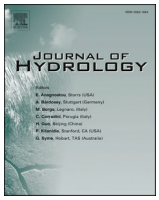
Journal of Hydrology 620 (2023) 129511

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol



Research papers

Quantifying flood model accuracy under varying surface complexities

W. Addison-Atkinson [a,*], A.S. Chen [a], M. Rubinato [b], F.A. Memon [a], J.D. Shucksmith [c]

[a] Centre for Water Systems, Harrison Building, University of Exeter, EX4 4QF, UK
[b] Centre for Agroecology, Water and Resilience, Coventry University, CV8 3LG, UK
[c] Department of Civil and Structural Engineering, University of Sheffield, S1 3JD, UK

ARTICLE INFO

This manuscript was handled by S. Sally Elizabeth Thompson, Editor-in-Chief, with the assistance of Tessa Maurer, Associate Editor

Keywords:
Dual drainage
Flow exchange
Model validation
Surface flow

ABSTRACT

Floods in urban areas which feature interactions between piped and surface networks are hydraulically complex. Further, obtaining in situ calibration data, although necessary for robust simulations, can be very challenging. The aim of this research is to evaluate the performance of a commonly used deterministic 1D-2D flood model, calibrated using low resolution data, against a higher resolution dataset containing flows, depths and velocity fields; which are replicated from an experimental scale model water facility. Calibration of the numerical model was conducted using a lower resolution dataset, which consisted of a simple rectangular profile. The model was then evaluated against a dataset that was higher in spatial resolution and more complex in geometry (a street profile containing parking spaces). The findings show that when the model increased in scenario complexity model performance was reduced, though most of the simulation error was $< 10\%$ (NRMSE). Similarly, there was more error in the validated model that was higher in spatial resolution than lower. This was due to calibration not being stringent enough when conducted in a lower spatial resolution. However, overall the work shows the potential for the use of low-resolution datasets for model calibration.

1. Introduction

Over the past decades, heavy rainfall events have become more frequent due to a warming climate (Jongman et al., 2014; Hall et al., 2014; Stewart et al., 2015; Rubinato et al., 2019), and had a worldwide significant impact causing social (Dewan, 2015) and economic losses (Svetlana et al., 2015). Moreover, multiple small-scale urban flooding events repeatedly occur in cities where the existing drainage is very old, subject to blockages (Hillas, 2014), or was designed to convey smaller volumes of flows now increased due to urbanisation. Many hydrodynamic models have been developed to understand the risks of urban flooding (Estrela and Quintas, 1994; Bates and De Roo, 2000; Liang, 2010; Rubinato et al, 2019). However, due to the complex flow processes in urban areas when inundation occurs, there is still space for improvements and further developments (Bates, 2004; Aronica et al., 2002; Rubinato et al., 2021). In particular, the difficulty in acquiring appropriate calibration/validation data is an ongoing concern. Interaction of flows between sewer systems and urban streets have three dimensional features and rapidly change in time; hence hydrodynamic models have the challenge to represent a variety of alternating unsteady flow regimes identifying flood routes within complex topography. Depth averaged two-dimensional (2D) models solving the full shallow water equations are commonly developed to simulate flooding events and estimate the surface runoff component. While pipe network models (1D) are commonly applied to quantify the sewer flows (Chen et al., 2007; Vojinovic and Tutiilic, 2009; Leandro et al., 2009; Maksimovic et al., 2009; Cea et al., 2010; Pan et al., 2012).

Previous studies have been conducted to assess the accuracy of these 2D models to replicate surface flow dynamics typical of flooding scenarios by comparing numerical simulations against laboratory experiments of flooded urban crossroads (Mignot et al., 2008a,b; Ghostine et al., 2009, 2010; El Kadi Abderrezzak et al., 2011; Bazin et al., 2017), or schematic flooded urban areas (Soares-Frazão and Zech, 2008; Van Emelen et al., 2012; Mignot et al., 2006) or against surveillance cameras and flooding information extracted from CCTV footage (Leitão et al., 2018; Moy de Vitry et al., 2017, 2018). However, systematic studies investigating flood model error utilising measured data under different scenarios are uncommon. Such studies are expensive and the accuracy of the analysis is reliant on the quality of the video footage and technician skill (Wirahadikusumah et al., 1998).

Calibrating and validating any hydrodynamic model is a crucial aspect of the modelling framework. However, an important debate that

* Corresponding author.
E-mail address: wa259@exeter.ac.uk (W. Addison-Atkinson).

https://doi.org/10.1016/j.jhydrol.2023.129511
Received 30 August 2022; Received in revised form 1 February 2023; Accepted 6 April 2023
Available online 13 April 2023

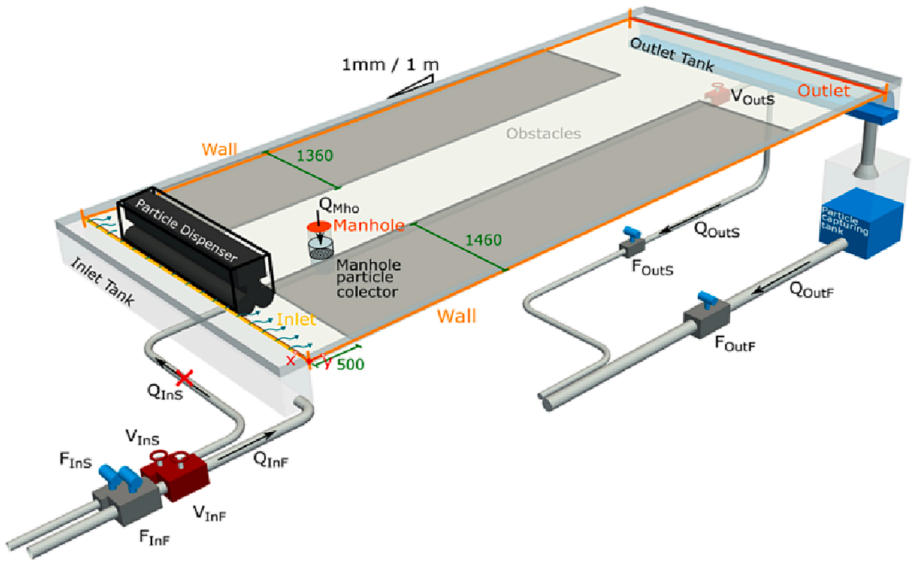


Fig. 1. 3D representation of configuration 1 in the experimental facility. All dimensions are in mm. (Image cited from Rubinato et al. (2022)).

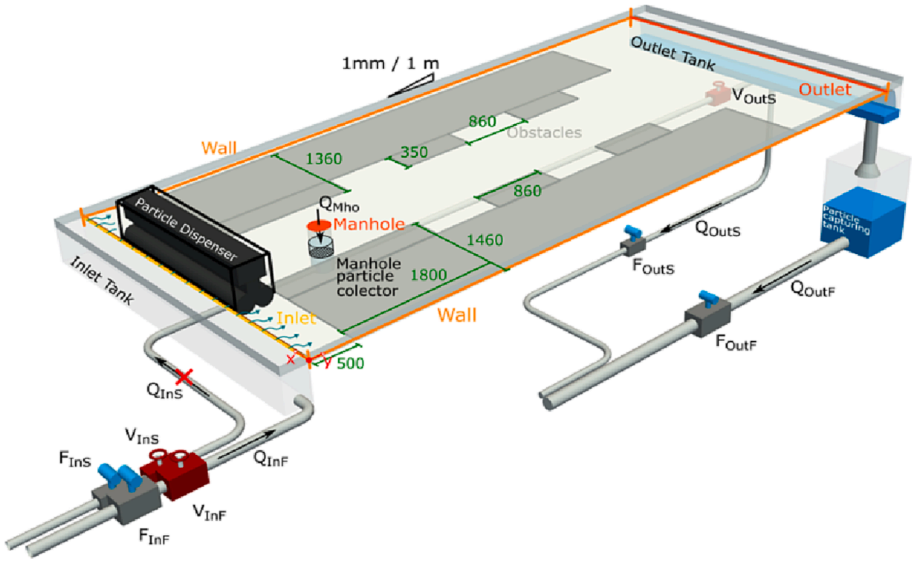


Fig. 2. 3D representation of configuration 2 in the experimental facility. All dimensions are in mm. (Image cited from Rubinato et al. (2022)).

previous studies possess include a lack of empirical data that is needed for calibrating and validating hydraulic simulations (Hunter et al., 2008; Collender et al., 2016). Consequently, the accuracy of any flood model is directly related to the possession of high-quality data (Molinari et al., 2019). Relevant surface flow data includes flood depths and extents, inundation times and flow velocities. Data in a piped systems includes pipe flows and depths,. Though, there are many challenges when regarding data availability and acquisition.

Therefore scale model facilities are often used effectively to investigate hydraulic behaviour of flood flows. Further, calibrating a flood model which can entail running several simulations can be very computationally demanding. This is especially apparent when using 1D-2D models (Addison-Atkinson et al., 2022). However, the relationship between calibration procedure (i.e., complexity and detail available) and model performance has yet to be fully explored in the literature. Many hydrological studies have identified that higher resolution grids include a higher definition of small streams, roads (and other permeable and impermeable urban surfaces) and narrow flow pathways.

However, some previous studies (Mateo et al., 2017; Kim et al., 2021) have found that increasing a models spatial resolution may not be beneficial in every case due to increased computational times that are needed. Along similar lines, numerical complexity has been known to be a significant source of error in complex flow regimes. Willis et al. (2019) argued this point and found that simplified models can be just as useful as more complex models due to the error higher resolution DEM's offer. It is true that high resolution DEM's are an integral base layer for modelling flood probability maps. Yet, flood modelling is more convoluted than this. Another example has been argued by Avand et al. (2022). The authors point out that DEM's cannot simply improve model accuracy alone. Many other factors (e.g., altitude, precipitation, and distances from water courses/flood risk zones) affect simulation outputs.

This leads to a research question: how does flood model performance differ under varying surface complexities? And can a lower resolution model still simulate this process with acceptable error? This is important to consider because the spatial detail, or resolution, of a dataset will undoubtedly affect how well a model represents reality. Similarly computational times can significantly increase in more complex models.

To address this question, the current work aims to explore the

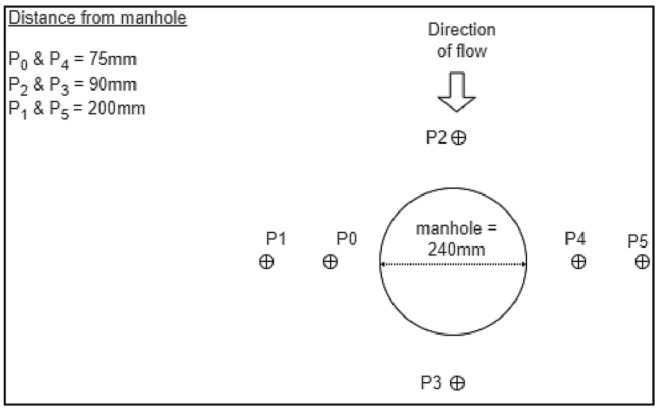


Fig. 3. Location of pressure transducers around manhole (distances taken from the edge of the manhole). Direction of Q1 surface flow and distance from manhole is shown.

relationship between 1D-2D flood simulation complexity and model error in the context of urban flooding; utilising high resolution datasets form a large-scale physical model representing piped and surface systems. And specifically access the performance of a model which is calibrated with data collected at a courser resolution than the validation dataset.

2. Methods

2.1. Water laboratory experiments

Flow rates, water depths and velocity fields were obtained using a physical scale model of a sewer pipe system (minor system) linked to a hypothetical urban surface (major system) via a single manhole. Built at the University of Sheffield, and 1:6 to scale, this unique water laboratory was used for multiple studies (Martins et al., 2017; Rubinato et al., 2017, 2018a,b; Beg et al., 2018) for the calibration and validation of many numerical models. Open access datasets associated with the facility are available at https://zenodo.org/communities/floodinteract/.The full facility is constructed from acrylic, it is 8.2 m long and 4 m wide, and the street profile replicated has width 1.18 m. The minor system does not have a slope, although the major system has a slope of 0.001. The manhole that links both systems together have a diameter of 0.240 m (simulating a 1.440 m manhole at full scale). The sewer pipe has a diameter of a 0.075 m (internal) and made out of clear acrylic (simulating a 0.450 m pipe at full scale). Electro-magnetic flow meters are used to record flows at the surface flow inlet (Q_1), the sewer flow inlet (Q_3) and at the outlets of the surface (Q_2). Two surface configurations were used; the first was a rectangular street profile (configuration 1: See Fig. 1), the second was a street with parking spaces (configuration 2: See Fig. 2). Pressure head was recorded by several pressure transducers, which were placed at different locations around the manhole. From the recorded output depths could be obtained (Fig. 3).

2.2. Hydraulic testing conditions

Three different manhole testing conditions were selected to create hydraulic conditions in the facility, for both surface configurations. Manhole conditions vary to consider different degrees of interactions between the major and minor systems. These are as follows:

1. The manhole is covered by a solid lid that significantly restricts the flow entering the minor system from the surface. Some flow is however bypassing a small gap around the edge of the lid and entering the piped system below from the surface. There is no surcharge from the pipe network, hence interaction between major and minor systems is minimal (S1).
2. The lid on the top of the manhole is removed, Flow enters the minor system freely from the surface, as the flow in the pipe is minimal (S2).
3. The lid on the top of the manhole is removed. There is significant flow in the sewer pipe, resulting in surcharge from the minor to major system (S3).

Each manhole condition was tested under three different hydraulic conditions. This involved increasing either the flow running on the surface (Q_1) or the flow entering the sewer pipe (Q_3). Measured (time averaged) hydraulic testing conditions under each test are shown in Table 1.

In this research all the tests were generated under steady conditions. At the upstream boundary of the flood plain, flows (Q_1) were discharged into the system by passing over the inlet weir moving towards the downstream section of the surface outlet tank (Q_2). The validation of the flow meters was assessed by volumetric discharge readings that were obtained with the laboratory measurement tank. LabviewTM software-controlled butterfly flow control valves, which allowed the flow release in the pipes that fed the minor and major system. This therefore

Table 1
Time averaged hydraulic parameters collected for each experimental test. Surface inflow (Q_1), surface outflow (Q_2), pipe inflow (Q_3), and pipe outflow (Q_4) and six surface depths (P_0-P_5) are shown.

