cluster ID	S_u (-)	S_d (-)	L_k (m)	Q_{in} (l/s)	F_{OT} (-)
-9	1.5	2	0.15	40	0.125
-8	2	3	0.3	25	0.086
-7	2	3	0.3	55	0.191
-6	1.5	2	0.15	55	0.208
-5	2	2	0	55	0.221
-4	2	2	0	25	0.103
-3	2	2	0	40	0.165
-2	2	2	0.6	55	0.208
-1	2	3	0.15	25	0.100
0	2	2	0.1	20	0.075
				21	0.078
1	1.5	2	0.15	25	0.078
	2				0.086
	2				0.092
2	2	2	0.1	28	0.108
				30	0.112
				31	0.115
3	2	2	0.15	40	0.150
				40	0.152
				55	0.173
4	2	2	0.1	47	0.178
				50	0.184
				51	0.189
				55	0.209
5	2	2	0.1	40	0.147
				40	0.147
				40	0.149
				41	0.153
6	1.5	1.5	0.15	25	0.078
	2				0.106
	1.5			40	0.127
7	2	1.5	0.15	40	0.157
	1.5			55	0.177
8	2	1.5	0.15	55	0.210
					0.213
9	2	2.5	0.15	25	0.089
				40	0.142
				55	0.195
10	2	3	0.15	40	0.168
				55	0.216
11	2	2	0.3	25	0.087
				40	0.146
				55	0.209

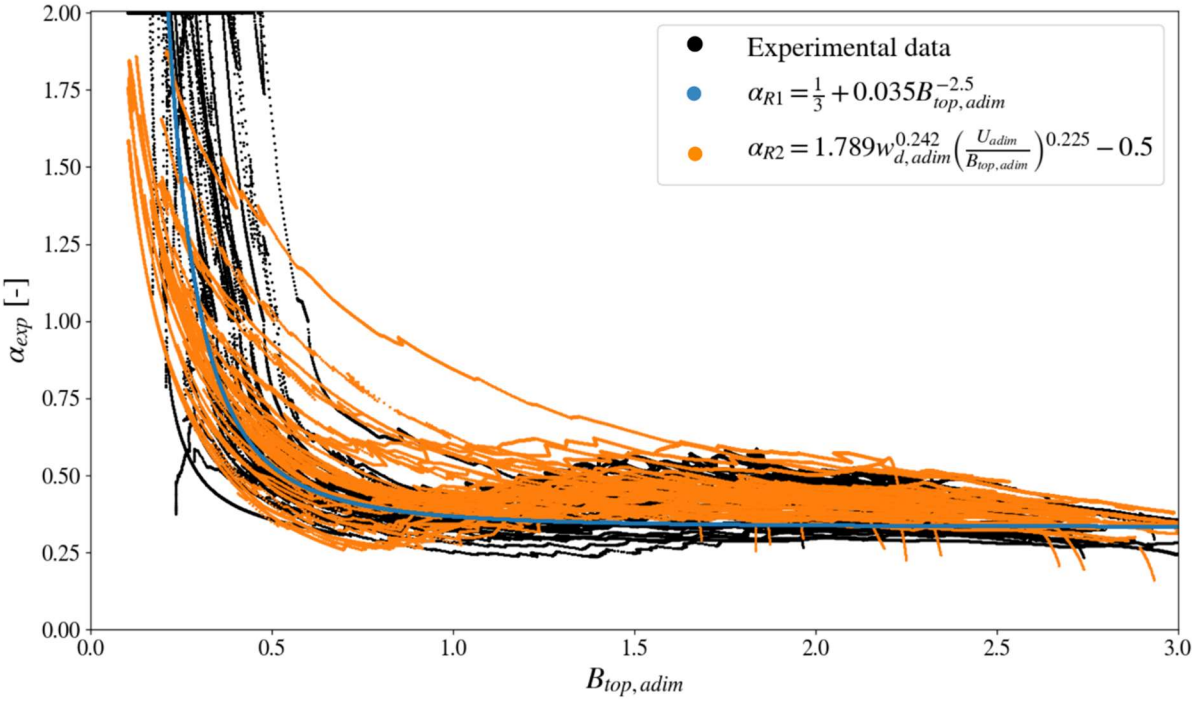


Figure S1. Fitting of the α regression formulas.

Chapter 6

Comparison of Numerical Modelling Approaches for Fluvial Dike Breaching

This chapter is partly inspired from to the conference paper "Comparison of numerical modelling approaches for fluvial dike breaching" by V. Schmitz, P. Archambeau, M. Pirotton, S. Erpicum, Dewals, that will be published in the Proceedings of the 12th International Conference on Fluvial Hydraulics (River Flow 2024, Liverpool). The PhD candidate developed the methodology, implemented the numerical models, analyzed the results, wrote the manuscript and generated the figures.

Comparison of Numerical Modelling Approaches for Fluvial Dike Breaching

V. Schmitz[1], P. Archambeau[1], M. Pirotton[1], S. Erpicum[1] & B. Dewals[1]

[1] *Research Group of Hydraulics in Environmental and Civil Engineering (HECE), University of Liège, Liège, Belgium.*

Key Points:

- A lumped dike breaching model is compared to its coupling with a 2D hydrodynamic model.
- The coupled model parametrization is adapted to better predict the breach discharge and expansion.
- Field-scale experiments show that the adjusted model provides more accurate and conservative predictions for both breach discharge and widening.
- The adjusted coupled model is applied to a real-world case study and proved to be an effective tool for supporting safer evacuation planning in floodplains.

Chapter 6: Comparison of Numerical Modelling Approaches

ABSTRACT

Failures of fluvial dikes often lead to devastating consequences in floodplains, with overtopping being one of the most frequent causes of dike failure. Numerical models are instrumental for assessing the consequences of failure and guide emergency plans. Multiple modelling strategies exist for representing dike morphology and flow description, but also the coupling between them. The selected modelling approach significantly influences the model's accuracy and computational performance. This study investigates two modelling strategies for fluvial dike breaching induced by overtopping. The first approach employs a spatially non-discretized dike breaching model to describe the flow and morphologic evolution of the breach. The second approach uses the dynamic coupling of a 2D hydrodynamic model with the lumped dike breaching model, whose role is then limited to describing the dike breach morphodynamics.

To evaluate their predictive capabilities, both modelling approaches were tested against field-scale experiments. Results indicate that the lumped model alone performed better than the coupled model in predicting the peak breach, while both models systematically underestimated the breach expansion. An in-depth analysis of the experimental data suggested that part of the discrepancies may be attributed to non-negligible inaccuracies in experimental measurements. Based on experimental observations, three flow resistance and sediment transport parameters were adjusted. Additionally, numerical and experimental tests supported reducing the effective breach width, i.e., the fraction of the breach width that conveys most of the flow.

This adjusted parametrization led to significantly more accurate and conservative predictions for both breach discharge and widening, which are critical for flood risk assessment. Conversely, using the numerical model with default parameters and in dam configuration, i.e., with the effective breach width equal to the total breach width, underestimated both breach expansion and discharge. This demonstrates the unsuitability of this model type for dike breaching scenarios. In a real-world case study, the dike configuration with the reduced effective breach width and adjusted parameters predicted larger short-term flood extents, thereby supporting safer evacuation planning in densely urbanized areas with critical infrastructures. Future work should focus on better characterizing the breach velocity field through 3D simulations and large-scale experiments to further enhance model accuracy.

1. Introduction

With the surge of extreme meteorological events, intensification of urbanization in the floodplains and aging infrastructures, fluvial dikes get more and more prompt to breaching. Hence, the need for predictive dike breaching models has become of paramount importance for establishing safe land-use planning and emergency response procedures (Zhong et al., 2021). Multiple modelling strategies exist regarding dike morphology and flow descriptions. The selected strategy significantly influences the model accuracy and computational performance.

The simplest and most computational-efficient models do not require any discretization of the domain and are therefore called "lumped models". Among this category, statistical (or parametric) models are exclusively based on regression analysis of data from past events, without considering underlying physics (De Lorenzo and Macchione, 2014; Chen et al., 2019;

Lee, 2019). In contrast, lumped physics-based models enable simulating hydraulic and dike breach variables (e.g., time-evolution of breach discharge and dimensions) by describing selected physical processes (Wu, 2016; Peter et al., 2018; Tsai et al., 2019; Li et al., 2020; Van Damme, 2020). For this reason, they may be considered as more general and trustworthy than purely statistical models. Conversely to lumped models, fully distributed physically based models can describe the phenomenon in greater detail, as they solve the flow and sediment governing equations using a computational mesh of the domain (e.g., Dazzi et al. (2019) or Hu et al. (2023)). Provided that the correct physical processes (including geotechnical aspects, e.g., side slope collapses) are properly incorporated in the model structure, their results are expected to be more accurate than lumped models, to the expense of a substantial increase in their computational cost. Nevertheless, combining a spatially distributed sediment transport model of dike breaching and a hydrodynamic model is usually avoided. 2D models based on shallow-water equations can only represent horizontal surface erosion, although geotechnical three-dimensional effects cannot be neglected. However, 3D models are computationally too demanding. Therefore, lumped models for dike breaching are often preferred. This coupling may be dynamic, i.e., data are shared between modules during computation, or static, i.e., the modules are run sequentially.

Here, two modelling strategies were implemented. The first modelling approach consists in describing the flow and the morphologic evolution of the dike breach using our adapted version of the lumped physically based model "DLBreach" (Schmitz et al., 2023a) that was initially developed by Wu (2013) for dam breaching, i.e., frontal configuration. In the other modelling approach, an in-house-developed 2D hydrodynamic model based on shallow-water equations (WOLF-2D) is dynamically coupled with our adapted dike breach model, whose goal is now limited to the description of the dike breach morphodynamics.