S	Test ID	Q_1 (l/s)	Q_2 (l/s)	Q_3 (l/s)	Q_4 (l/s)	P_0 (mm)	P_1 (mm)	P_2 (mm)	P_3 (mm)	P_4 (mm)	P_5 (mm)
Rectangular Configuration (*Conf1*)											
S1	1	3.68	2.67	3.59	4.60	20.6	19.9	22.7	18.8	21.6	19.9
	2	5.02	3.96	3.59	4.64	20.9	21.2	25.0	20.1	21.7	21.3
	3	6.34	5.38	3.59	4.55	21.8	22.4	27.4	20.7	22.8	22.7
S2	4	3.68	1.04	0.59	3.23	15.6	16.2	16.8	15.2	17.6	16.8
	5	5.87	2.83	0.59	3.62	18.0	18.3	18.6	17.2	20.3	19.1
	6	6.34	3.09	0.58	3.84	18.4	19.0	19.0	17.8	20.9	19.7
S3	7	3.66	4.49	7.56	6.73	25.3	22.3	24.0	19.2	24.1	21.5
	8	3.66	5.26	8.53	6.93	27.4	24.4	26.7	20.8	25.9	23.2
	9	3.69	5.84	9.28	7.13	28.5	25.7	27.8	20.4	26.6	25.6
Parking Slots Configuration (*Conf2*)											
S1	10	3.65	2.42	3.52	4.75	20.8	19.5	21.7	18.5	21.3	18.9
	11	4.99	3.74	3.52	4.77	21.3	20.7	24.3	20.3	22.1	20.4
	12	6.32	5.07	3.52	4.77	22.1	21.5	26.3	21.4	23.2	21.7
S2	13	3.65	1.33	3.43	5.75	14.9	15.0	16.2	11.4	17.9	15.5
	14	5.00	2.90	3.45	5.55	17.6	17.9	18.5	17.7	20.4	18.2
	15	6.35	3.75	3.41	6.01	18.4	19.0	19.1	18.0	20.6	19.4
S3	16	3.67	4.40	7.50	6.77	25.0	22.2	23.7	18.9	23.7	21.0
	17	3.67	5.19	8.50	6.98	26.9	24.1	26.2	20.4	25.4	22.7
	18	3.66	5.80	9.28	7.14	28.0	25.3	27.9	21.3	26.5	24.1

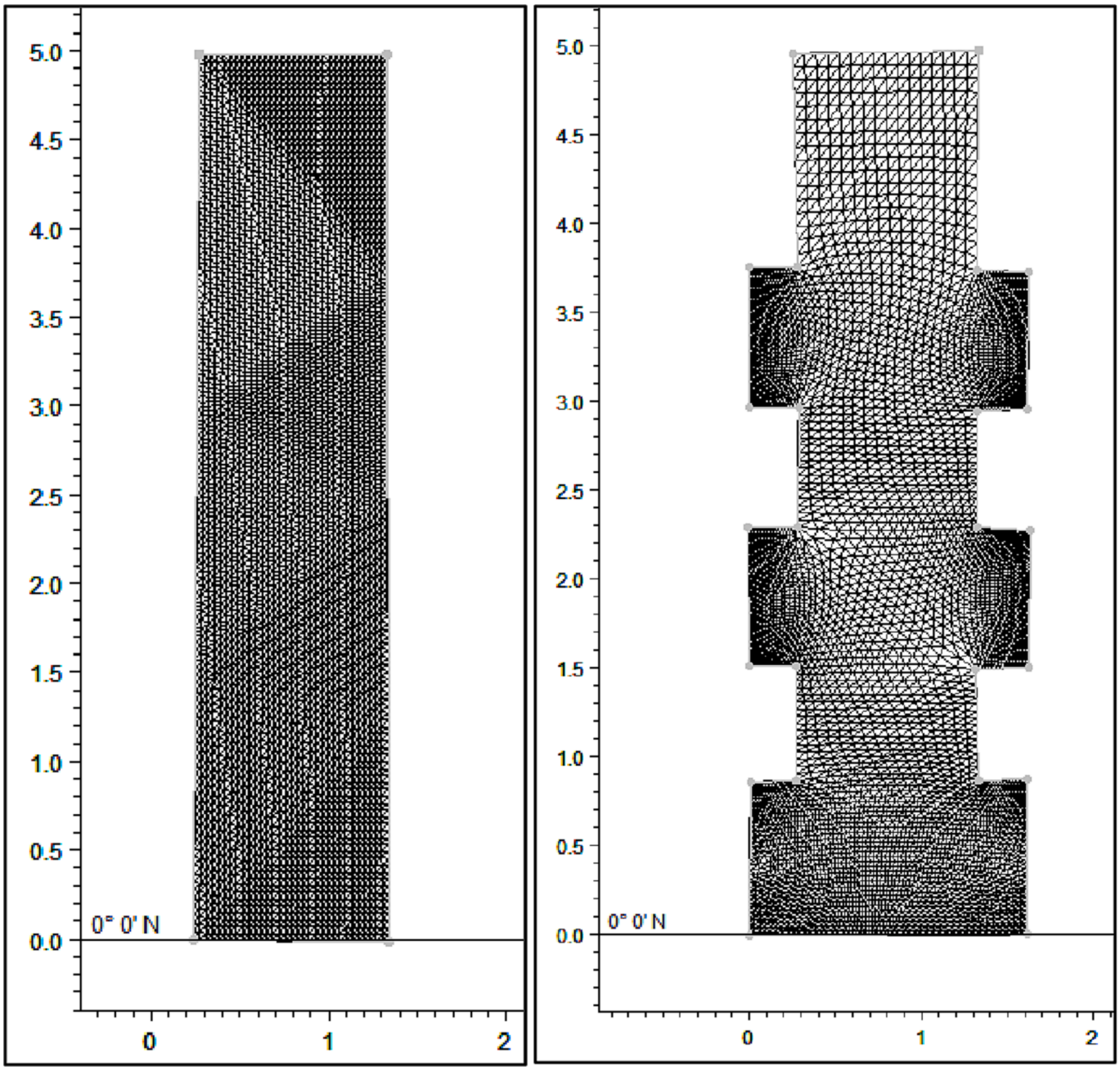


Fig. 4. The meshes used for each configuration. The left shows configuration 1 and the right shows configuration 2. Both have a geometry of 1 m × 5 m.

set a range of steady inflows. For all the tests conducted, flows were first created and left to stabilise before values were recorded. Once stabilised, data were collected for a period of 3 min to define reliable temporally averaged values (technique previously tested in Rubinato et al. (2017) and Martins et al. (2018)).

2.3. 2D velocity experimental measurements

Surface flow Velocity fields were obtained using surface Particle Image Velocimetry (sPIV). This is a common way to characterise 2D velocity fields of surface flow (Carmer et al., 2009). To achieve this during this study the water facility was equipped with a seeding particle dispenser. The seeding particles (polypropylene, 2–3 mm diameter, and have a density of 0.90 g/cm^3 (Weitbrecht et al., 2002)) were dispensed in uniform distribution over the flood plain via a roller brush attached to a vibrating particle hopper (varying between 0 and 20 rpm to control the release rate). This was recorded via three GoPro Hero 4 Black Edition cameras (set to record video frames of size 1440x1920 pixels). The cameras were hung independently to avoid vibrations. The cameras were fitted at a height of 1.5 m, to acquire the video frames. A 1 mm per pixel resolution at the centre of the images was obtained with a consequent maximum frame rate of 80 Hz (Martins et al., 2018). This established that each sPIV seeding particle was characterised by a cluster of at least 5 pixels. This was considered good particle definition and provided rigorous detection by the PIV software (Dynamic Studio by Dantec Dynamics Ltd). Go Pro cameras can have issues with lens distortion; thus, it was removed after spatial calibration by utilising chequerboard images. For full description please refer to Rojas Arques et al. (2018). The pixels outside the measurement area were then cropped for each image. Spatial calibration was carried out for all three Go Pro cameras, for the range of flow depths studied in this paper. Prior to each test the mean "background" (i.e., with no seeding particles) image was recorded over 5 mins

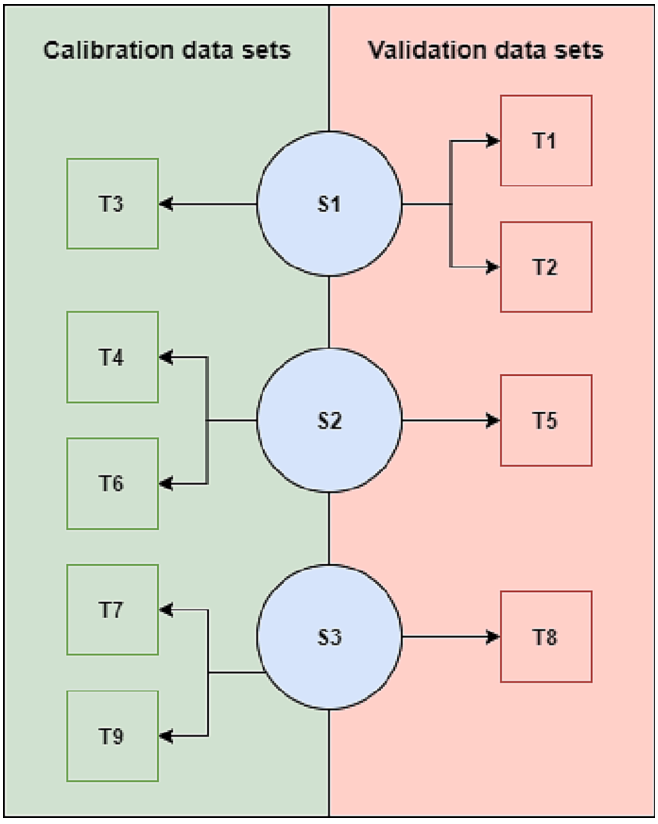


Fig. 5. The three scenarios (S1, S2 & S3) and hydraulic tests (T1-T9) that were split for calibration and validation purposes with configuration 1.

Table 2
Average computational times (in seconds) for both configurations and surfaces (G32 and G16) for each of the three scenarios.

Scenario Test ID	Configuration 1 computational times (seconds)		Configuration 2 Computational times (seconds)	
	G32	G16	G32	G16
S1	39	78	41	82
S2	40	79	42	83
S3	45	113	46	114

and then sPIV instantaneous images were recorded for a period of 3 min for each test. These images were subtracted from this background, such that the background would turn black while the particles would remain white. After having obtained the images with the instantaneous location of the seeding particles, these images were analysed using the commercial PIV software Dynamic Studio. Initially an adaptive correlation was performed to determine the velocity field for each time adjacent image pair and then a range of validation was applied to remove unrealistic high velocities and zero velocities resulting from interrogation areas with no seeding particles. For each flow condition tested within the facility, the filter removed<5% of the velocity vectors and those removed were then replaced via a 3 × 3 moving time average routine.

2.4. Numerical model

MIKE URBAN+ (DHI, 2020), was used in this research, as it is relatively simple to use (which is further supported with its new interface), with widely accessible academic licenses. Also, it is a very popular software so this research may benefit its many users. MIKE URBAN + can simulate both 2D overland and pipe flows by coupling 1D (MOUSE) and 2D (MIKE 21) processes together. The 1D hydrodynamic pipe flow model uses Saint-Venant 1D equations to solve pipe hydraulics. Whereas Depth Average Shallow Water equations are used to solve depths and velocities over the 2D surface. In this study the exact measurements of the physical model were replicated in MIKE URBAN + for the numerical simulations.

2.5. Domain discretisation

The domain for each surface was created using the MIKE ZERO toolbox. However, a smaller domain was simulated as nothing significantly changed past this area. So, the new domain geometry was 1 m × 5 m for the main flood plain channel. Two resolutions were used for each configuration. A lower resolution of 32x32 (G32) containing 25,488 cells, and a higher resolution of 16x16 (G16) containing 103,660 cells. Like the water facility, both surfaces were assigned a Manning's roughness coefficient of 0.009 ($sm^{-1/3}$), and a slope of 0.001. A square mesh was created for each configuration (Fig. 4 shows a mesh example). A broad crested weir was shaped along the downstream extent of the surface. The weir had a height of 0.05 m and width of 0.1 m. A weir coefficient of 0.8 and a weir exponential coefficient of 0.5 was set.

2.6. Boundary and initial conditions

As default in MIKE URBAN+, all grid boundaries are closed. An open boundary was therefore assigned to the downstream end of the grid file, so flow could leave the domain. The Q_1 boundary inflows were created by adding 14 single point sources along the highest edge on the 2D surface. The physical models Q_1 inflows was divided by 14 at each point for the simulation. Mean flows from the water laboratory test results were used as inflow boundary conditions at Q_1. Similarly, the Q_3 pipe inflow was assigned to the node in the piped system. Water laboratory mean flows were used as Q_3 inflow boundary conditions.

2.7. Model calibration and validation

For configuration 1, simulations were split between calibration and validation. Fig. 5 illustrates which test was used for each scenario. Calibration test numbers were chosen at random (in RStudio using random sample vector selection). It was considered that this would be sufficient and non-bias way to represent how the model performed under steady flow conditions. Computational times varied. As seen below in Table 2. All calibrations took place using the lower resolution surface (G32) of configuration 1.

All inflow conditions were simulated as continuous flows. For the first set of experiments, using configuration 1, the simulations included all nine observed water laboratory tests, from each scenario (S1, S2, S3). Pipe calibrations (adjusting pipe roughness values) and manhole discharge coefficients (adjusting coefficients in the manhole exchange) and changing manning's roughness values were done manually. During the calibration that replicated S1 (lid on manhole) a weir discharge coefficient of 0.2 was assigned to the manhole. This allowed some surface flow to bypass the lid and enter the manhole, as seen in the laboratory model. When calibrating S2 and S3 a discharge coefficient of 0.9 was applied to the manhole. Similarly, this allowed comparable flows as the laboratory model to enter the manhole. The pipes, Q_3 and Q_4 were both assigned a Hazen-Williams roughness value of 140. To simulate velocity fields, both grid surfaces were separately tested with the initial boundary conditions. Configuration 2 was used to further validate the modelling procedure using tests T10-T18; therefore, all parameters were kept the same.

2.8. Data analysis

Modelled manhole flow exchange (Q_e), sewer pipe outflow (Q_4) surface outflow (Q_2) and the surface depths (at P_0, P_1, P_2, P_3, P_4 and P_5)

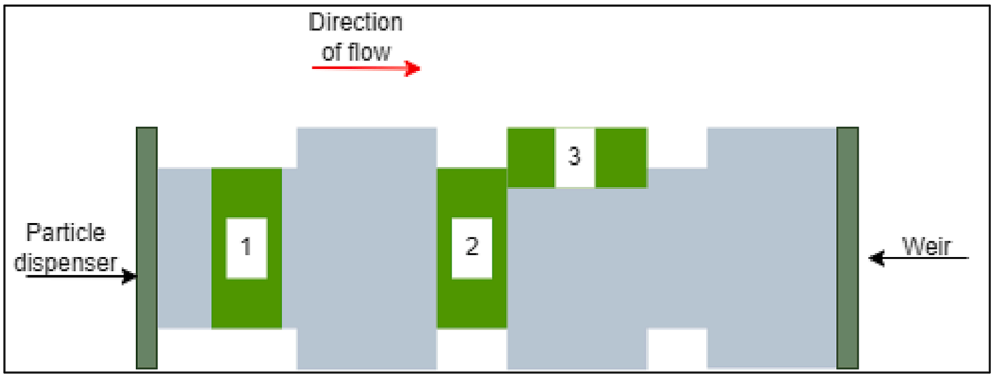


Fig. 6. Locations of vector field analysis. The green band (1) shows where the vector fields were taken around the manhole, the green band (2) shows where the vector fields were taken 2 m below the manhole, and the green band (3) shows the investigated parking space. The red arrow identifies the direction of flow. Both the weir at the downstream end and particle dispenser are shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Combined scenario flow exchange (Q_e), pipe outflow (Q_4) and overland flows (Q_2) from configuration 1. Results of the calibration and validation are shown, quantified by multiple R^2, normalised root mean square error (NRMSE) and statistical significance.

	Flow ID	Multiple R^2	NRMSE	*P*-value
Calibration	Q_e	0.989	0.072	<0.001
	Q_4	0.954	0.028	<0.001
	Q_2	0.944	0.036	<0.001
Validation	Q_e	0.992	0.079	<0.001
	Q_4	0.978	0.027	<0.001
	Q_2	0.968	0.037	<0.001

Table 4
Calibration* and validation results for observed (obs) and simulated (sim) flow exchange (Q_e), outflow at sewer pipe (Q_4) and outflow at flood plain surface (Q_2) from configuration 1.