Two experimental field tests conducted by Kakinuma et al. (2013) on a 735-m-long stretch of the Chiyoda Test Channel (Japan) are used to assess the predictive capability of each modelling strategy. The focus is set on the evolution of the breach discharge and expansion. Subsequently, the coupled approach is applied to a hypothetical dike breach along the Albert Canal in southern Belgium.

The remainder of the paper is structured as follows: Section 2 introduces the modelling strategies. Section 3 describes the field experiments used as test cases and the Albert Canal case study. Results are presented and discussed in Section 4 to assess the accuracy and performance of each modelling approach. Additionally, the applicability of the developed coupled modelling approach is demonstrated through a real-world case study, and a sensitivity analysis of different model parametrizations is presented. Finally, conclusions are drawn in Section 5.

2. Modelling strategies

2.1 Lumped model

An existing semi-analytical physically-based dam breaching model was adapted to the case of fluvial dike breaching (Schmitz et al., 2023a). The original model is called DLBreach and was developed by Wu (2013). It is composed of three coupled modules: a hydrodynamic module, a sediment transport module, and a dike morphodynamic module (Figure 1). The sediment transport and dike morphodynamic modules are parts of a so-called "breach evolution module". The flow is assumed uniformly distributed through the entire trapezoidal breach and the breach expansion is symmetric. However, in fluvial configurations, the flow inertia in the main channel direction plays an important role in the breaching process. Laboratory experiments (Rifai et al., 2017), field tests (Kakinuma and Shimizu, 2014) and detailed 2D computations (Charrier, 2015) show that the flow through the breach is not uniformly distributed along the breach width, but rather concentrated in the most downstream part of the breach. This phenomenon induces a greater water velocity near the downstream breach extremity (Charrier, 2015), which leads to more intense erosion at this location. Conversely, limited erosion appears on the upstream extremity, leading to non-symmetrical breach expansion in the case of fluvial dikes (Schmitz et al., 2023a).

In our modified implementation of the model of Wu (2013, 2016), the concept of "effective breach width" was introduced to describe the fraction of the total breach width that is effectively used to convey water (Schmitz et al., 2023a). Shortly after the breaching onset (i.e., Stage Kakinuma et al. (2013)1 in Rifai et al. (2017)), flow velocity fields along both breach extremities tend to be similar due to the small breach width, leading to a symmetrical breach expansion. During this period, the initial breach expansion description proposed by Wu (2013) is thus used. Once the breach bottom reaches the channel bottom, asymmetrical breach widening is enforced by considering an effective breach width, b_{eff}, equal to half of the total breach width in the present work. This leads to a higher flow velocity, U_b, for a similar breach discharge, Q_b, so that:

$$U_b = Q_b / A_{eff} \quad , \tag{1}$$

with $A_{eff} = f(b_{eff})$ the effective breach flow section. Erosion is computed based on this increased velocity and the related effective breach area. Erosion only appears on the downstream breach extremity and breach expansion intensifies in that direction.

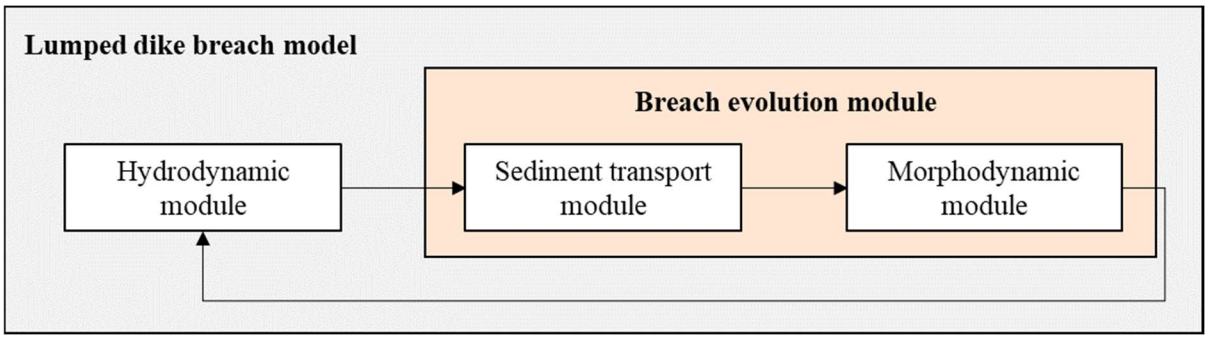


Figure 1. Flow chart of the lumped dike breach model.

2.2 Dynamic coupling with a 2D-horizontal hydrodynamic model

Figure 2 provides a flow chart of the second modelling strategy, which consists in dynamically coupling the breach evolution module of the lumped model with WOLF-2D, a 2D-horizontal hydrodynamic model running on GPU.

At each time step, $t_{hydro,n}$, WOLF-2D computes the entire hydrodynamic field. At specific time steps defined by the user, $t_{topo,i}$, the breach evolution module of the lumped model is provided with WOLF-2D hydrodynamic data, namely the breach discharge, Q_b, and the mean water depth across the breach section located along the dike crest centerline, h_{top}. The lumped model then computes the resulting breach topography, which will be imposed at the next user-specified time step in the hydrodynamic model, $t_{topo,i+1}$. This updated topography is sent back to WOLF-2D, which uses a linear interpolation to estimate the topography between successive topographies provided by the lumped model. The entire hydrodynamic field is updated accordingly, until reaching the next user-specified time step, $t_{topo,i+1}$. In the present case, the breach topography was updated every 0.5 seconds in the simulations related to the tests conducted by Kakinuma et al. (2013) and every 30 seconds in the real-world case study, i.e., $t_{topo,i+1} - t_{topo,i} = 0.5$ s or 30 s.

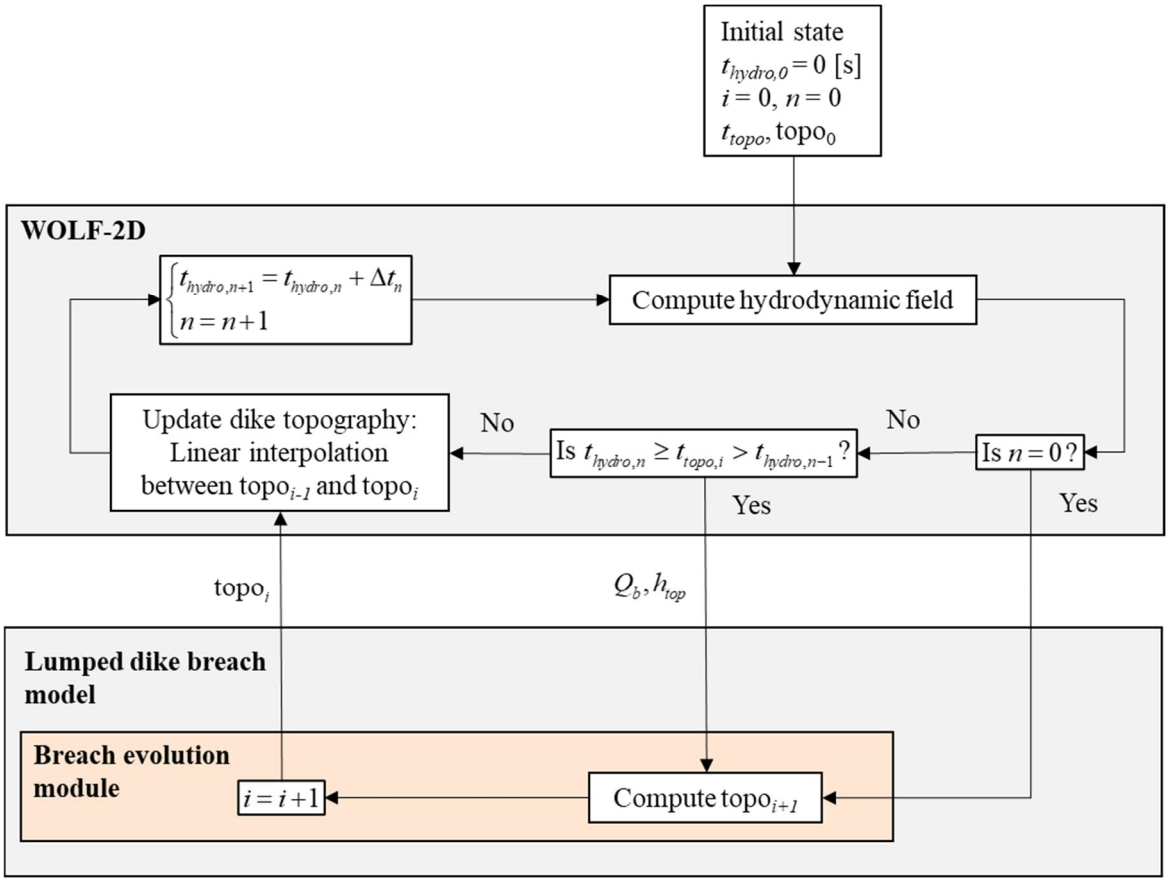


Figure 2. Flow chart of the dynamic coupling between WOLF-2D and the breach evolution module of the lumped dike breach model.

2.3 Lumped model and coupled approach provided with experimental topography

Two alternative approaches were adopted by using the experimental breach width at the dike crest level as an input for the lumped model and WOLF-2D. These approaches allow focusing on the capability of each model to predict the breach discharge, independently of the performance of the erosion module. In those cases, the observed breach geometry is simplified as a flat bottom and two constant lateral slopes. Also, the breach deepening is obtained by assuming a uniform erosion rate over the entire breach surface because only the breach width at the dike crest level was recorded during the experiments. In practice, a linear interpolation is applied between two successive experimental topographies.