S	Test ID	$Q_{e\ (obs)}$ l/s)	$Q_{e\ (sim)}$ (1/s)	$Q_{4\ (obs)}$ (1/s)	$Q_{4\ (sim)}$ (l/s)	$Q_{2\ (obs)}$ (l/s)	$Q_{2\ (sim)}$ (l/s)
S1	T1	−1.01	−1.13	4.60	4.72	2.67	2.55
	T2	−1.05	−1.19	4.64	4.78	3.96	3.83
	T3*	−0.96	−1.11	4.55	4.70	5.38	5.23
S2	T4*	−2.64	−2.77	3.23	3.36	1.04	0.91
	T5	−3.03	−3.17	3.62	3.76	2.83	2.70
	T6*	−3.30	−3.41	3.84	3.99	3.09	2.93
S3	T7*	0.80	0.96	6.73	6.60	4.49	4.62
	T8	1.60	1.73	6.93	6.80	5.26	5.42
	T9*	2.15	2.30	7.13	6.98	5.84	5.99

Table 5
Combined scenario validation results shows flow exchange (Q_e), pipe outflow (Q_4) and overland flows (Q_2) from configuration 2. Results are quantified by multiple R^2, normalised root mean square error (NRMSE) and statistical significance.

Flow ID	Multiple R^2	NRMSE	*P*-value
Q_e	0.997	0.082	<0.001
Q_4	0.994	0.024	<0.001
Q_2	0.995	0.036	<0.001

were directly taken from the model for comparison with experimental data. The velocity fields around the manhole and two meters within the manhole were investigated for configuration 1. Fig. 6 shows the vector field analysis locations including the parking space for configuration 2. Velocity field data was abstracted from the modelling output in a grid around each observation area, (location 1, 2 and 3) and compared with

Table 6
Observed (obs) and simulated (sim) flow exchange (Q_e), outflow at sewer pipe (Q_4) and outflow at flood plain surface (Q_2) for configuration 2.

S	Test ID	$Q_{e\ (obs)}$ (1/s)	$Q_{e\ (sim)}$ (1/s)	$Q_{4\ (obs)}$ (1/s)	$Q_{4\ (sim)}$ (1/S)	$Q_{2\ (obs)}$ (1/s)	$Q_{2\ (sim)}$ (1/s)
S1	T10	−1.23	−1.36	4.75	4.88	2.42	2.29
	T11	−1.25	−1.39	4.77	4.91	3.74	3.60
	T12	−1.25	−1.40	4.77	4.92	5.07	4.92
S2	T13	−2.32	−2.45	5.75	5.88	1.33	1.20
	T14	−2.10	−2.24	5.55	5.69	2.90	2.79
	T15	−2.60	−2.75	6.01	6.16	3.75	3.60
S3	T16	0.73	0.86	6.77	6.64	4.40	4.53
	T17	1.52	1.65	6.98	6.85	5.19	5.32
	T18	2.14	2.29	7.14	6.99	5.80	5.95

Table 7
Combined scenario surface depths (P_0, P_1, P_2, P_3, P_4 and P_5). Results of the calibration and validation are shown for configuration 1, quantified by multiple R^2, normalised root mean square error (NRMSE) and statistical significance.

	Depth ID	Multiple R^2	NRMSE	*P*-value
Calibration	P_0	0.993	0.053	<0.001
	P_1	0.957	0.058	0.004
	P_2	0.949	0.059	0.005
	P_3	0.767	0.067	0.052
	P_4	0.973	0.054	0.002
	P_5	0.927	0.059	0.009
Validation	P_0	0.993	0.046	0.001
	P_1	0.984	0.050	0.008
	P_2	0.965	0.048	0.018
	P_3	0.900	0.056	0.051
	P_4	0.993	0.049	0.001
	P_5	0.965	0.051	0.017

the experimental measured velocities. The results were interpolated within the mesh to that same grid then plotted each pair of points. In vector fields analysis location 1 velocities from the manhole centre were recorded at 125 mm, 250 mm, 375 mm and 500 mm. In vector fields analysis location 2 velocities from the centre of the flood plain were recorded at 125 mm, 250 mm, 375 mm and 500 mm.

A linear regression model was conducted in RStudio, which quantified how close the fit was regarding calibration and validation results with the physical testing. Multiple R^2 (the absolute fraction of variance), normalised root mean square errors (NRMSE) and *P*-values were chosen for quantifying the robustness of the numerical model. The normalised root mean square error (or sometimes known as relative root mean square error) is a quantitative statistical indicator that normalises the

Table 8
Calibration* and validation results for observed (obs) and simulated (sim) surface depths (P_0, P_1, P_2, P_3, P_4 and P_5) from configuration 1.

	Test ID	$P_{0\ (obs)}$ (mm)	$P_{0\ (sim)}$ (mm)	$P_{1\ (obs)}$ (mm)	$P_{1\ (sim)}$ (mm)	$P_{2\ (obs)}$ (mm)	$P_{2\ (sim)}$ (mm)	$P_{3\ (obs)}$ (mm)	$P_{3\ (sim)}$ (mm)	$P_{4\ (obs)}$ (mm)	$P_{4\ (sim)}$ (mm)	$P_{5\ (obs)}$ (mm)	$P_{5\ (sim)}$ (mm)
S1	T1	20.63	19.82	19.88	19.05	22.65	21.78	18.78	17.95	21.62	20.80	19.87	19.04
	T2	20.93	20.01	21.24	20.28	25.02	24.04	20.11	19.15	21.65	20.73	21.29	20.33
	T3*	21.81	20.77	22.37	21.30	27.42	26.34	20.69	19.66	22.82	21.77	22.74	21.71
S2	T4*	17.57	16.56	18.20	17.18	18.80	17.71	17.24	16.22	19.59	18.49	18.83	17.81
	T5	18.03	16.91	18.30	17.17	18.64	17.45	17.21	16.09	20.34	19.22	19.11	17.99
	T6*	18.42	17.27	18.95	17.30	18.97	17.21	17.84	16.33	20.87	19.72	19.74	18.22
S3	T7*	25.32	26.44	22.34	23.54	23.98	25.30	19.24	20.44	24.06	25.28	21.48	22.69
	T8	27.36	28.49	24.38	25.61	26.67	28.02	20.84	22.15	25.88	27.30	23.25	24.54
	T9*	28.53	30.04	25.74	26.99	27.81	29.34	20.39	21.90	26.59	28.14	25.58	27.09

Table 9
Combined scenario surface depths (P_0, P_1, P_2, P_3, P_4 and P_5) for configuration 2. Quantified by multiple R^2, normalised root mean square error (NRMSE) and statistical significance.

Depth ID	Multiple R^2	NRMSE	*P*-value
P0	0.990	0.052	<0.001
P1	0.966	0.058	<0.001
P2	0.961	0.057	<0.001
P3	0.906	0.064	<0.001
P4	0.967	0.053	<0.001
P5	0.966	0.058	<0.001

root mean square error (RMSE). In this case it was normalised due to different flow inputs being used in each test at Q_1 and Q_3 (Table 1). Although there are different ways to normalise RMSE, this study divided RMSE by the absolute mean of measured data. As described by Despotovic et al. (2016), NRMSE < 10% is considered excellent, good if 10% < NRMSE < 20%, fair if 20% < NRMSE < 30%, and poor if NRMSE > 30%.

3. Results

Accuracy is determined by comparing the outputs from each and determining the amount of error. The results include errors from combined scenarios. And observed and simulated results of manhole, pipe and surface flows, surface depths and velocity fields. The numerical model fit can be quantified by the multiple R^2, normalised root mean square error (NRMSE) and statistical significance (*P*-values). Configuration 1 was split for calibration (tests T3, T4, T6, T7 and T9) and validation (T1, T2, T5 and T8). All of configuration 2 results were used to validate the models performance.

3.1. Manhole exchange, pipe and surface flows results of surface configuration 1

3.1.1. Calibration

The results are given in Table 3 and in Appendix A.1. Ranges of R^2 were between 0.944 and 0.989. Ranges of NRMSE were between 0.028 and 0.072. The highest error was recorded at Q_e. As seen in Table 4 below the calibrated tests T3, T4 and T6 overestimated flows entering the manhole (Q_e) by 0.148 (1/s), 0.125 (1/s) and 0.150 (1/s) respectively. During surcharged conditions T7 had an additional flow of 0.125 (1/s) leaving the manhole, and T9 had an increase of 0.148 (1/s).

3.1.2. Validation

The model's validation results are reported in Table 3 and Appendix A.2. Ranges of R^2 were between 0.968 and 0.992. Ranges of NRMSE were between 0.027 and 0.079. Similarly, Q_e had the highest error. Observed and simulated flows are shown in Table 4. For T1, T2 and T5 flows entering Q_e were over calculated by 0.124 (1/s), 0.136 (1/s) and 0.138 (1/s) respectively. However, for T8 there was an over estimated surcharge of 0.132 (1/s). The statistical significance implied a robust model for each split data sets (<0.05). The model accuracy illustrates that the mean flows from the numerical model was closely related to the measured water facility flows. The results of flow exchange and outflows from configuration 1 are considered excellent (NRMSE < 10%).

3.2. Manhole exchange, pipe and surface flows results of surface configuration 2

3.2.1. Validation

Results for the validated flows are shown in Table 5 and Appendix A.3. Ranges of R^2 were between 0.994 and 0.997, and ranges of NRMSE were between 0.024 and 0.082. Corresponding to configuration 1, flow error for configuration 2 was greatest at Q_e. All tests had a significant outcome (<0.05). The results of flow exchange and outflows from

Table 10
Validation results for observed (obs) and simulated (sim) surface depths (P_0, P_1, P_2, P_3, P_4 and P_5) from configuration 2.

S	Test ID	$P_{0\ (obs)}$ (mm)	$P_{0\ (sim)}$ (mm)	$P_{1\ (obs)}$ (mm)	$P_{1\ (sim)}$ (mm)	$P_{2\ (obs)}$ (mm)	$P_{2\ (sim)}$ (mm)	$P_{3\ (obs)}$ (mm)	$P_{3\ (sim)}$ (mm)	$P_{4\ (obs)}$ (mm)	$P_{4\ (sim)}$ (mm)	$P_{5\ (obs)}$ (mm)	$P_{5\ (sim)}$ (mm)
S1	T10	20.80	19.97	19.50	18.65	21.70	20.82	18.50	17.66	21.30	20.47	18.90	18.06
	T11	21.30	20.37	20.70	19.73	24.30	23.31	20.30	19.33	22.10	21.17	20.40	19.43
	T12	22.10	21.05	21.50	20.41	26.30	25.20	21.40	20.35	23.20	22.13	21.70	20.65
S2	T13	14.90	13.88	15.00	13.97	16.20	15.10	11.40	10.37	17.90	16.79	15.50	14.46
	T14	17.60	16.46	17.90	16.76	18.50	17.30	17.70	16.57	20.40	19.27	18.20	17.06
	T15	18.40	17.24	19.00	17.34	19.10	17.33	18.00	16.48	20.60	19.44	19.40	17.87
S3	T16	25.00	26.13	22.20	23.42	23.70	25.03	18.90	20.12	23.70	24.93	21.00	22.22
	T17	26.90	28.04	24.10	25.34	26.20	27.56	20.40	21.72	25.40	26.84	27.70	29.02
	T18	28.00	29.52	25.30	26.56	27.90	29.44	21.30	22.82	26.50	28.06	24.10	25.62

Table 11
Calibration* and validation results of velocities for configuration 1 (analysis area 1). Illustrates the vector type (longitudinal and transversal) for each grid used (G32 and G16) and test (Tl- T9). Multiple R^2, normalised root mean square error (NRMSE) and *P*-values are shown for total model outcome. NRMSE is also shown at 4 distances from the manhole centre.

Test ID	Total model NRMSE	Multiple R^2	*P*-value	NRMSE - Distance from manhole			
				125 mm	250 mm	375 mm	500 mm
G32 Longitudinal							
T1	0.058	0.995	<0.001	0.066	0.054	0.045	0.025
T2	0.075	0.990	<0.001	0.094	0.075	0.060	0.042
T3*	0.098	0.983	<0.001	0.134	0.091	0.072	0.056
T4*	0.058	0.997	<0.001	0.089	0.059	0.047	0.028
T5	0.082	0.993	<0.001	0.122	0.084	0.066	0.046
T6*	0.101	0.991	<0.001	0.136	0.098	0.088	0.063
T7*	0.117	0.981	<0.001	0.140	0.112	0.102	0.090
T8	0.125	0.972	<0.001	0.141	0.120	0.110	0.097
T9*	0.130	0.958	<0.001	0.149	0.131	0.117	0.101
G32 Transversal							
T1	0.069	0.998	<0.001	0.073	0.068	0.053	0.032
T2	0.089	0.997	<0.001	0.110	0.083	0.076	0.052
T3*	0.126	0.994	<0.001	0.143	0.124	0.123	0.109
T4*	0.081	0.998	<0.001	0.101	0.073	0.061	0.036
T5	0.119	0.994	<0.001	0.138	0.099	0.083	0.073
T6*	0.155	0.990	<0.001	0.159	0.140	0.126	0.122
T7*	0.158	0.989	<0.001	0.175	0.152	0.131	0.115
T8	0.166	0.986	<0.001	0.219	0.155	0.141	0.124
T9*	0.173	0.985	<0.001	0.226	0.163	0.154	0.127
G16 Longitudinal							
T1	0.068	0.991	<0.001	0.079	0.064	0.051	0.028
T2	0.082	0.987	<0.001	0.102	0.083	0.069	0.051
T3*	0.113	0.971	<0.001	0.142	0.107	0.082	0.067
T4*	0.071	0.994	<0.001	0.093	0.072	0.061	0.033
T5	0.091	0.987	<0.001	0.130	0.094	0.073	0.060
T6*	0.113	0.984	<0.001	0.144	0.110	0.096	0.074
T7*	0.132	0.972	<0.001	0.154	0.125	0.110	0.097
T8	0.140	0.958	<0.001	0.164	0.133	0.122	0.102
T9*	0.141	0.945	<0.001	0.167	0.140	0.125	0.109
G16 Transversal							
T1	0.076	0.996	<0.001	0.086	0.077	0.062	0.039
T2	0.107	0.995	<0.001	0.133	0.104	0.085	0.073
T3*	0.139	0.992	<0.001	0.169	0.134	0.133	0.119
T4*	0.101	0.996	<0.001	0.117	0.086	0.071	0.043
T5	0.137	0.988	<0.001	0.148	0.113	0.097	0.083
T6*	0.163	0.988	<0.001	0.174	0.147	0.138	0.125
T7*	0.172	0.985	<0.001	0.184	0.163	0.144	0.132
T8	0.189	0.986	<0.001	0.238	0.178	0.157	0.146
T9*	0.192	0.979	<0.001	0.243	0.188	0.168	0.150

Table 12
Validation results of velocities around the manhole for configuration 2 (analysis area 1). Illustrates the vector type (longitudinal and transversal) for each grid used (G32 and G16) and test (T10-T18). Multiple R^2, normalised root mean square error (NRMSE) and *P*-values are shown for total model outcome. NRMSE is also shown at 4 distances from the manhole centre.

Test ID	Total model NRMSE	Multiple R^2	*P*-value	NRMSE - Distance from manhole			
				125 mm	250 mm	375 mm	500 mm
G32 Longitudinal							
T10	0.060	0.994	<0.001	0.068	0.056	0.049	0.027
T11	0.087	0.985	<0.001	0.100	0.079	0.063	0.048
T12	0.107	0.974	<0.001	0.139	0.096	0.078	0.059
T13	0.063	0.997	<0.001	0.094	0.065	0.051	0.032
T14	0.088	0.992	<0.001	0.126	0.087	0.068	0.049
T15	0.108	0.990	<0.001	0.143	0.110	0.095	0.081
T16	0.135	0.964	<0.001	0.151	0.128	0.111	0.100
T17	0.139	0.953	<0.001	0.158	0.136	0.117	0.107
T18	0.146	0.923	<0.001	0.163	0.142	0.128	0.116
G32 Transversal							
T10	0.080	0.996	<0.001	0.085	0.076	0.062	0.040
T11	0.106	0.992	<0.001	0.122	0.098	0.087	0.063
T12	0.138	0.989	<0.001	0.150	0.146	0.136	0.120
T13	0.093	0.996	<0.001	0.110	0.081	0.070	0.043
T14	0.137	0.992	<0.001	0.146	0.111	0.099	0.086
T15	0.156	0.988	<0.001	0.170	0.156	0.146	0.142
T16	0.173	0.985	<0.001	0.182	0.160	0.151	0.148
T17	0.180	0.982	<0.001	0.232	0.168	0.155	0.152
T18	0.183	0.980	<0.001	0.253	0.173	0.164	0.131
G16 Longitudinal							
T10	0.069	0.991	<0.001	0.083	0.066	0.055	0.032
T11	0.091	0.991	<0.001	0.110	0.087	0.071	0.054
T12	0.113	0.971	<0.001	0.149	0.108	0.086	0.072
T13	0.072	0.996	<0.001	0.093	0.081	0.066	0.036
T14	0.095	0.991	<0.001	0.132	0.103	0.076	0.064
T15	0.114	0.986	<0.001	0.159	0.119	0.106	0.080
T16	0.139	0.962	<0.001	0.168	0.128	0.124	0.102
T17	0.142	0.950	<0.001	0.170	0.138	0.126	0.107
T18	0.147	0.926	<0.001	0.176	0.143	0.132	0.113
G16 Transversal							
T10	0.090	0.995	<0.001	0.094	0.086	0.079	0.040
T11	0.115	0.991	<0.001	0.136	0.111	0.089	0.077
T12	0.149	0.986	<0.001	0.176	0.149	0.142	0.129
T13	0.112	0.994	<0.001	0.129	0.093	0.084	0.058
T14	0.145	0.991	<0.001	0.154	0.117	0.110	0.089
T15	0.167	0.985	<0.001	0.196	0.174	0.146	0.135
T16	0.190	0.979	<0.001	0.191	0.183	0.152	0.143
T17	0.195	0.976	<0.001	0.244	0.187	0.162	0.154
T18	0.201	0.974	<0.001	0.267	0.192	0.174	0.162

configuration 2 are considered excellent (NRMSE < 10%). This signified that the model accuracy was good and closely related to the measured water facility flows.

Table 6 shows observed and measured flow results from configuration 2. Flows were slightly over calculated at Q_e for configuration 2; also, errors during all scenarios increased slightly when parking spaces were incorporated into the study. For T10, T11 and T12 flows entering the manhole at Q_e were over calculated by 0.125 (1/s), 0.137 (1/s) and 0.149 (1/s) respectively. For T13, T14 and T15 the hydrodynamic model over calculated flows entering the manhole by 0.126 (1/s), 0.139 (1/s), and 0.151 (1/s) respectively. Under surcharged conditions flow during T16, T17 and T18 were over calculated by 0.126 (1/s), 0.133 (1/s) and 0.149 (1/s). The increased error of flows in configuration 2 occurred due to an increase in complexity in flow patterns encompassing the manhole. Although the difference in error is small, it is highlighted by greater errors at Q_e for configuration 2 (NRMSE = 0.082) than at configuration 1 (NRMSE = 0.072).

3.3. Surface depth results for configuration 1

3.3.1. Calibration

The calibrated values for surface depths are shown in Table 7 and Appendix B.1. Ranges of R^2 were between 0.767 and 0.993. Ranges of NRMSE were between 0.053 and 0.067. Location P_0 had the lowest error and P_3 had the highest. The results show a good statistical significance. This indicates the numerical model managed to replicate the observed depths well. Table 8 illustrates observed and simulated depths during each hydraulic testing scenario.

When identifying depth location P_2, which is right in front of the manhole, the depths during T3, T4 and T6 were underestimated. They were within 1.08 mm, 1.09 mm, and 1.76 mm of the experimental data set respectively. Model performance decreased with increasing Q_1 inflow. Depths during T7 and T9 were overestimated at P_2 and were within 1.32 mm and 1.53 mm of the experimental data set respectively. Model performance decrease with increasing Q_3 inflow.

Table 13
Calibration* and validation results of velocities for configuration 1 (analysis area 2). Illustrates the vector type for each grid used (G32 and G16) and test number (Tl-T9). Multiple R^2, normalised root mean square error (NRMSE) and *P*-values are shown for total model outcome. NRMSE is also shown at 4 distances from the grid centre.

Test ID	Total model NRMSE	Multiple R^2	*P*-value	NRMSE - Distance from grid centre			
				125 mm	250 mm	375 mm	500 mm
G32 Longitudinal							
T1	0.048	0.983	<0.001	0.061	0.049	0.032	0.022
T2	0.061	0.933	<0.001	0.078	0.060	0.049	0.039
T3*	0.073	0.871	<0.001	0.098	0.074	0.062	0.050
T4*	0.057	0.991	<0.001	0.067	0.052	0.036	0.027
T5	0.070	0.979	<0.001	0.087	0.061	0.051	0.040
T6*	0.099	0.965	<0.001	0.113	0.083	0.072	0.060
T7*	0.105	0.816	<0.001	0.124	0.112	0.098	0.083
T8	0.109	0.792	<0.001	0.128	0.115	0.108	0.087
T9*	0.114	0.874	<0.001	0.139	0.124	0.111	0.091
G32 Transversal							
T1	0.054	0.998	<0.001	0.067	0.059	0.050	0.030
T2	0.083	0.993	<0.001	0.083	0.075	0.069	0.044
T3*	0.095	0.990	<0.001	0.106	0.096	0.081	0.061
T4*	0.075	0.990	<0.001	0.072	0.063	0.057	0.034
T5	0.094	0.990	<0.001	0.092	0.089	0.073	0.047
T6*	0.114	0.976	<0.001	0.115	0.102	0.088	0.066
T7*	0.117	0.991	<0.001	0.121	0.111	0.096	0.076
T8	0.123	0.976	<0.001	0.127	0.118	0.106	0.080
T9*	0.125	0.989	<0.001	0.138	0.131	0.114	0.092
G16 Longitudinal							
T1	0.052	0.972	<0.001	0.065	0.053	0.035	0.026
T2	0.067	0.910	<0.001	0.081	0.070	0.056	0.047
T3*	0.085	0.815	<0.001	0.101	0.090	0.082	0.059
T4*	0.064	0.985	<0.001	0.076	0.057	0.040	0.033
T5	0.084	0.957	<0.001	0.096	0.080	0.065	0.055
T6*	0.111	0.942	<0.001	0.127	0.090	0.085	0.068
T7*	0.119	0.745	<0.001	0.138	0.124	0.114	0.091
T8	0.123	0.727	<0.001	0.147	0.131	0.120	0.098
T9*	0.133	0.808	<0.001	0.156	0.146	0.132	0.106
G16 Transversal							
T1	0.069	0.998	<0.001	0.086	0.075	0.066	0.038
T2	0.101	0.990	<0.001	0.108	0.094	0.082	0.060
T3*	0.122	0.987	<0.001	0.133	0.128	0.110	0.085
T4*	0.099	0.981	<0.001	0.094	0.083	0.070	0.045
T5	0.113	0.984	<0.001	0.113	0.107	0.095	0.068
T6*	0.149	0.954	<0.001	0.142	0.137	0.125	0.090
T7*	0.152	0.987	<0.001	0.170	0.152	0.144	0.104
T8	0.174	0.970	<0.001	0.184	0.168	0.157	0.119
T9*	0.183	0.981	<0.001	0.207	0.191	0.162	0.134

Table 14
Validation results of velocities for configuration 2 (analysis area 2). Illustrates the vector type (longitudinal and transversal) for each grid used (G32 and G16) and test (T10-T18). Multiple R^2, normalised root mean square error (NRMSE) and *P*-values are shown for total model outcome. NRMSE is also shown at 4 distances from the grid centre.

Test ID	Total model NRMSE	Multiple R^2	*P*-value	NRMSE – Distance from grid centre			
				125 mm	250 mm	375 mm	500 mm
G32 Longitudinal							
T10	0.052	0.975	<0.001	0.064	0.051	0.036	0.026
T11	0.067	0.953	<0.001	0.090	0.064	0.055	0.044
T12	0.088	0.909	<0.001	0.108	0.086	0.075	0.056
T13	0.065	0.992	<0.001	0.089	0.061	0.048	0.030
T14	0.087	0.978	<0.001	0.101	0.065	0.056	0.046
T15	0.100	0.962	<0.001	0.116	0.096	0.080	0.073
T16	0.112	0.853	<0.001	0.128	0.115	0.105	0.092
T17	0.119	0.892	<0.001	0.140	0.132	0.110	0.094
T18	0.125	0.837	<0.001	0.160	0.137	0.114	0.095
G32 Transversal							
T10	0.067	0.994	<0.001	0.077	0.065	0.060	0.039
T11	0.085	0.994	<0.001	0.161	0.086	0.078	0.062
T12	0.106	0.992	<0.001	0.113	0.104	0.103	0.101
T13	0.081	0.996	<0.001	0.105	0.074	0.065	0.039
T14	0.105	0.937	<0.001	0.106	0.096	0.084	0.067
T15	0.115	0.994	<0.001	0.132	0.113	0.110	0.103
T16	0.129	0.984	<0.001	0.138	0.121	0.110	0.106
T17	0.138	0.980	<0.001	0.134	0.130	0.123	0.111
T18	0.145	0.983	<0.001	0.149	0.135	0.134	0.125
G16 Longitudinal							
T10	0.060	0.957	<0.001	0.072	0.061	0.039	0.030
T11	0.085	0.903	<0.001	0.106	0.081	0.061	0.049
T12	0.093	0.866	<0.001	0.106	0.090	0.085	0.066
T13	0.065	0.991	<0.001	0.090	0.069	0.050	0.034
T14	0.091	0.964	<0.001	0.097	0.092	0.069	0.056
T15	0.112	0.939	<0.001	0.134	0.107	0.099	0.072
T16	0.126	0.825	<0.001	0.140	0.134	0.120	0.100
T17	0.131	0.808	<0.001	0.156	0.135	0.125	0.104
T18	0.137	0.792	<0.001	0.168	0.140	0.134	0.111
G16 Transversal							
T10	0.081	0.991	<0.001	0.091	0.085	0.072	0.040
T11	0.103	0.990	<0.001	0.115	0.099	0.092	0.074
T12	0.131	0.984	<0.001	0.155	0.141	0.124	0.092
T13	0.085	0.996	<0.001	0.099	0.090	0.083	0.048
T14	0.125	0.957	<0.001	0.124	0.113	0.101	0.082
T15	0.145	0.989	<0.001	0.175	0.148	0.137	0.118
T16	0.178	0.960	<0.001	0.191	0.165	0.157	0.130
T17	0.194	0.970	<0.001	0.194	0.187	0.171	0.145
T18	0.227	0.960	<0.001	0.254	0.192	0.172	0.151

3.3.2. Validation

The validated model results are shown in Table 7 and Appendix B.2 and include the tests for T1, T2, T5 and T8. Ranges of R^2 were between 0.900 and 0.993. Similarly, P_0 had the lowest error. Statistical significance also explains a good model validation of depths (<0.05). Ranges of NRMSE were between 0.046 and 0.056. When identifying depth location P_2, depths at T1, T2, T5 were underestimated and stood within 0.87 mm, 0.98 mm and 1.19 mm of the experimental data set respectively. Model performance decreased with increasing Q_1 inflow for these three tests. However, at location P_2, T8 over estimated depths and was within 1.35 mm of the experimental data set. During this test model performance decreased with increasing Q_3 inflow. Table 8 illustrates observed and simulated depths during each hydraulic testing scenario.

3.4. Surface depth results for configuration 2

3.4.1. Validation

The accuracy of the validated model is illustrated in Table 9 below and Appendix B.3. Ranges of R^2 were between 0.906 and 0.990. Ranges of NRMSE were between 0.052 and 0.064. All tests were significant (<0.05). Like configuration 1, the error of depths in configuration 2 were highest at P_3 and lowest at P_0. This highlights that the wave reflections of the flow had the most disparity at these two points.

Table 10 illustrates observed and simulated depths during each hydraulic scenario. During S1 and S2 depths were slightly under calculated and S3 depths were over calculated. Differences in depths were < 2 mm, which was comparable to configuration 1; though the differences were slightly larger when parking spaces were added to the flood plain. Similarly, errors in depths increased in relation to increasing Q_1 inflow for S1 and S2. For S3 errors in depths were related to Q_3 inflow conditions. The error in depths during S1 were the smallest in all hydraulic scenarios, as seen in the previous street profile. When identifying location P_2, depths during T10, T11 and T12 were within 0.88 mm, 0.99 mm and 1.10 mm, respectively. S2 also saw underestimated depths. At P_2 depths during T13, T14 and T15 were within 1.10 mm, 1.20 mm and 1.77 mm, correspondingly. Depths during S3, where pipe inflow dominated flows,

Table 15
Validation results of velocities for configuration 2 (analysis area 3). Illustrates the vector type (longitudinal and transversal) for each grid used (G32 and G16) and test (T10-T18). Multiple R^2, normalised root mean square error (NRMSE) and *P*-values are shown for total model outcome.

Test ID	Total model NRMSE	Multiple R^2	*P*-value
G32 Longitudinal			
T10	0.090	0.998	<0.001
T11	0.119	0.996	<0.001
T12	0.132	0.994	<0.001
T13	0.099	0.994	<0.001
T14	0.125	0.995	<0.001
T15	0.144	0.993	<0.001
T16	0.157	0.994	<0.001
T17	0.167	0.993	<0.001
T18	0.177	0.991	<0.001
G32 Transversal			
T10	0.111	0.997	<0.001
T11	0.133	0.994	<0.001
T12	0.173	0.990	<0.001
T13	0.117	0.994	<0.001
T14	0.172	0.991	<0.001
T15	0.186	0.988	<0.001
T16	0.205	0.981	<0.001
T17	0.210	0.987	<0.001
T18	0.215	0.986	<0.001
G16 Longitudinal			
T10	0.098	0.995	<0.001
T11	0.126	0.993	<0.001
T12	0.138	0.991	<0.001
T13	0.104	0.998	<0.001
T14	0.133	0.990	<0.001
T15	0.149	0.988	<0.001
T16	0.164	0.988	<0.001
T17	0.173	0.987	<0.001
T18	0.182	0.984	<0.001
G16 Transversal			
T10	0.115	0.996	<0.001
T11	0.137	0.993	<0.001
T12	0.181	0.986	<0.001
T13	0.121	0.993	<0.001
T14	0.176	0.988	<0.001
T15	0.190	0.983	<0.001
T16	0.211	0.980	<0.001
T17	0.219	0.984	<0.001
T18	0.225	0.979	<0.001

increased with increasing inflow. At P_2 depths were within 1.33 mm during T16, 1.36 mm during T17 and 1.54 mm during T18.

3.5. Results of velocity fields for surface configuration 1 (analysis area 1)

3.5.1. Calibration

The agreement between the observed and simulated velocities from analysis area 1 are shown in Table 11. The results of the velocity fields in analysis area 1 are plotted in Appendix C.1, C.2, C.3 and C.4. The distances from the manhole centre are shown in each. The results show that the lower resolution grid (G32) had less error than the higher resolution grid (G16). Longitudinal flow had smaller error than transversal flows. Velocity error also increased the nearer to the manhole.