3. Case studies

3.1 Field experiments on the Chiyoda Test Channel

3.1.1 Experimental setup

Large scale dike breaching tests were performed on the Chiyoda Test Channel (Tokachi River, Japan) between 2010 and 2011 (Kakinuma et al., 2013). The setup consists of a 735-m-long slightly curved main channel with a bed slope of 1/500 and a variable width (Figure 3). The bottom of the channel is erodible. Here, Tests 1 and 2 are considered.

Over the first 363-m-stretch of the river, both channel sides are limited by a 4-m-high trapezoidal non-erodible dike, whose slopes are equal to 2 (H/V). The channel bottom is 22 m wide over this section. To reach higher water levels using a limited inflow discharge, vertical steel pile sheets were installed on the left side of the rest of the channel to reduce its bottom width to 8 m. A 3-m-high trapezoidal erodible dike was built along part of the right side of the narrow channel (Figure 4). Beside it, an 80-m-large floodplain is present. Its bottom is at the same level as the bottom of the main channel. Finally, there is a wall at the downstream extremity of the main channel to reduce the flow section and regulate the water level. Shortly after that wall, the channel bottom widens to reach 30 m and water is freely released.

The inflow discharge Q_{in} is supplied by an upstream regulating gate and increases progressively until reaching a target inflow discharge of 70 m^3/s in Test 1 and 35 m^3/s in Test 2. During the filling phase, the water level rises in the main channel. Before starting the tests, a 0.5-m-deep notch was dug in the crest of the erodible dike to trigger overtopping. Its top and bottom widths are equal to 3 m and 1 m, respectively. The wall at the channel downstream extremity was designed and the upstream gate operated so that the water level overflows the initial notch bottom by 30 cm. Once the water level overtops the notch, surface erosion and breaching develop.

The main characteristics of the erodible dike and the position of the initial notch center are provided for each tested configuration in

Table ***1***. In both tests, the erodible dike is made of gravel and coarse sand.

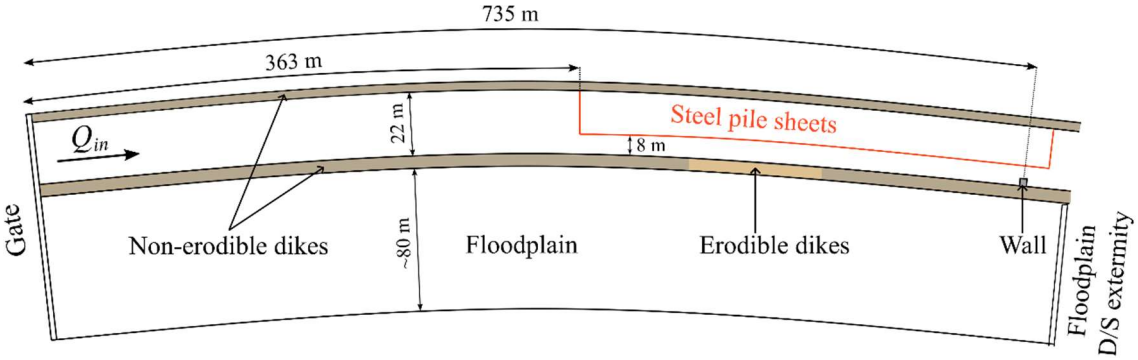


Figure 3. General scheme of the experimental setup.

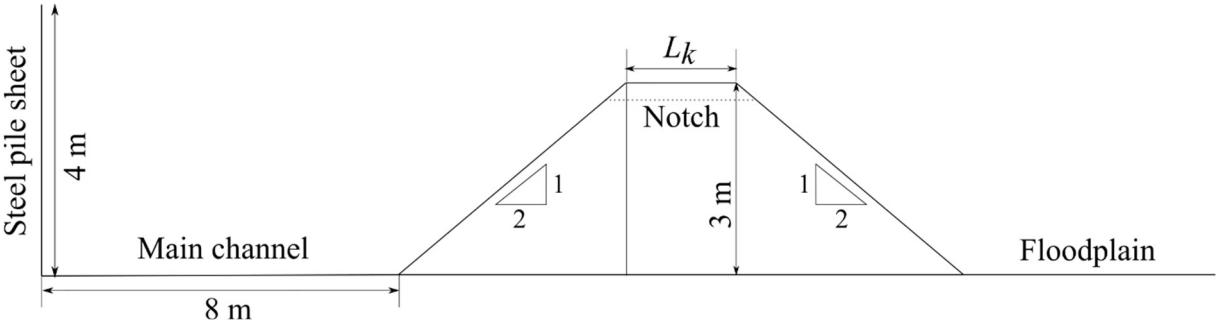


Figure 4. Erodible dike cross-section.

Table 1. Tested configurations. L_k = dike crest width; S = dike side slopes (H/V); d_{50} = median grain size of the dike material; F = Froude number in the narrow channel section; x = distance between the upstream gate and the center of the initial notch; w_d = location of the erodible dike section with respect to the initial notch.

Test ID	L_k (m)	S (-)	d_{50} (mm)	Target Q_{in} (m^3/s)	Target F	x (m)	w_d (m)
1	3	2	5.4	70	0.47	463	[-18; 62]
2	3		4.9	35	0.23	583	[-40; 44]

3.1.2 Experimental data

Water levels were recorded in the main channel (along the steel sheet piles) and in the floodplain (50 m from the sheet piles) using radio and hydraulic water gauges. Acceleration sensors were placed in the dike body to monitor its morphological evolution. At the end of each test, when the main channel got empty, the final dike morphology was recorded in more detail using three-dimensional laser scanners.

Water surface velocities were measured by radio flowmeters at the center of the channel. Data were collected every second. Average values over periods of 60 seconds were considered to smoothen the signal. Flow rates were derived from the flow cross-sections and the recorded

surface velocities using a corrective factor that considers the non-uniform velocity distribution along the river depth. In both tests, the flow rate was recorded 50 m upstream of the initial notch. Additionally, it was recorded 120 m and 50 m downstream of the notch in Test 1 and Test 2, respectively.

The discharge through the breach was deduced by Kakinuma et al. (2013) from mass balance applied to a control volume around the breach. The upstream and downstream extremities of the control volume corresponded to the flow rate measurement locations.

3.1.3 Parametrization of the test cases in the numerical models

In both modelling approaches, the adopted geometry corresponds to the one displayed in Figure ***3***, except the slight channel curvature that is discarded. In WOLF-2D, a non-erodible bed with a constant slope of 1/500 is considered in the main channel and the floodplain. The floodplain is 80 m wide and stretches along the entire dike length. A boundary condition on the Froude number, F, is prescribed at the downstream extremity of the floodplain (F = 0.75). The dike breaching event is not expected to be sensitive to this value as the flow through the breach is mostly supercritical. The other extremities of the floodplain are considered as waterproof. A Manning coefficient of 0.023 [s/m$^{1/3}$] is applied to the whole domain, as suggested by Kakinuma et al. (2013). Backwater effect may be considered in the lumped model, but it is discarded here as too little information is available in the floodplain. The spatial discretization is set to 1 m

The same inflow discharge is used in both modelling strategies. At the beginning of each simulation, no water is present in the domain. The inflow discharge is injected at the upstream extremity of the main channel, i.e., at the gate location. The total water volume injected in the main channel measured at the gate is 12% smaller than that measured 50 m upstream from the notch in Test 1 (error of 54,600 m^3), whilst it is 10% larger in Test 2 (error of 24,600 m^3). This highlights significant uncertainties in the experimental data as mass conservation is not ensured in both cases. The flow rate 50 m upstream from the notch was used by Kakinuma and Shimizu (2014) to derive the experimental breach discharge. For this reason, it was considered as the inflow discharge in all simulations. A time shift of 9 minutes is applied to take into account the time required for water to leave the gate and reach the position 50 m upstream from the notch. This time shift was determined based on numerical tests.

An explicit relation between the water level and the outflow discharge is prescribed as a downstream boundary condition in both modelling approaches. An experimental rating curve was used to fit a "weir" discharge coefficient, C_d, in the following relation:

$$Q_{out} = \frac{2}{3} C_d l_w \sqrt{g(h-w)^3}, \quad (2)$$

with Q_{out} the outflow discharge, l_w the width of the "weir" (i.e., the channel bottom width here), g the gravity acceleration, h the mean water free surface elevation over the "weir", and w the "weir" crest elevation (i.e., the channel bed elevation in this case). This relation is applied at the location where the experimental rating curves were measured, i.e., 120 m (Test 1) or 50 m (Test 2) downstream of the initial notch, so that the computational domain is trimmed at this location. No explicit representation of the downstream wall is required in this case as the experimental rating curves are systematically located upstream of it.

3.2 Hypothetical dike breaching along the Albert Canal in Wallonia (Belgium)

3.2.1 General description of the case study

The modelled area is about 50 km long and stretches from Ivoz-Ramet dam located upstream from the city of Liège (Belgium) to Genk locks, in the northern part of Belgium (Figure 5). Starting from Ivoz-Ramet, the Meuse River flows towards Liège. The Ourthe River, controlled by the Grosses Battes weir, joins the Meuse River at their confluence just upstream of Liège. After the city, most of the flow is diverted northward into the Albert Canal. A 200-m-long emergency side weir is located along the right bank of the canal, shortly after the channel division. In Lanaye, part of the flow is directed towards Maastricht (The Netherlands) via Lanaye locks, while the rest continues to the locks at Genk.