For G32, calibrated NRMSE ranged between 0.058 and 0.130 for longitudinal flow. Calibrated transversal flows had slightly more error. NRMSE ranged between 0.081 and 0.173. When the calibrated model was run using G16, NRMSE ranged between 0.071 and 0.141 for longitudinal flow. Transversal flows had an NRMSE range of 0.101 and 0.192. R^2 values were > 0.945, with a significant outcome in all calibrated tests (*P*-value < 0.05).

3.5.2. Validation

The validated model with the G32 surface had a ranging NRMSE of between 0.058 and 0.125 for longitudinal velocities and between 0.069 and 0.166 for transversal velocities. The validated model with the G16 surface had a ranging NRMSE of 0.068 and 0.140 for longitudinal flows and a range of 0.076 and 0.189 for transversal flows. In all the tests R^2 values were > 0.958 and tests were significant (*P*-value < 0.05).

3.6. Results of velocity fields for surface configuration 2 (analysis area 1)

3.6.1. Validation

The accuracy of the model is shown for both G32 and G16 in longitudinal and transversal vectors in Table 12. These results are plotted in Appendix D.1, D.2, D.3 and D.4. For G32 longitudinal flows NRMSE ranged from between 0.06 and 0.146. Transversal flows had slightly higher error. They ranged from between 0.08 and 0.183. Similarly, all tests were significant with $R^2 > 0.923$. G16 had larger errors; NRMSE for longitudinal flows ranged between 0.069 and 0.147 and transversal flow error ranged between 0.090 and 0.201.

3.7. Results of velocity fields for configuration 1 (analysis area 2)

The agreement between the observed and simulated velocities from analysis area 2 are shown in Table 13 and in Appendix C.5, C.6, C.7 and C.8. The distances from the manhole centre are shown in each. The results show that the lower resolution grid (G32) had less error than the higher resolution grid (G16). Longitudinal flow had smaller error than transversal flows. Velocity error also increased the nearer to the grid centre.

3.7.1. Calibration

In the calibrated model within G32, the NRMSE ranged between 0.057 and 0.114 for longitudinal velocity fields and for transversal velocity fields ranged between 0.075 and 0.125. Errors were slightly larger for the higher resolution data set. Ranges of NRMSE were between 0.064 and 0.133 for longitudinal velocity fields and between 0.099 and 0.183 for transversal velocity fields. R^2 values were > 0.816 and all tests were significant.

3.7.2. Validation

The validated model using G32 showed a range of NRMSE ranging between 0.048 and 0.109 for longitudinal velocity fields and for transversal velocity fields it ranged between 0.054 and 0.123. R^2 values were > 0.792. Similarly, these errors increased in the higher resolution grid. Ranges of NRMSE were between 0.052 and 0.123 for longitudinal velocity fields and between 0.069 and 0.174 for transversal velocity fields. R^2 values in the validated data set were > 0.727. Table 13 shows the error with regards to their distance to the centre of the flood plain. Similarly, velocity errors increased the nearest to the centre of the flood plain.

3.8. Results of velocity fields for configuration 2 (analysis area 2)

3.8.1. Validation

The results show that less error occurred with increasing distance from the manhole, (see Table 14). For G32 NRMSE ranged between 0.052 and 0.125 for longitudinal flows and between 0.067 and 0.145 for lateral flows. Multiple R^2 values were also high (>0.909). For G16 NRMSE for longitudinal flows ranged between 0.060 and 0.137 and between 0.081 and 0.227 for transversal flows. Table 13 shows that error increased with decreasing distance from the grid centre. These results are plotted in Figs. D.5–D.8 in Appendix D. Like configuration 1, the results show that for configuration 2 errors increased with increasing inflows. This error increased when under surcharge conditions. For example, tests being conducted under S1 and S2 had less error than tests under S3. Along similar lines the higher resolution grid also had larger

error than the lower resolution grid.

3.9. Results of velocity fields of configuration 2 (analysis area 3)

3.9.1. Validation

These results show the velocity field error for analysis area 3, which was situated inside a parking space. They are illustrated in Table 15 below and Figs. D.9–D.12 in Appendix D. The results suggest that the model outcome was acceptable, as the %NRMSE was below the 30% limit and R^2 was high (>0.979). The results suggest that NRMSE ranged between 0.09 and 0.177 for longitudinal flows, and for transversal flows it ranged between 0.111 and 0.215 in G32. For G16 NRMSE ranged between 0.098 and 0.182 for longitudinal flows and for transversal error ranged between 0.115 and 0.225.

4. Discussion

4.1. Manhole exchange, pipe, and surface flows

As flow rates in the model increased the model performance decreased. For example, during S1 and S2 the inflow at Q_1 dominated the flow conditions. This meant that as the inflow increased, so did the simulated error. Along the same lines, this was apparent for S3 where the model was influenced by Q_3 input flows. This meant that the error of S3 increased with increasing Q_3 inflows. Errors in the manhole exchange, pipe and surface flows were higher in configuration 2 than configuration 1. Similarly, this is very likely to do with the increased complexity of the floodplain; plus, calibration took place using a lower spatial resolution.

Throughout the hydraulic testing scenarios, the numerical model overestimated manhole exchange for all scenarios. This meant that the model either overestimated pipe outflow or overestimated surface outflow. For S1 and S2, simulated pipe outflow (Q_4) was greater in the numerical model than the experimental datasets. As a result, 2D overland flow was smaller than observed values under these scenarios. Along similar lines for S3, over calculated surcharged flow meant that pipe outflows were smaller in the numerical model than the experimental datasets. Such errors have been experienced in previous studies. For example, Fraga et al. (2017) found that their model over estimated surcharge by 2%, which is slightly in better agreement than the model in the present study. Rubinato et al. (2017) calibrated their flood model under steady flow conditions and found that their model over estimated unsteady surcharge flow rates. This occurred due to turbulent conditions created by unsteady flows. In the present study the manhole exchange error comes to pass due to S1 and S2 being heavily influenced by Q_1 inflow, whereas S3 is determined by Q_3 pipe inflow.

4.2. Surface depths

The errors of simulated surface depths increased with increasing flow rate. This was also greater in configuration 2 when compared to configuration 1. Differences in depths were < 2 mm for all depth locations during the three hydraulic testing scenarios. This is consistent to other studies. Rubinato et al. (2016) had a l-3 mm disparity after changing downstream boundary conditions (adjusting a weir) when modelling depths around a manhole. Similarly, Martins et al. (2017) had a range of depths just above 2 mm when validating 2D shock capturing flood models around a surcharging manhole. The results from the present study suggest manhole exchange dynamics are attributable to changes in surface depths. For example, during S1 and S2, Q1 inflow influenced the flow at Q_e and Q_2. For both scenarios the results confirm that there was more error at depth locations with higher Q_1 inflow. These results are consistent with a previous study. For example, Martins et al. (2017) found that higher Q_1 increased the variation in depths. The biggest depth discrepancy in their study came from P_2, as this point is upstream of the manhole. This meant that this location was most effected by overland inflow. During the present study, we conclude that P_2 was also strongly affected by Q_1 during S1 and S2, though most error came from depth location P_3, which is situated downstream of the manhole. During S3 depths were slightly overestimated. This was due to pipe inflow at Q_3 dominating discharges through the manhole and over the flood plain. Thus, during simulation the manhole magnified this exchange. Similarly, the variation in depths increased at each location with higher inflows at Q_3. This meant that model error increased with increasing Q_3.

In the current study, depths were related to the dominate boundary inflow. Depths during S1 and S2 increased with increasing Q_1 and S3 was increased with Q_3. The variation of depths also increased with the increasing inflow boundary. Though, the results do indicate that there was a stable and balanced flow during the simulations, which is agreeable with many of the observed values. In the observed data, P_1 had a greater depth than P_5 under S3. Even though the locations were the same distance from the manhole. This was due to P_1 being closer to the boundary wall of the physical model. This changes the wave reflections of the flow, causing some parts of the 2D surface to have greater depths than others. The numerical model was able to replicate this well. This is seen by the relationship between the depth trends and Q_e flow exchange in both the observed and simulated results. This has been accurately replicated in other numerical modelling studies, for example in the work of Kesserwani et al. (2015). The authors found that depths increased at P_1 and P_5 with increasing Q_e. P_1 also had a greater depth than P_5, due to the geometry of the model.

4.3. Velocity fields (analysis area 1)

For configuration 1 the results show that for longitudinal flows in G32 model performance decreased with increasing inflows. Surface flows and depths were underestimated during S1 and S2, as discussed in the exchange flow results above; this meant that velocities were also slightly under calculated. For S3 the opposite occurred, as flows were over calculated. Longitudinal flows had more error than transversal flows. This meant that the model performed better in flows that were more directional and less turbulent, as seen in cross flows. Along similar lines, an increase in error was directly linked to inflow conditions. Velocity error increased with increasing Q_1 inflow for S1 and S2. For S3, velocity error increased with Q_3 inflow.

The model did not perform as well with the validated higher resolution grid, albeit the modelling outcome was considered good. The highest %NRMSE in G16 was 19.2%, which is well within the acceptable model performance percentage (<30%). This agrees to other outcomes, for example in the works of Willis et al. (2019). They found that increasing errors can occur in higher resolution data sets due to an increase in model complexity.

During configuration 2 the model performance also decreased with increasing inflows, which mirrored configuration 1 results. Due to surface flows and depths being under calculated for S1 and S2, this meant that velocities were also under simulated. Q_1 inflow was responsible for these errors. S3 had higher simulated velocities, due to over simulated flows and depths due to Q_3 inflows. For G16 the highest %NRMSE was during lateral flows under T18 (20.1%), yet this still was still below the acceptable limit (<30%). The model performed less accurately (26.7% error) in analysis area 1 during surcharge conditions in the higher resolution grid. This was at a range of 125 mm to the manhole centre during T18, which was consistent to the first street profile.

4.4. Velocity fields (analysis area 2 and 3)

Model performance improved the further away velocities were recorded from the manhole. For example, the model performed better in analysis area 2 than 1. This is due to the manhole being a more complex area to simulate as verified in other works (Martins et al., 2018; Willis et al., 2019; Shrestha et al., 2022). Similarly model performance improved in longitudinal flows than transversal, as they were more

dominant and less complex. As seen in analysis area 1, model performance also decreased in the higher resolution dataset. The highest % NRMSE during surcharge at G16 was 18.3%, which is well within the he acceptable model performance percentage (<30%). Out of all velocity field analysis locations model performance in analysis area 3 was the lowest. This was likely the most complex region to simulate velocity fields in configuration 2. Here, flows were more turbulent. Similarly, longitudinal velocity field performance was greater than transversal velocity fields, and the performance of the model decreased with increasing spatial resolution. Which was likely to do with the model being calibrated in a lower resolution.

4.5. Spatial resolution, calibration, and computational time trade-off.

The key debate that this paper covers is model performance and the associated errors with regards to increasing inflows, and surface complexities. The keys findings found that model performance did decrease with increasing pipe and surface inflows, spatial resolution of the 2D surfaces used, and when parking spaces were introduced. Also, the model performed better in the calibrated lower resolution grids. Model error linked to numerical complexity has been discussed in previous research for example Mateo et al. (2017); Willis et al. (2019); Kim et al. (2021). The authors found that if parameter calibrations are conducted in lower spatial resolution model performance can decrease when the resolution is increased in future simulations. Though this uncertainty is far less when flows are less complex. They argue that simplified models can be just as useful as more complicated models, which agrees with the present paper. These differences are likely to be reduced through the calibration process. Meaning that calibration should take place using the highest resolution possible. However, one could also argue that there is a trade-off between resolution and computational times. This is something that the modeller needs to consider for themselves. For example, Table 2 illustrates that in the higher resolution grid computational times were about twice as long than the lower resolution grid. Therefore, we argue that we found it acceptable in our example to use a lower resolution model. As errors were not large enough to reduce model performance enough and simulations were far less time consuming.

4.6. Limitations of study

The biggest limitation of this study is that it only looked at two hypothetical street profiles that were both relatively simple. Therefore, the differences in errors were quite small between each configuration and opposing resolution. It is proposed that future work may look at creating more obstacles on the floodplain. This may resemble street furniture in the real world. For example, benches, plant containers or curbs. Another important point to consider is that this study only looked at MIKE URBAN + as the modelling software and made no comparison with any other. It is hypothesised that if this work was carried out in a similar commercial package (e.g., Info works) the results would be similar, as both MIKE and info works use shallow water equations and finite volume numerical schemes. Though results maybe slightly different if using UIM (Chen et al., 2007), as UIM uses diffusive wave equations and a finite difference numerical scheme.

5. Conclusion

The need to understand the accuracy of 2D velocity fields across flooded urban surfaces has motivated this research. Water facilities such as the one used in this study are a useful way to obtain high resolution data. Yet, calibrating high-resolution models can be computationally demanding. Therefore, it may be necessary to initially calibrate flood models in lower resolutions, reducing simulation times. To increase the accuracy of flood models after calibration it is possible to increase spatial resolution. This allows for a more detailed investigation of inundation over an area, paying closer attention to small streams and narrow flow conduits. Yet, error is likely to increase from the elevated model complexity. As such, this paper analysed this process by comparing high resolution data sets collected from a water laboratory with a numerical model. The level of complexity was increased within two hypothetical street profiles to assess if it could influence the performance of the numerical simulations. The model was calibrated before spatial resolution was increased. This study suggests that hydrodynamic 1D-2D models with a lower spatial resolution may still be appropriate for modelling urban inundation. This is because computational times are reduced, and errors may still be acceptable. Our findings provide a valuable understanding of urban flooding scenarios using a high spatial resolution experimental dataset. We conclude that the data set is reproducible for future studies and is suitable for the calibration and validation of other current numerical models. The model was able to replicate highly accurate flows, surface depths and velocities, when compared to the physical experimental data sets. Model error increased with increasing inflows. Similarly, error increased when the spatial resolution became finer due to the model being calibrated in the less complex system. The conclusions can be simplified as follows-.

In both configurations-

- Model performance (flows, depths, and velocity fields) decreased with increasing pipe and surface inflows.
- Model performance decreased in the finer spatial resolution due to calibration taking place with a courser resolution.
- Velocity field error increased with decreasing distance to the manhole and grid centre.

Configuration 1 where model performance was improved-

- Validated flow errors were $\leq$ 7.9%
- Validated surface depth errors were $\leq$ 5.6%
- Validated velocity field errors were $\leq$ 19.2%

Configuration 2 where model performance was reduced (due to parking spaces being added)-

- Validated flow errors were $\leq$ 8.2%
- Validated surface depth errors were $\leq$ 6.4%
- Validated velocity field errors were $\leq$ 22.7%

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgements

The work presented in this paper was carried out as part of PhD research and was supported by the UK Engineering and Physical Sciences Research Council Centre for Doctoral Training in Water Informatics Science and Engineering (grant ID: EP/L016214/1). The experimental datasets used in this work were produced by the UK Engineering and Physical Sciences Research project - ID: EP/K040405/1. Open Access experimental datasets used in this paper are available at https://zenodo.org/communities/floodinteract/.