In this work, we consider the hypothetical breaching of an earthfill dike located along the left bank of the upstream part of the Albert Canal. The breach location was selected due to the high risk posed to the densely urbanized area in the floodplain and because this dike is erodible, unlike most canal embankments in this area, which are made of concrete. A scheme of the dike and Albert Canal cross-sections is provided in Figure 6. An initial notch measuring 1.2 m in depth and 4 m in width was created on the dike crest to trigger overtopping and initiate breaching at the designated location. The dike is assumed erodible, homogeneous (d_{50} = 1 mm), and non-cohesive, while its foundations are considered non-erodible.

3.2.2 Boundary conditions

The modelled scenario represents normal river discharge conditions but involves an exceptional operational state in Monsin dam, where only one out of six weirs is open. A critical flow is assumed over this weir and the emergency side weir located along the Albert Canal. An inflow discharge of 900 m^3/s and 600 m^3/s is imposed at Ivoz-Ramet dam (Meuse River) and Grosses Battes weir (Ourthe River), respectively. For comparison, the 25-year flood discharge on the downstream part of the Ourthe River is about 900 m^3/s, while the 5-year flood discharge after the Meuse-Ourthe confluence is about 1,800 m^3/s. The downstream boundary conditions correspond to an emergency operation of the Lanaye and Genk locks, each draining a prescribed discharge of 75 m^3/s, while all other locks along the Albert Canal are assumed to be closed. Using historical data from the Wallonia public administration (SPW), a Manning coefficient of 0.027 [$s/m^{1/3}$] is applied to the whole domain. The spatial discretization is set to 2 m, resulting in approximately 2.5×10^6 active mesh cells.

Figure 5. Domain and boundary conditions of the numerical simulations.

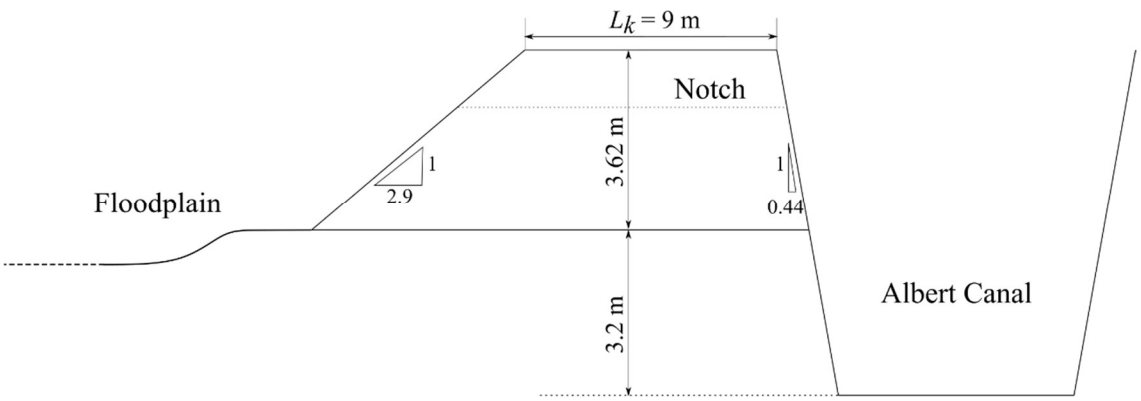


Figure 6. Cross-section of the erodible dike and Albert Canal.

4. Results and discussion

4.1 Comparison and validation of the modelling strategies

In this section, results generated by the lumped dike breach model alone and the dynamic coupling between WOLF-2D and the breach evolution module of the lumped model are discussed and compared. Additionally, two alternative approaches were adopted by using the experimental breach width at the dike crest level as an input for the lumped model and WOLF-2D.

Figure 7 shows the evolution of the breach discharge and breach expansion in the field tests conducted by Kakinuma et al. (2013). The breach discharge in Test 2 exhibits a plateau, as the inflow discharge gradually increases, stabilizes, and then decreases. In contrast, the inflow discharge in Test 1 increases initially, then steadily decreases, without forming a plateau. All modelling strategies predict the peak discharge with less than 20% error, except for the dynamic coupling in Test 1. Both modelling approaches based on the experimental breach topography provide acceptable results.

When the lumped model computes the breach expansion itself (yellow curves in Figure 7), the results are satisfactory, especially in Test 2. However, the breach width is systematically underestimated (Figure 7c and 7d). These observations indicate that, in this model, the breach conveys too much water for a given breach width, while erosion is underestimated for a given breach discharge. This trend is particularly evident when the lumped model uses the experimental breach topography as an input, i.e., the peak breach discharge is systematically overestimated. A similar observation holds when considering the dynamic coupling approach. In this case, shallow-water equations are used to compute the flow through the breach. This results in a smaller breach discharge than what is obtained with the lumped model, leading to even lower breach widening. Consequently, the breach discharge decreases as well.

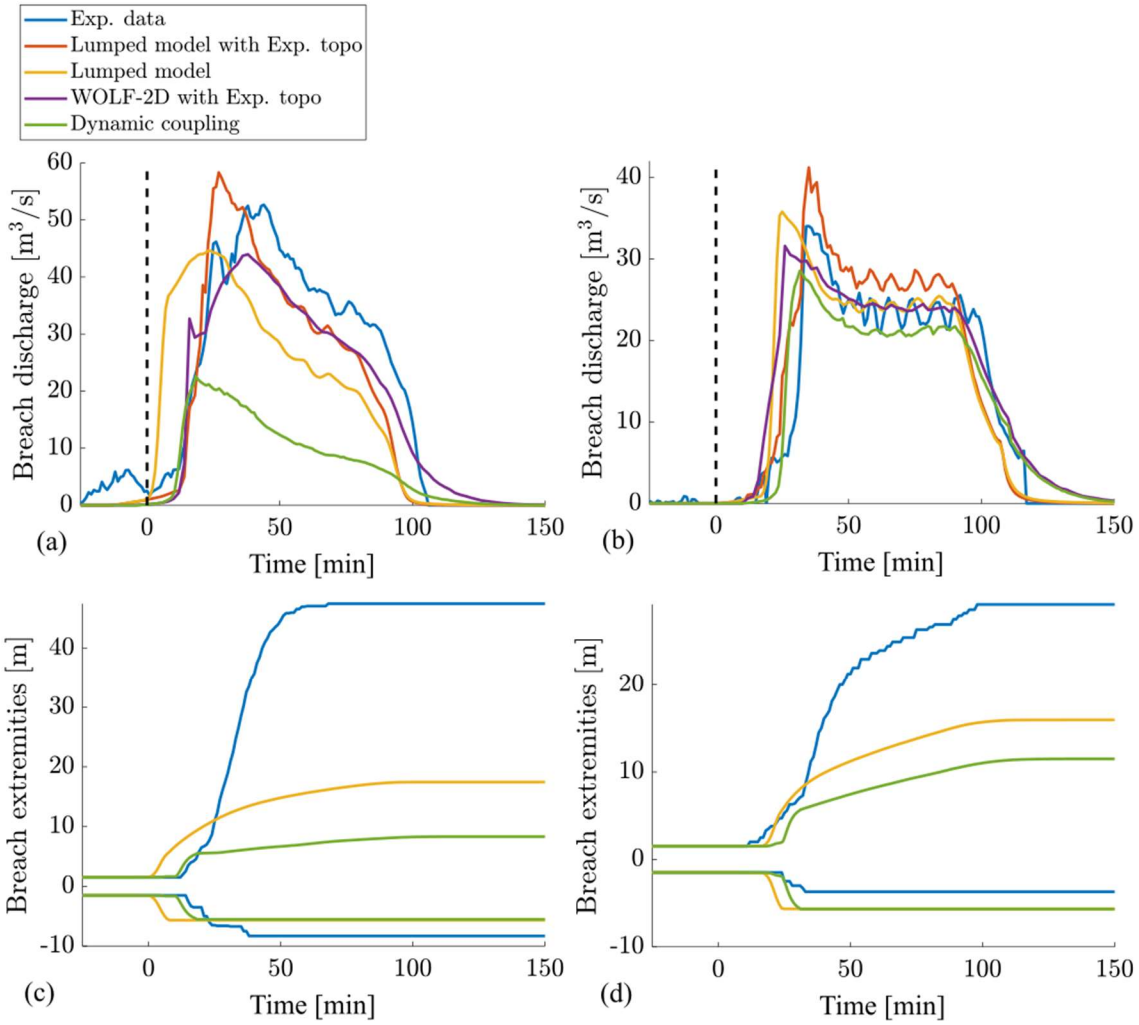


Figure 7. Time evolution of the breach discharge and expansion for Test 1 (a and c) and Test 2 (b and d) obtained experimentally and with the different modelling strategies.

To tackle these limitations, the breach evolution module of the lumped model should be revised so that the computed erosion rates intensify for a given breach width and discharge. A solution may consist in modifying the value of the effective breach width that dictates the breach erosion rate. Kakinuma and Shimizu (2014) highlighted a breach main stream width that was significantly smaller than half of the total breach width. Similarly, in the calibration of the analytical model for the prediction of the breach discharge introduced in Chapter 5 of this thesis, an effective breach width close to one-third of the total breach width was identified (Figure 4 in Chapter 5).

Alternatively, the value of erosion-related model parameters may be modified. Schmitz et al. (2023b) (Chapter 1) highlighted three parameters whose uncertainty has a critical impact on the prediction of the breach widening and, consequently, the breach discharge. These parameters were identified as a parameter involved in Strickler's formula (A_n), an empirical coefficient involved in the mixing length computation (λ), and the critical Shields parameter (θ_{cr}). In the case of fluvial dike breaching, highly turbulent flow and complex 3D flow phenomena take place close to the breach (Michelazzo et al., 2015), which may significantly increase turbulence and justify the use of adapted values for flow resistance and sediment transport parameters.

Consequently, it was decided to fix the value of the breach effective width to 33% of the total breach width. Additionally, the value of A_n and θ_{cr} was decreased by 20%, whilst the value of λ dropped from 3 to 0, corresponding to sediment concentration reaching equilibrium instantaneously.

Figure 8 compares the results obtained for the different parametrizations of the lumped model that is coupled with WOLF-2D. In this figure, the dam configuration refers to results obtained with the initial parametrization of the lumped model, i.e., the model without the use of an effective breach width definition and with the default values of the erosion parameters. The configuration with the adjusted parametrization refers to the case where the value of the flow resistance and sediment transport parameters was modified.

Overall, the dike configurations systematically perform better and are more conservative than the dam configuration, i.e., the initial model proposed by Wu (2013), both for the breach discharge and expansion. Among the configurations tested, the dike configuration with b_{eff} = 33% and adapted parameters yields the best performance in both test cases. Numerical and experimental results are in particularly good agreement in Test 2, where the breach widening and breach peak discharge are accurately captured. The velocity field across the breach obtained for this test further supports the choice of reducing the effective width to a value below 50% (Figure 9).

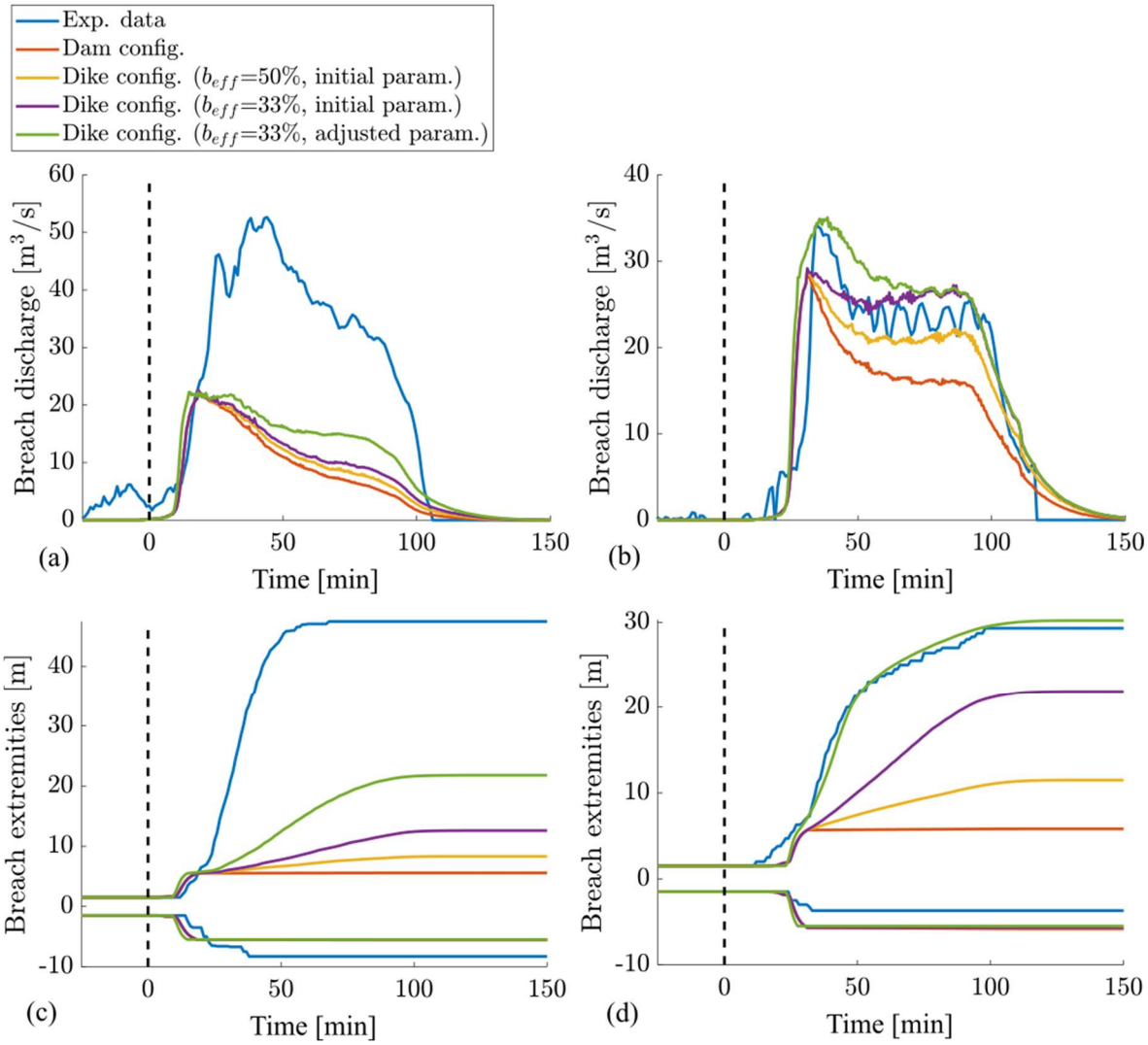


Figure 8. Time evolution of the breach discharge and expansion for Test 1 (a and c) and Test 2 (b and d) obtained for different parametrizations of the lumped model coupled with WOLF-2D.

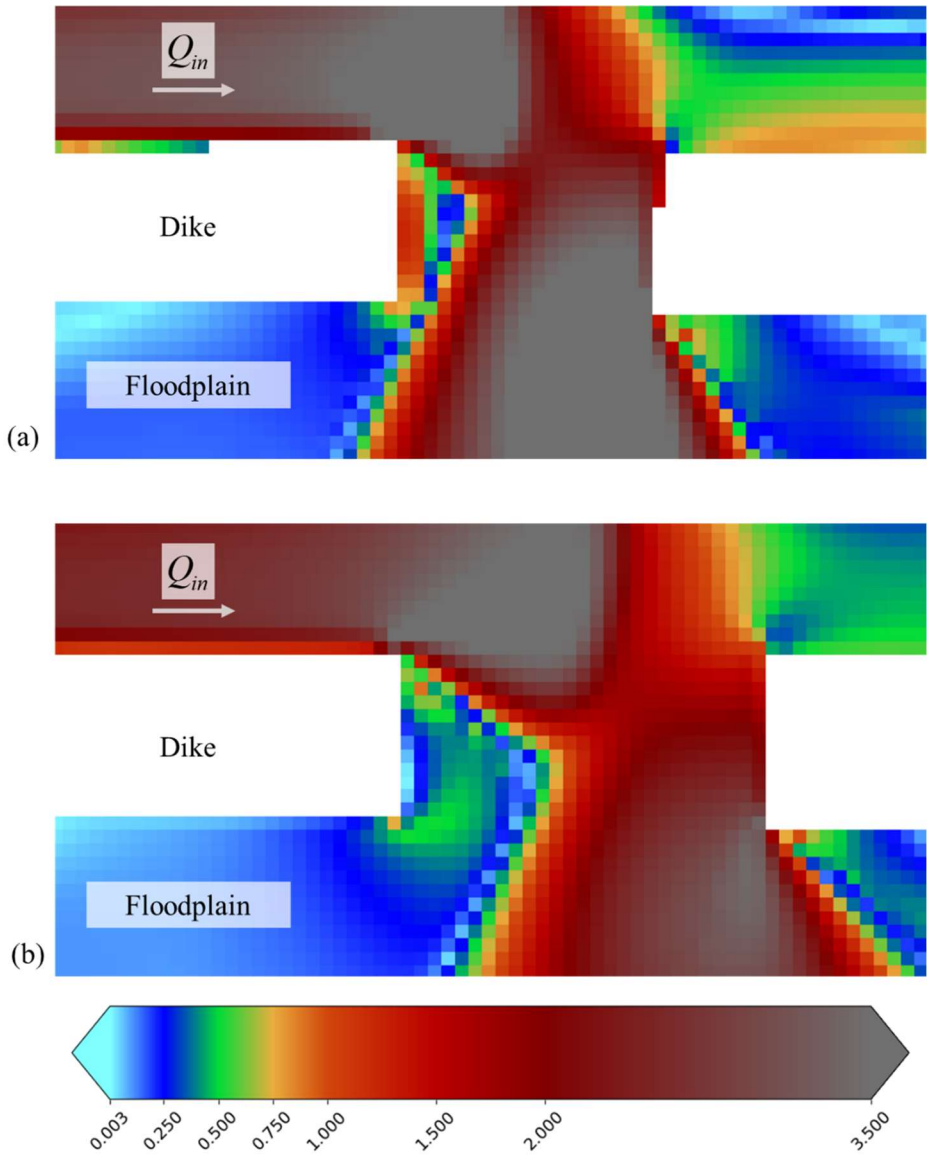


Figure 9. Velocity field in the breach vicinity at (a) time = 50 [min] and (b) time = 100 [min]. The norm of the velocity is displayed in [m/s].

Although improving the results, the new parametrization of the lumped model still fails to deliver fully satisfactory outcomes, as demonstrated by the numerical results for Test 1. It is worth noting that inaccuracies are inherent to experimental tests, especially when conducted at a large scale (e.g., complex monitoring of the entire experimental setup, inaccuracy in the measuring devices, uncontrolled parameters such as water leaks…). These limitations should thus be considered when comparing numerical results and experimental data. In particular, the uncertainty related to the experimental breach discharge and downstream rating curve might have a significant impact on the results, leading to larger discrepancies between numerical and experimental data than what might occur under ideal conditions.