Appendix A - Pipe flow plots

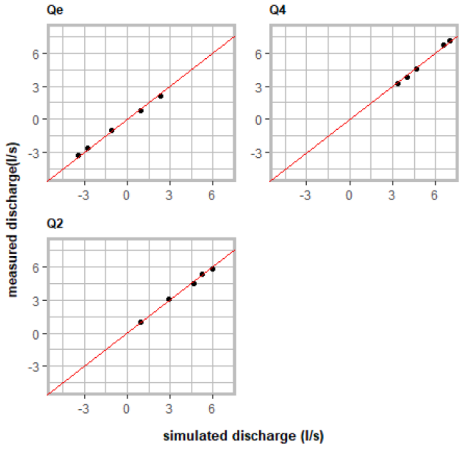


Fig. A1. Combined scenario pipe flows of configuration 1 - calibration data set.

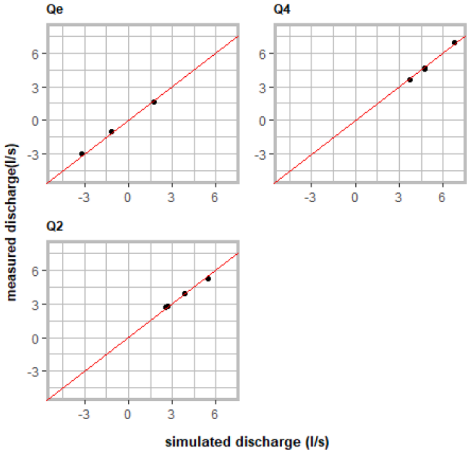


Fig. A2. Combined scenario pipe flows of configuration 1 - validation data set.

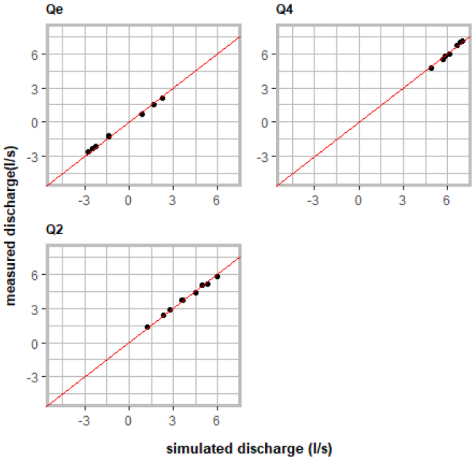


Fig. A3. Combined scenario pipe flows of configuration 2 - validation.

Appendix B – Surface depth plots

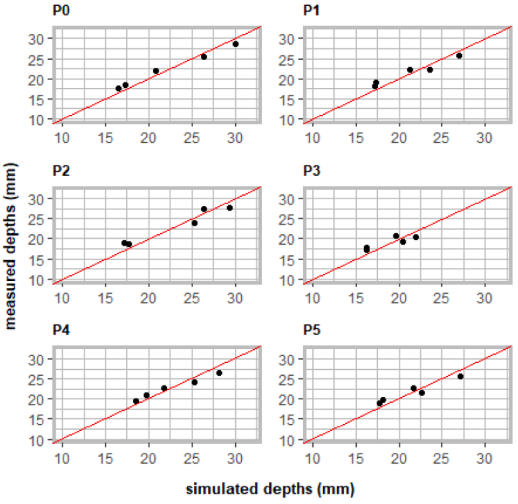


Fig. B1. Combined scenario surface depths of configuration 1- calibration data set.

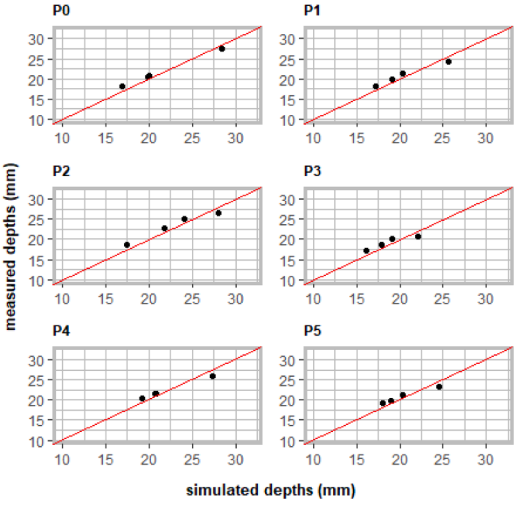


Fig. B2. Combined scenario surface depths of configuration 1- validation data set.

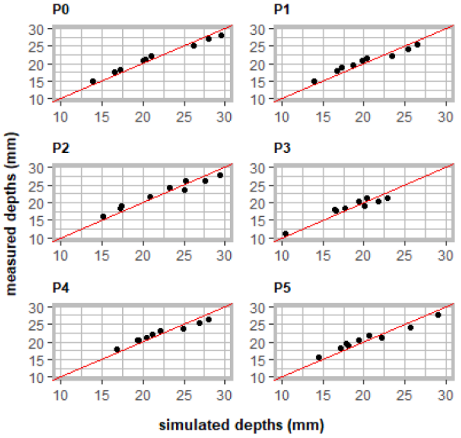


Fig. B3. Combined scenario surface depths of configuration 2- validation.

Appendix C – Configuration 1 velocity plots

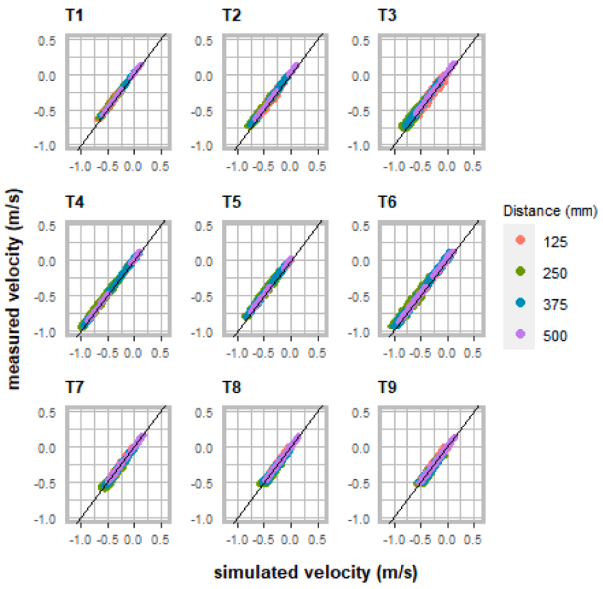


Fig. C1. Longitudinal velocity fields of configuration 1 with G32 in analysis area 1. Showing distances from the manhole centre at each test (T1-T9).

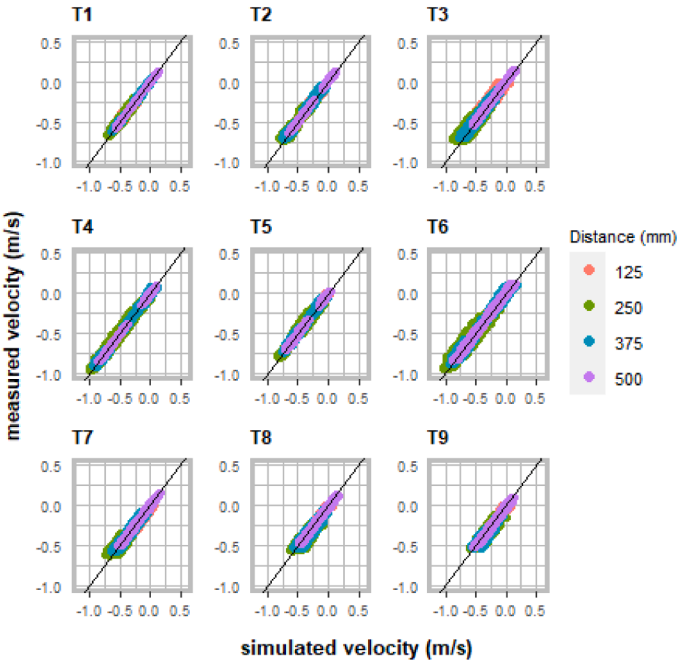


Fig. C3. Longitudinal velocity fields of configuration 1 with G16 in analysis area 1. Showing distances from the manhole centre at each test (T1-T9).

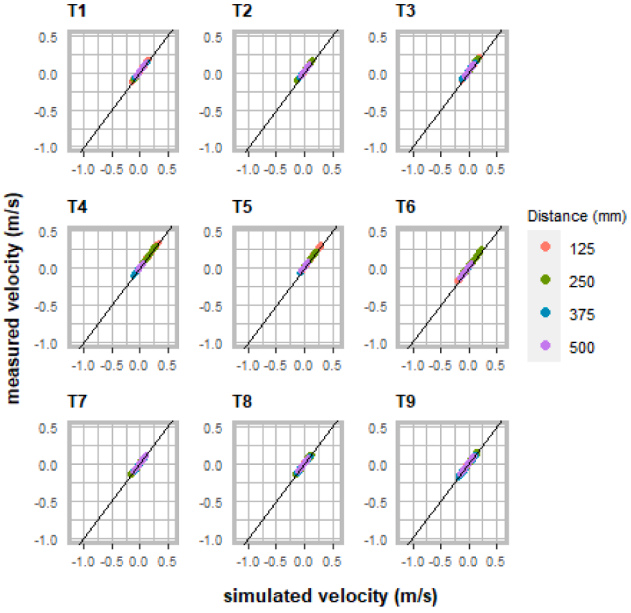


Fig. C2. Transversal velocity fields of configuration 1 with G32 in analysis area 1. Showing distances from the manhole centre at each test (T1-T9).

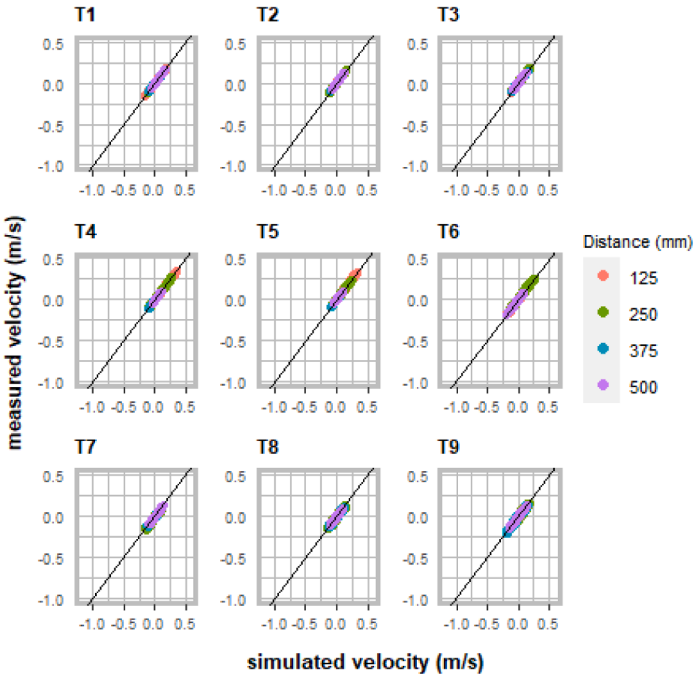


Fig. C4. Transversal velocity fields of configuration 1 with G16 in analysis area 1. Showing distances from the manhole centre at each test (T1-T9).

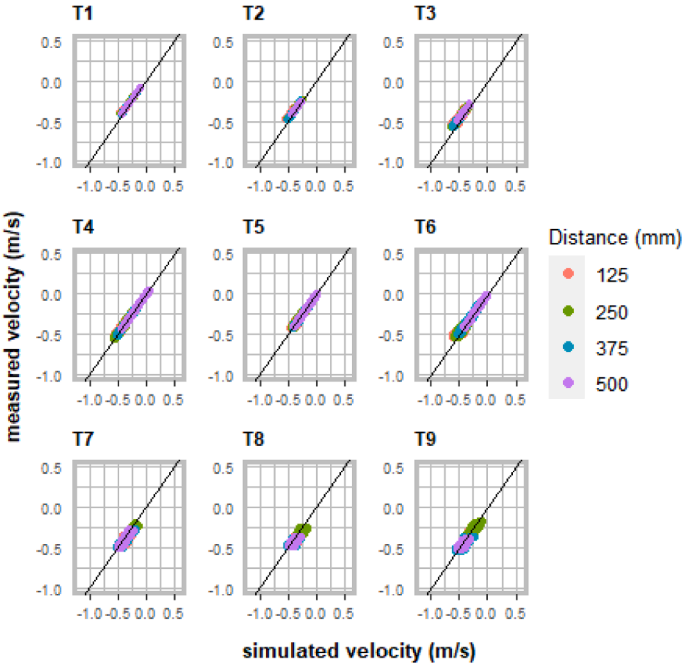


Fig. C5. Longitudinal velocity fields of configuration 1 with G32 in analysis area 2. Showing distances from the grid centre at each test (T1-T9).

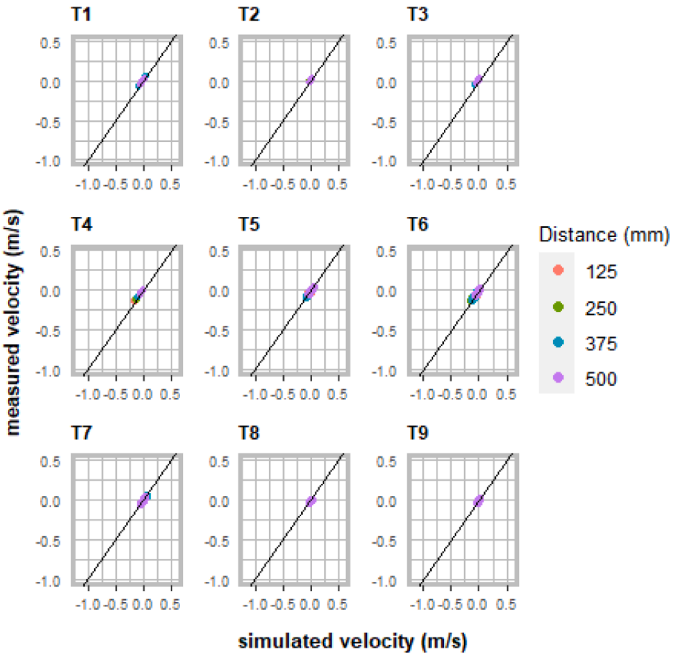


Fig. C6. Transversal velocity fields of configuration 1 with G32 in analysis area 2. Showing distances from the grid centre at each test (T1-T9).

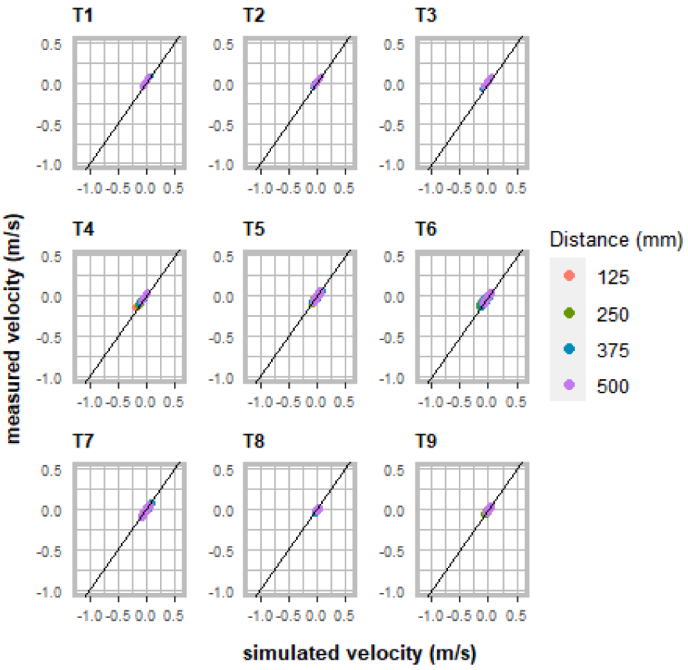


Fig. C8. Transversal velocity fields of configuration 1 with G16 in analysis area 2. Showing distances from the manhole centre at each test (T1-T9.