In Test 1, the total water volume measured at the gate was significantly lower than the volume measured just upstream of the dike, which Kakinuma et al. (2013) used to compute the experimental breach discharge (396,000 m^3 versus 451,000 m^3). This discrepancy suggests that

the breach discharge may have been overestimated. Conversely, in Test 2, the water volume measured at the gate exceeded the one measured upstream of the dike (278,000 m^3 versus 253,000 m^3), indicating a potential slight underestimation of the breach discharge estimated by Kakinuma et al. (2013). These observations may help explain why the numerical results tend to underestimate the breach discharge in Test 1 and overestimate it in Test 2.

4.2 Dike breaching on the Albert Canal in Wallonia (Belgium)

In this section, a hypothetical breach along the Albert Canal (Section 3.2) is modelled using the lumped dike breaching model coupled with WOLF-2D. Results obtained with the four lumped model parametrizations presented in Section 4.1 are compared, namely (A) dam configuration, i.e., the initial model proposed by Wu (2013), (B) dike configuration with b_{eff} = 50%, (C) dike configuration with b_{eff} = 33%, and (D) dike configuration with b_{eff} = 33% and adjusted parameters. The breach discharge evolution, breach expansion and water volume in the floodplain are plotted in Figure 10. Flood maps obtained 1.5 hours, 3 hours, and 6 hours after breaching initiation for each parametrization are provided in Figure 11.

When b_{eff} = 33%, the breach discharge experiences a sharp peak discharge quickly after breaching onset, which is concomitant with a quick expansion of the breach. The peak discharge is maximum when using adjusted flow resistance and sediment transport parameters, and the breach reaches about twice the width of the configuration with the default parameters values. Shortly after breaching initiation, both models predict a plateau in the breach discharge, which then drops to a stabilized value. Similarly, the breach expands much more slowly after the peak discharge, and then gets fixed about 3 hours after breaching onset. The dike configuration with default model parameter values and b_{eff} = 50% predicts similar results, except that no sharp global maximum in the breach discharge is observed. The resulting breach expansion is also significantly reduced.

The dam configuration leads to very different results. In this case, the breach expands very slowly, which leads to a gentle increase in the breach discharge, before dropping back to a stabilized value similar to the one reached by the dike configurations. This may be attributed to the fact that this model considers the whole breach section as “effective”, i.e., b_{eff} = 100%. The differences in the flood maps can be easily interpreted based on the evolution of the breach discharge and the resulting water volume stored in the floodplain. Among the dike configurations, the flooded area is much larger shortly after the breaching initiation when selecting b_{eff} = 33% and adjusted model parameters. However, the difference fades over time and almost disappears after about 6 hours. This observation is in good agreement with the evolution of the water volume stored in the floodplain, which becomes similar after some time for the three dike configurations.

The dam configuration predicts a much smaller increase in the water level in the floodplain. This difference fades with increasing time, but much more slowly than it does between dike configurations. As suggested in Section 4.1, this approach underestimates the breach expansion, which leads to the underestimation of the breach discharge and the related dynamic evolution of the flood extent. To remain conservative, it is greatly advised to opt for a dike configuration model, with b_{eff} = 33% and adapted model parameters. This approach allows for a safer planning

of the evacuation zone around the breach on a short-term basis, whilst leading to similar results as other dike breach model parametrizations over longer periods.

In the present case study, the floodplain is densely urbanised. Several critical infrastructures are exposed to the flood, such as health facilities, police and firefighter stations, as well as water management infrastructures. Figure 12 provides an overlap of a vulnerability map with the flooded area (shaded area) obtained at equilibrium, when the water level in the floodplain reaches the level of a dike crest located further downstream on the Albert Canal. At this stage, the water depth exceeds 5 m locally, and many critical infrastructures are impacted. This situation is reached about 95 hours and 101 hours after breaching initiation for the dike configurations and the dam configuration, respectively.

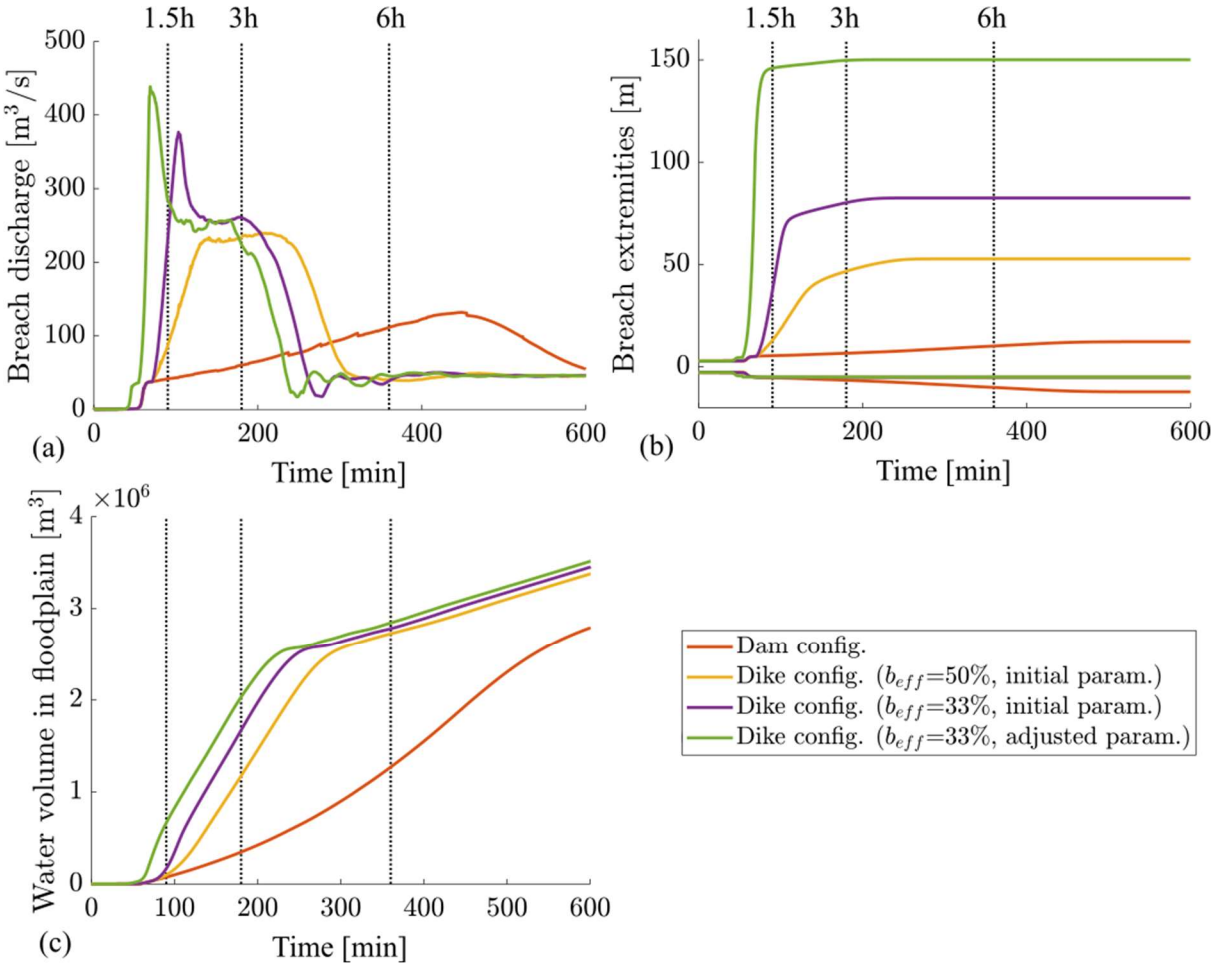


Figure 10. Evolution of (a) the breach discharge, (b) breach top width, and (c) water volume in the floodplain as a function of the coupled model parametrization.

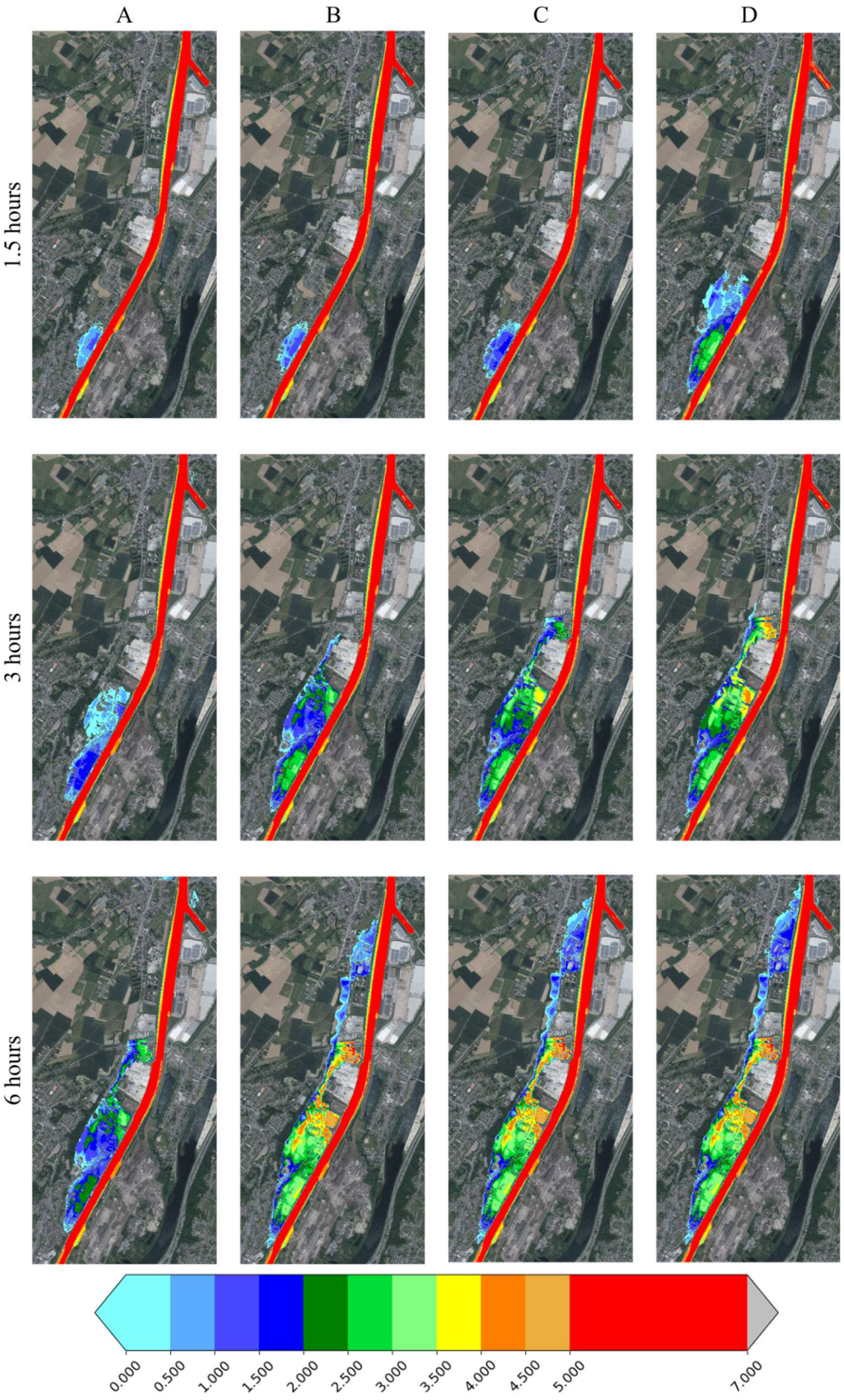


Figure 11. Evolution of the flooded area and water depths when the coupled model is used (A) in dam configuration, (B) dike configuration with default model parameters and b_{eff} = 33%, (C) dike configuration with default model parameters and b_{eff} = 50%, and (D) dike configuration with adjusted parameters and b_{eff} = 33%. The colormap corresponds to the water depth in [m].

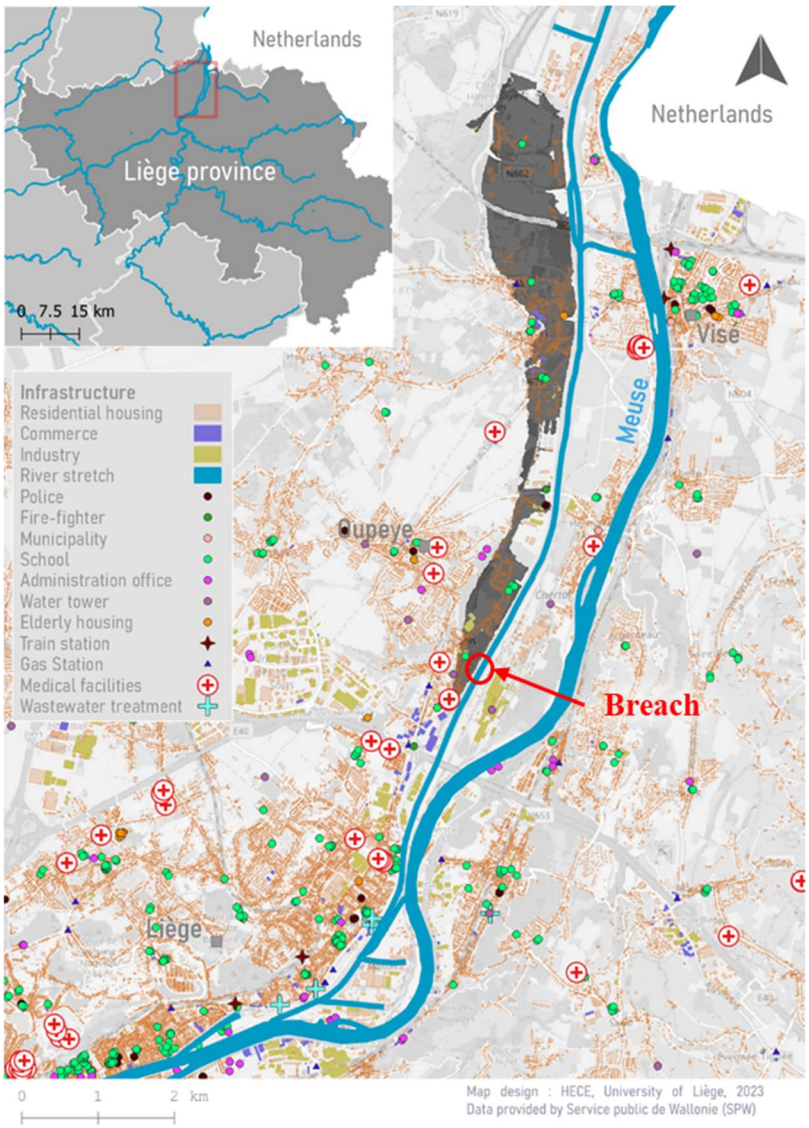


Figure 12. Vulnerability map with flooded area (shadow) induced by the considered dike breach.

5. Conclusion

This study has explored two modelling strategies for dike breaching. The first modelling approach consists in describing the flow and the morphologic evolution of the dike breach using a lumped dike breach model. In the second modelling approach, a 2D hydrodynamic model (WOLF-2D) is dynamically coupled with the breach evolution module of the lumped dike breach model, whose goal is then limited to the description of the dike breach morphology. Additionally, two alternative approaches were adopted by using the experimental breach width at the dike crest level as an input for the lumped model and WOLF-2D. These approaches allowed focusing on the capability of each model to predict the breach discharge, independently of the performance of the erosion module.

The predictive capability of each modelling strategy was evaluated using two field-scale experiments conducted by Kakinuma et al. (2013), with a focus on the evolution of the breach

discharge and expansion. The coupled approach between the lumped model and WOLF-2D was then applied to a hypothetical full-scale breach scenario along the Albert Canal in Belgium.

When compared to the field-scale experiments conducted by Kakinuma et al. (2013), all modelling approaches led to satisfactory results regarding the peak discharge at the breach, except for the coupled approach, which exhibited a poorer performance. Furthermore, all models underestimated the breach widening.

Based on experimental observations that highlighted the presence of complex 3D flow structures in the vicinity of the breach (Michelazzo et al., 2015), the model parametrization was adapted by adjusting the value of three flow resistance and sediment transport parameters whose uncertainty had a critical impact on the prediction of the breach widening (Schmitz et al., 2023b). In addition, the value of the effective breach width, b_{eff}, i.e., the fraction of the breach width that conveys most of the flow, was decreased from 50% to 33%. This was supported by experimental and numerical results demonstrating that the flow significantly concentrates close to the downstream extremity of the breach (Kakinuma and Shimizu, 2014; Charrier, 2015).

The lumped dike breaching model coupled with WOLF-2D demonstrated that the dike configuration with b_{eff} = 33% and adjusted parameters provided the most reliable and conservative results. This configuration accurately captured the peak discharge and breach widening, which are critical for assessing flood risk. The dam configuration, i.e., assuming b_{eff} = 100%, with default model parameter values significantly underestimated the breach expansion and discharge, highlighting its limited predictive capability for fluvial dike breaching cases.

Although improving the results, the new parametrization of the lumped dike breaching model still failed to deliver fully satisfactory outcomes. These discrepancies between experimental and numerical results may be partially explained by the substantial uncertainties in the experimental measurements, which were highlighted by a large mass imbalance in the main channel.

In the hypothetical breach scenario along the Albert Canal, the lumped dike breaching model coupled with WOLF-2D showed that all dike configurations led to similar flood extents over a long-term period, whilst results significantly differed for shorter periods after breaching initiation. Among the dike configurations, the one using b_{eff} = 33% and adjusted model parameters ensured more conservative flood predictions. In contrast, the flood extent and breach widening were much smaller when predicted by the dam configuration model with default parametrization. Given the densely urbanized floodplain and the presence of critical infrastructures, such as health facilities and emergency services, the dike configuration with b_{eff} = 33% and adjusted parameters is recommended to support safer evacuation planning. These findings emphasize the importance of selecting appropriate breach modelling approaches to enhance flood risk management in highly vulnerable areas.

To further improve the model predictive capabilities, future work should focus on defining in more detail the velocity field through the breach to better calibrate the value of the effective breach width and the model parametrization, which both dictate the breach erosion rate. 3D numerical simulations and laboratory tests could help guiding this characterization.

References

Charrier, G., 2015. Etude expérimentale des ruptures de digues fluviales par surverse (Doctoral dissertation, Aix-Marseille).

Chen, Z.Y., Ping, Z.Y., Wang, N.X., Yu, S., Chen, S.J., 2019. An approach to quick and easy evaluation of the dam breach flood. Science China Technological Sciences 62, 1773–1782.

Dazzi, S., Vacondio, R., Mignosa, P., 2019. Integration of a Levee Breach Erosion Model in a GPU-Accelerated 2D Shallow Water Equations Code. Water Resources Research 55, 682–702.

De Lorenzo, G., Macchione, F., 2014. Formulas for the peak discharge from breached earthfill dams. Journal of Hydraulic Engineering 140, 56–67.