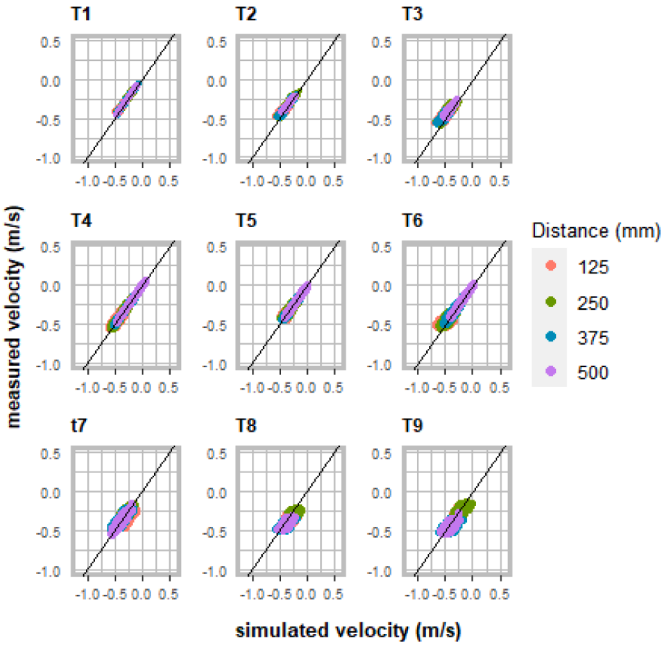


Fig. C7. Longitudinal velocity fields of configuration 1 with G16 in analysis area 2. Showing distances from the grid centre at each test (T1-T9).

Appendix D – Configuration 2 velocity plots

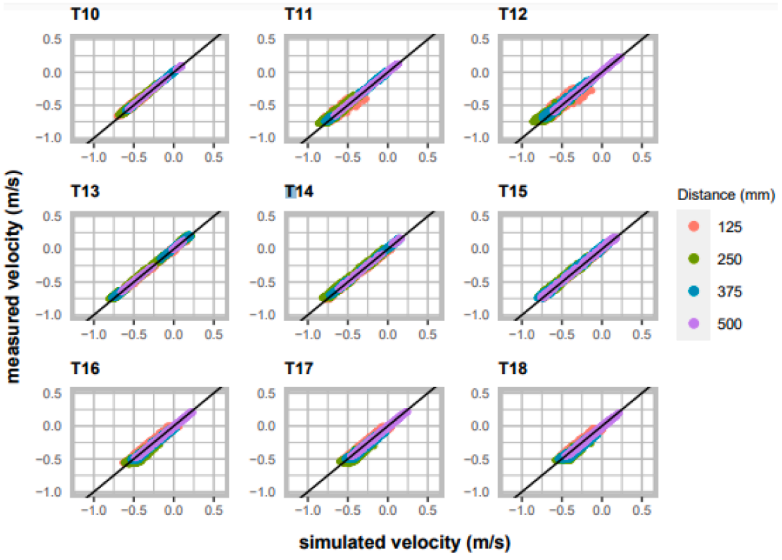


Fig. D1. Longitudinal velocity fields of configuration 2 with G32 in analysis area 1. Showing distances from the manhole centre at each test (T10-T18).

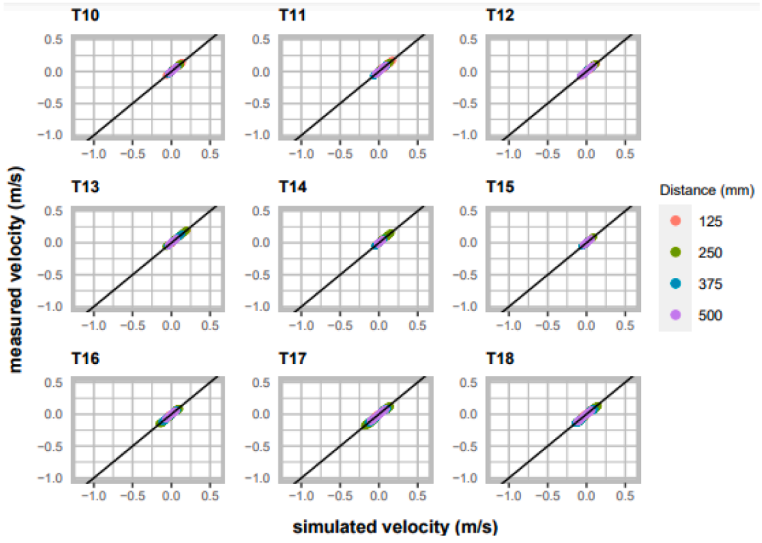


Fig. D2. Transversal velocity fields of configuration 2 with G32 in analysis area 1. Showing distances from the manhole centre at each test (T10-T18).

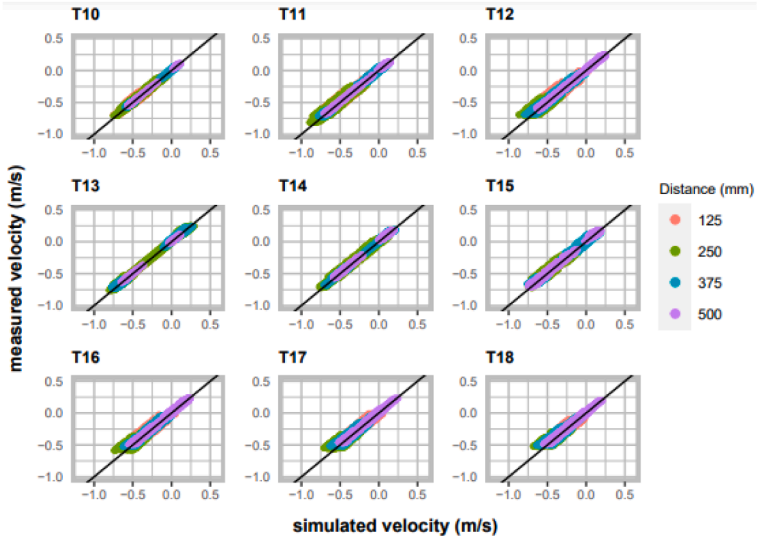


Fig. D3. Longitudinal velocity fields of configuration 2 with G16 in analysis area 1. Showing distances from the manhole centre at each test (T10-T18).

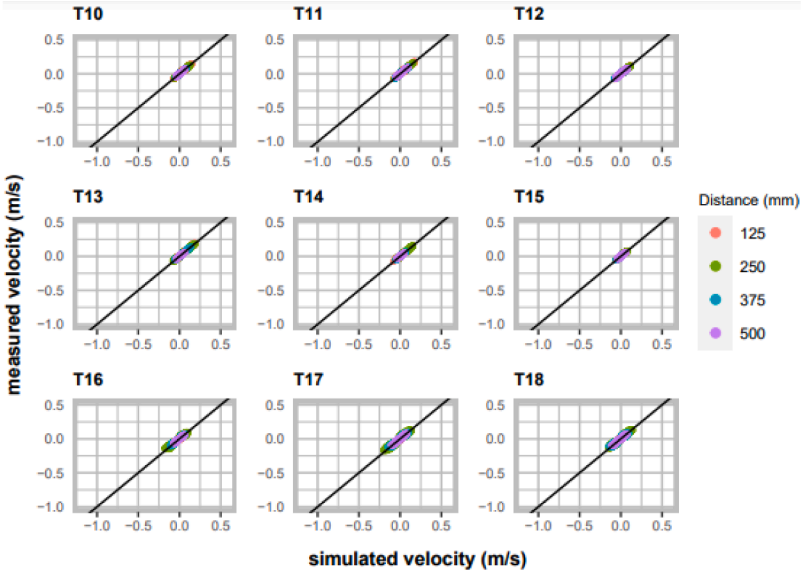


Fig. D4. Transversal velocity fields of configuration 2 with G16 in analysis area 1. Showing distances from the manhole centre at each test (T10-T18).

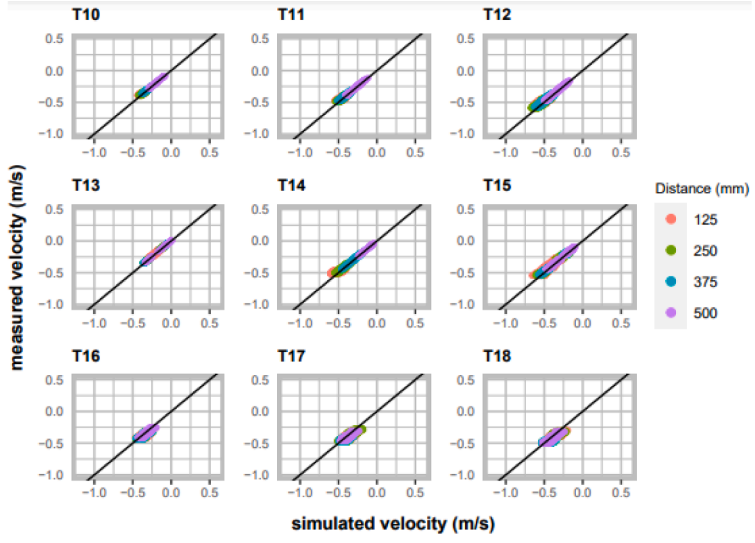


Fig. D5. Longitudinal velocity fields of configuration 2 with G32 in analysis area 2. Showing distances from the grid centre at each test (T10-T18).

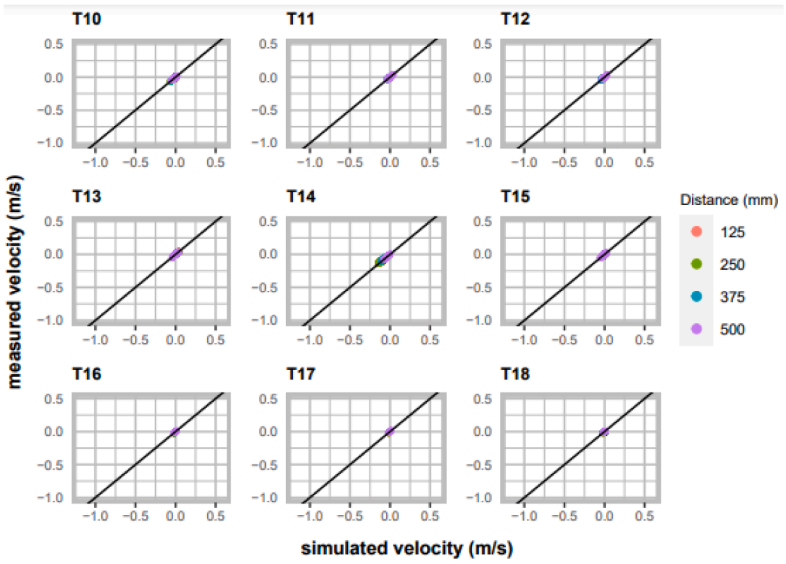


Fig. D6. Transversal velocity fields of configuration 2 with G32 in analysis area 2. Showing distances from the grid centre at each test (T10-T18).

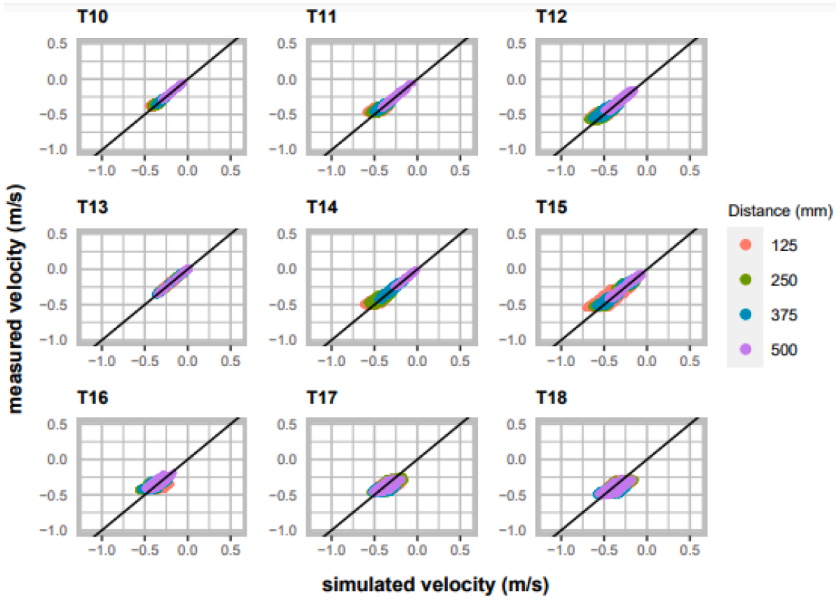


Fig. D7. Longitudinal velocity fields of configuration 2 with G16 in analysis area 2. Showing distances from the grid centre at each test (T10-T18).

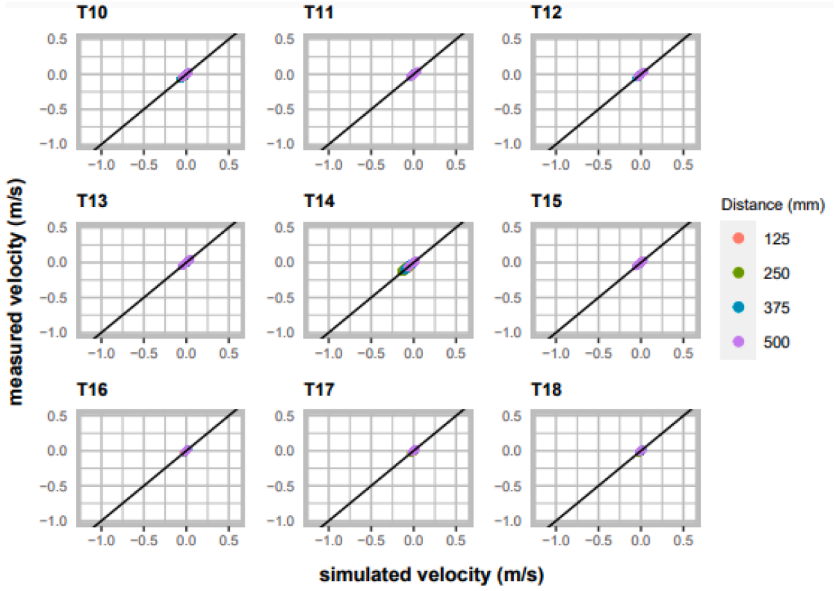


Fig. D8. Transversal velocity fields of configuration 2 with G16 in analysis area 2. Showing distances from the grid centre at each test (T10-T18).

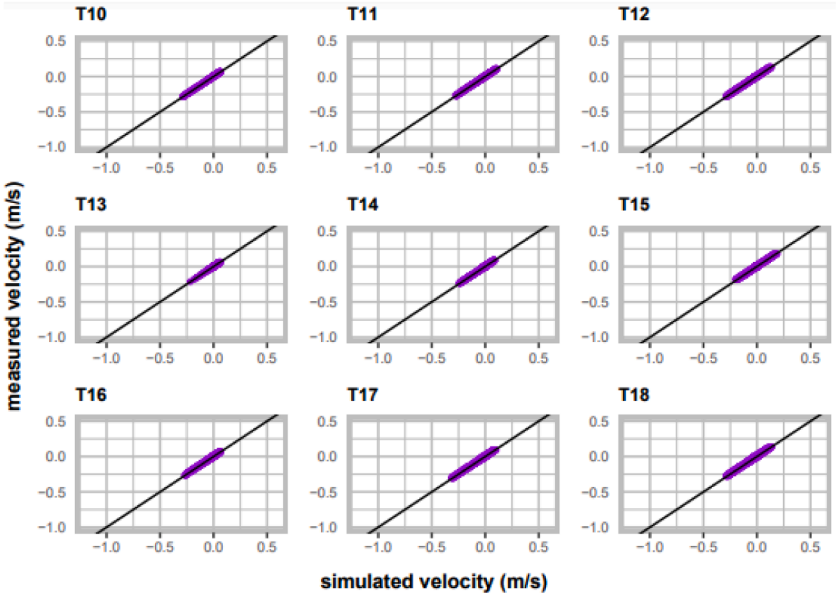


Fig. D9. Longitudinal velocity fields of configuration 2 with G32 in analysis area 3. Showing each test (T10-T18).