Hu, P., Ji, A., Li, W., Cao, Z., 2023. Numerical modelling of levee breach with an improved slope-failure operator. Journal of Hydraulic Research 1–13.

Kakinuma, T., Shimizu, Y., 2014. Large-scale experiment and numerical modeling of a riverine levee breach. Journal of Hydraulic Engineering 140.

Kakinuma, T., Tobita, D., Yokoyama, H., Takeda, A., 2013. Levee breach observation at Chiyoda experimental flume, in: 12thInternationalSymposium River Sedimentation (ISRS), IRTCES, Kyoto, Japan.

Lee, K., 2019. Simulation of Dam-Breach Outflow Hydrographs Using Water Level Variations. Water Resources Management 33, 3781–3797.

Li, Y., Chen, A., Wen, L., Bu, P., Li, K., 2020. Numerical simulation of non-cohesive homogeneous dam breaching due to overtopping considering the seepage effect. European Journal of Environmental and Civil Engineering.

Michelazzo, G., Oumeraci, H., Paris, E., 2015. Laboratory Study on 3D Flow Structures Induced by Zero-Height Side Weir and Implications for 1D Modeling. Journal of Hydraulic Engineering 141, 04015023.

Peter, S.J., Siviglia, A., Nagel, J., Marelli, S., Boes, R.M., Vetsch, D., Sudret, B., 2018. Development of Probabilistic Dam Breach Model Using Bayesian Inference. Water Resources Research 54, 4376–4400.

Rifai, I., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., El Kadi Abderrezzak, K., Dewals, B., 2017. Overtopping induced failure of noncohesive, homogeneous fluvial dikes. Water Resources Research 53, 3373–3386.

Schmitz, V., Arnst, M., Abderrezzak, K. El kadi, Pirotton, M., Erpicum, S., Archambeau, P., Dewals, B., 2023b. Global sensitivity analysis of a dam breaching model: To which extent is parameter sensitivity case-dependent? Water Resources Research e2022WR033894.

Schmitz, V., Rifai, I., Kheloui, L., Erpicum, S., Archambeau, P., Violeau, D., Pirotton, M., El Kadi Abderrezzak, K., Dewals, B., 2023a. Main channel width effects on overtopping-induced non-cohesive fluvial dike breaching. Journal of Hydraulic Research 61, 601–610.

Tsai, C.W., Yeh, J., Huang, C., 2019. Development of probabilistic inundation mapping for dam failure induced floods. Stochastic Environmental Research and Risk Assessment 33, 91–110.

Van Damme, M., 2020. An analytical process-based approach to predicting breach width in levees constructed from dilatant soils. Natural Hazards 101, 59–85.

Wu, W., 2013. Simplified physically based model of earthen embankment breaching. Journal of Hydraulic Engineering 139, 837–851.

Wu, W., 2016. Introduction to DL Breach–A simplified physically based dam/levee breach model. Clarkson University, NY.

Zhong, Q., Wang, L., Chen, S., Chen, Z., Shan, Y., Zhang, Q., Ren, Q., Mei, S., Jiang, J., Hu, L., Liu, J., 2021. Breaches of embankment and landslide dams - State of the art review. Earth-Science Reviews 216.

Conclusion

1. Summary

Modelling the breaching of fluvial dikes involves many challenging aspects. In this thesis, we build upon an existing spatially non-discretized dam breaching model (DLBreach) and adapted it to the fluvial case. It was then coupled with a 2D hydrodynamic model (WOLF-2D) that better represents the flow characteristics.

When adapting a numerical model, a deep understanding of the model structure and the dependencies between inputs and outputs is essential to properly interpret results, assess their reliability in light of uncertainties, and spot model features that can be improved. Chapter 1 introduces a global sensitivity analysis and an uncertainty quantification that helped address these questions. This analysis ranked input variables based on their contribution to the variability of outputs, namely the maximum breach discharge and the time to reach this peak. The analysis was conducted on various dam configurations at both laboratory and field scales. Depending on the considered case study, the impact of input parameter uncertainties on output variability varied, as did the magnitude of output uncertainties. The results were summarized in a flow chart enabling the practical determination of the input parameter whose uncertainty is the most critical in each configuration and what standard deviation in the output variables is expected. Among the 26 input parameters involved in the model, only five were identified as particularly critical when predicting the breach discharge. These parameters corresponded to the inflow discharge, the discharge coefficient used to estimate the breach discharge, a parameter involved in Strickler's formula, an empirical coefficient used in the mixing length computation, and the critical Shields parameter.

Based on an in-depth understanding of the initial model, we identified the components that should be altered to better reflect the breaching of fluvial dikes. To this end, experimental tests were analyzed to highlight key physical phenomena omitted or poorly represented in the initial model and in previous experimental campaigns. Although conducted at laboratory scale, these experiments allowed for a better understanding of the physical processes controlling dike breaching. Chapters 2 and 3 demonstrate the impact of the dike geometry and main channel width on breaching dynamics, which had not been systematically investigated before. Chapter 2 highlights that the temporal evolution of the breach discharge may follow three distinct patterns, depending on the non-dimensional dike cross-section and the Froude number in the main channel. It also proposes a simplified conceptual model suggesting that dike breaching might be seen as the superposition of a dam break flow and an open channel flow. Chapter 3 confirmed this trend by showing that the influence of river inertia and river dimensions, i.e., storage volume, compete during a dike breaching event. In particular, river momentum concentrates the flow across the breach toward its downstream extremity, leading to a highly non-uniform velocity field and resulting in asymmetrical breach expansion. The initial model was adapted to account for this observation by introducing the concept of effective breach width, which represents the fraction of the breach width that conveys most of the breach flow. This modification increased erosion intensity at the downstream extremity of the breach, leading to wider breaches. Although still underestimating breach width, this modification greatly improved the results compared to the ones obtained with the initial model. However, the breach

discharge was significantly overestimated by the model conveyance formula, indicating that a more accurate breach width prediction would introduce additional errors in discharge predictions.

Chapters 4 and 5 focus on improving the prediction of the breach discharge. In Chapter 4, eleven empirical and semi-empirical side weir equations were applied to various experimental configurations. Their performance varied with the test case, but none demonstrated a consistently better accuracy. These formulas were derived based on simplified experimental setups (prismatic weir shape, sharp-crested weir, hydrostatic pressure distribution…). However, real dike breaching feature complex breach shapes and 3D flow patterns that were not considered during the development of the tested empirical formulas. This may explain why they failed to properly capture the breach discharge evolution. To address this limitation, Chapter 5 introduces machine learning techniques for breach discharge prediction. When properly trained, machine learning proved particularly suitable for capturing highly non-linear physical phenomena. Three predictive approaches were compared:

1. direct prediction using decision-tree-based machine learning techniques,
2. direct prediction using empirical formulas developed for simplified configurations,
3. a new analytical approach incorporating an empirical parameter computed based either on machine learning or on empirical regressions.

Machine learning and the new analytical model outperformed the empirical formulas. Combining empirical models with a machine-learning-based corrective term considerably improved the accuracy of such models. Notably, the new analytical approach combined with machine learning yielded highly accurate results, with potentially better generalization beyond the training space compared to machine learning algorithms alone.

Chapter 6 concludes this work by comparing the performance of the lumped model adapted to the fluvial case with that of the dynamically coupled 2D hydrodynamic model. Additionally, both approaches were compared to the original dam breaching model. When compared to field-scale experiments, all dike breaching models consistently outperformed the dam breaching model in predicting breach discharge and expansion, while being systematically more conservative. However, all models underestimated breach widening. Based on experimental and numerical observations and insights from the global sensitivity analysis in Chapter 1, the model parametrization was adapted by adjusting the value of critical parameters affecting breach widening predictions. The adapted dike model provided the most reliable and conservative results when validated against field-scale experimental data. This approach was applied to a real-world hypothetical breach along the Albert Canal in Belgium and produced more conservative flood extent predictions, making it particularly valuable for supporting evacuation planning in floodplains.

2. Perspectives

In future work, the accuracy of the numerical tools should be improved by further revising the dike breaching model structure and parametrization. Erosion formulas and erosion-related parameters should be enhanced to better account for complex 3D flow phenomena near the breach. 3D simulations could be particularly valuable for this purpose. They could also help better characterize the flow field across the breach, further supporting the use of an effective breach width in the simpler models. Similarly to the initial dam breaching model, this work assumed a trapezoidal breach shape. The influence of this assumption could also be tested by considering more realistic breach cross-sections, e.g., ellipsoidal. When using the lumped dike breaching model alone, further improvements in breach discharge predictions could be achieved by developing new empirical formulas or employing advanced machine learning techniques, such as physics-guided deep learning.

Expanding experimental datasets will be key to enhancing these empirical formulations. Additionally, more experimental test campaigns would provide deeper insights into the physical processes underlying dike breaching. These could focus on parameters whose influence was not yet systematically investigated, such as dike compaction, erodibility of the riverbed/floodplain, etc. Experimental setups must also better reflect real-world dike breach dynamics. One often overlooked feature is the main channel's storage volume prior to breaching. In laboratory setups, this volume quickly depletes after breaching initiation if inflow discharge remains constant, leading to an artificial drop in water level. In real scenarios, river storage volume is much larger, reducing this drawdown. Future laboratory experiments should consider using adaptative inflow discharges for more realistic water level evolution.

To overcome scaling effect, field-scale experiments remain necessary. However, data from such experiments are limited, and their reliability is often constrained by inherent measurement uncertainties. Despite these challenges, large-scale tests are critical for understanding the scaling effects in dike breaching and improving the interpretation of lab-scale experimental results.