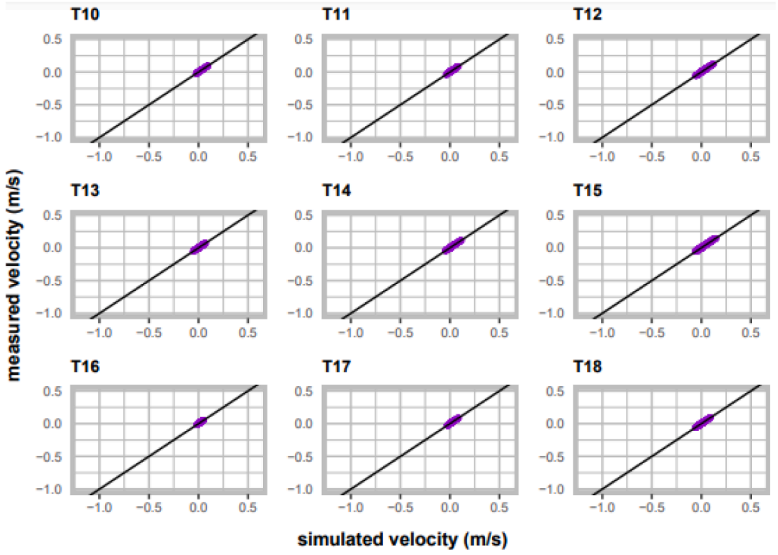


Fig. D10. Transversal velocity fields of configuration 2 with G32 in analysis area 3. Showing each test (T10-T18).

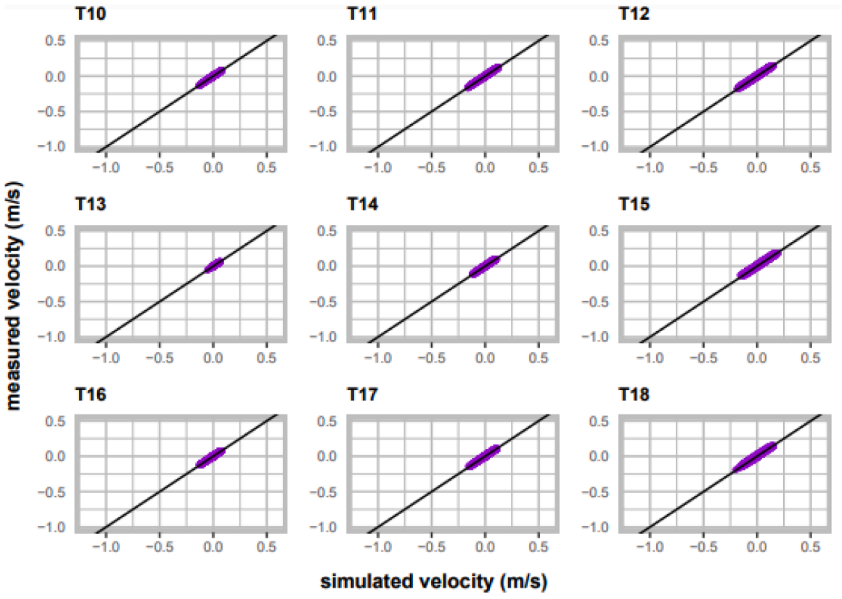


Fig. D11. Longitudinal velocity fields of configuration 2 with G16 in analysis area 3. Showing each test (T10-T18).

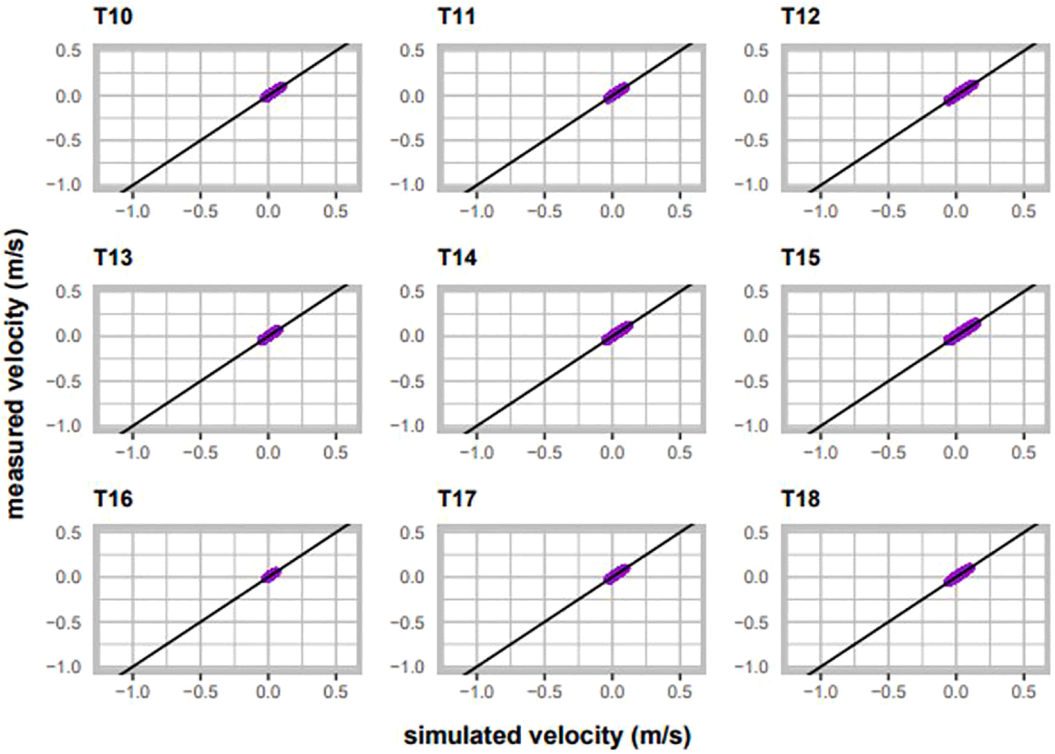


Fig. D12. Transversal velocity fields of configuration 2 with G16 in analysis area 3. Showing each test (T10-T18).

References

Addison-Atkinson, W., Chen, A.S., Memon, F.A., Chang, T., 2022. Modelling urban sewer flooding and quantitative microbial risk assessment: a critical review. J. Flood Risk Manage.

Aronica, G., Bates, P.D., Horritt, M.S., 2002. Assessing the un-certainty in distributed model predictions using observed binary pattern information within GLUE. Hydrol. Process. 16 (10), 2001–2016.

Avand, M., Kuriqi, A., Khazaei, M., Ghorbanzadeh, O., 2022. DEM resolution effects on machine learning performance for flood probability mapping. J. Hydro Environ. Res. 40, 1–16.

Bates, P.D., 2004. Remote sensing and flood inundation modelling. Hydrol. Process. 18 (13), 2593–2597.

Bates, P., De Roo, A., 2000. A simple raster-based model for flood inundation simulation. J. Hydrol. 236(l-2):54–77.

Bazin, P.-H., Mignot, E., Paquier, A., 2017. Computing flooding of crossroads with obstacles using a 2D numerical model. J. Hydraul. Res. 55 (1), 72–84.

Beg, M.N.A., Carvalho, R.F., Tait, S., Brevis, W., Rubinato, M., Schellart, A., Leandro, J., 2018. A comparative study of manhole hydraulics using stereoscopic PIV and different RANS models. Water Sci. Technol. 2017 (1), 87–98.

Carmer, C.F., Rummel, A.C., Jirka, G.H., 2009. Mass transport in shallow turbulent wake flow by planar concentration analysis technique. J. Hydraul. Eng. 135 (4), 257–270.

Cea, L., Garrido, M., Puertas, J., Jacorne, A., Del Rio, H., Suarez, J., 2010. Overland flow computations in urban and industrial catchments from direct precipitation data using a two-dimensional shallow water model. Water Sci. Technol. 62 (9), 1998–2008.

Chen, A., Djordjevic, S., Leandro, J., and Savic, D. (2007). The urban inundation model with bidirectional flow interaction between 2d overland surface and Id sewer networks. *NOVATECH 2007*, Lyon, France.

Collender, P., Cooke, O., Bryant, L., Kjeldsen, T., Remais, J., 2016. Estimating the microbiological risks associated with inland flood events: Bridging theory and models of pathogen transport. Crit. Rev. Environ. Sci. Technol. 46 (23–24), 1787–1833.

Despotovic, M., Nedic, V., Despotovic, D., Cvetanovic, S., 2016. Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation. Renew. Sustain. Energy Rev. 56, 246–260.

Dewan, T.H., 2015. Societal impacts and vulnerability to floods in Bangladesh and Nepal. Weather Clim. Extremes 7, 36–42.

Dhi, 2020. Mike Urban + Model Manager. User guide, Danish Hydraulic Institute, Horsholm, Denmark.

El Kadi Abderrezzak, K., Lewicki, L., Paquier, A., Rivière, N., Travin, G., 2011. Division of critical flow at three-branch open-channel intersection. J. Hydraul. Res. 49 (2), 231–238.

Estrela, T., Quintas, L., 1994. Use of a GIS in the modelling of flows on floodplains. In: 2nd International Conference on River Flood Hydraulics. John Wiley and Sons Ltd., Chichester, pp. 177–190.

Fraga, L., Cea, L., Puertas, J., 2017. Validation of a 1D–2D dual drainage model under unsteady part-full and surcharged sewer conditions. Urban Water J. 14 (l), 74–84.

Ghostine, R., Kesserwani, G., Vazquez, J., Rivière, N., Ghenaim, A., Mose, R., 2009. Simulation of supercritical flow in crossroads: Con-frontation of a 2D and 3D numerical approaches to experimental results. Comput. Fluids 38 (2), 425–432.

Ghostine, R., Mignot, E., Abdallah, M., Lawniczak, F., Vazquez, J., Mosé, R., Grégoire, C., 2010. Discontinuous galerkin finite-element method for simulation of flood in crossroads. J. Hydraul. Eng. 136 (8), 474–482.

Hall, J., Arheimer, B., Borga, M., Brâzdil, R., Claps, P., Kiss, A., Kjeldsen, T.R., Kriauciünienė, J., Kundzewicz, Z.W., Lang, M., Llasat, M.C., Macdonald, N., McIntyre, N., Mediero, L., Merz, B., Merz, R., Molnar, P., Montanari, A., Neuhold, C., Parajka, J., Perdigâo, R.A.P., Plav- cová, M., Rogger, M., Salinas, J.L., Sauquet, E., Schâr, C., Szolgay, J., Viglione, A., Blôschl, G., 2014. Understanding flood regime changes in Europe: a state-of-the-art assessment. Hydrol. Earth Syst. Sci. 18 (7), 2735–2772.

Hillas, T., 2014. Reducing the occurrence of flooding through the effective management of sewer blockage thesis (Ph.D.). University of Exeter.

Hunter, N., Bates, P., Neelz, S., Pender, G., Villanueva, L, Wright, N., Liang, D., Falconer, R., Lin, B., Waller, S., Crossley, A., and Mason, D. (2008). Benchmarking 2D hydraulic models for urban flooding. Proceedings of the Institution of Civil Engineers - Water Management, 161(l):13-30.

Jongman, B., Hochrainer-Stigler, S., Feyen, L., Aerts, J.C.J.H., Mechler, R., Botzen, W.J. W., Bouwer, L.M., Pflug, G., Roj as, R., and Ward, P. J., 2014. Increasing stress on disaster-risk finance due to large floods. Nat. Clim. Chang. 4 (4), 264–268.

Kesserwani, G., Lee, S., Rubinato, M., and Shucksmith, J. (2015). Experimental and numerical validation of shallow water flow around a surcharging manhole. *10th International Urban Drainage Conference September 20-23 Quebec Canada*, page 10.

Kim, S., Shen, H., Noh, S., Seo, D.-J., Welles, E., Pelgrim, E., Weerts, A., Lyons, E. and Philips, B. (2021) 'High-resolution modeling and prediction of urban floods using WRF-Hydro and data assimilation.' *Journal of Hydrology*, 598, July, p. 126236.

Leandro, J., Chen, A.S., Djordjevic, S., Savic, D.A., 2009. Comparison of 1D/1D and 1D/2D Coupled (Sewer/Surface) hydraulic models for urban flood simulation. J. Hydraul. Eng. 135 (6), 495–504.

Leitâo, J.P., Peha-Haro, S., Lüthi, B., Scheidegger, A., Moy de Vitry, M., 2018. Urban overland runoff velocity measurement with consumer-grade surveillance cameras and surface structure image ve- locimetry. J. Hydrol. 565, 791–804.

Liang, Q., 2010. Flood simulation using a well-balanced shallow flow model. J. Hydraul. Eng. 136 (9), 669–675.

Maksimovic, C., Prodanovic, D., Boonya-Aroonnet, S., Leitâo, J.P., Djordjevic, S., Allitt, R., 2009. Overland flow and pathway analysis for modelling of urban pluvial flooding. J. Hydraul. Res. 47 (4), 512–523.

Martins, R., Kesserwani, G., Rubinato, M., Lee, S., Leandro, J., Djordjevic, S., Shucksmith, J., 2017. Validation of 2D shock capturing flood models around a surcharging manhole. Urban Water J. 14 (9), 892–899.

Martins, R., Rubinato, M., Kesserwani, G., Leandro, J., Djordjevic, S., Shucksmith, J.D., 2018. On the characteristics of velocities fields in the vicinity of manhole inlet grates during flood events. Water Resour. Res. 54 (9), 6408–6422.

Mateo, C. M. R., Yamazaki, D., Kim, H., Champathong, A., Vaze, J. and Oki, T. (2017) 'Impacts of spatial resolution and representation of flow connectivity on large-scale simulation of floods.' *Hydrology and Earth System Sciences*, 21(10) pp. 5143–5163. Mignot, E., Paquier, A., and Ishigaki, T. (2006). Comparison of numerical and experimental simulations of a flood in a dense urban area. *Water Science and Technology*, 54(6-7):65 73.

Mignot, E., Paquier, A., Ishigaki, T., 2006. Comparison of numerical and experimental simulations of a flood in a dense urban area. Water Science and Technology 54 (6–7), 65–73.

Mignot, E., Paquier, A., Rivière, N., 2008a. Experimental and numerical modeling of symmetrical four-branch supercritical. J. Hydraul. Res. 46 (6), 723–738.

Mignot, E., Rivière, N., Perkins, R., Paquier, A., 2008b. Flow patterns in a four-branch junction with supercritical flow. J. Hydraul. Eng. 134 (6), 701–713.

Molinari, D., De Bruijn, K.M., Castillo-Rodrfguez, J.T., Aronica, G.T., Bouwer, L.M., 2019. Validation of flood risk models: Current practice and possible improvements. Int. J. Disaster Risk Reduct. 33, 441–448.

Moy de Vitry, M., Dicht, S., Leitão, J.P., 2017. floodX: urban flash flood experiments monitored with conventional and alternative sensors. Earth Syst. Sci. Data 9 (2), 657–666.

Moy de Vitry, M., Kramer, S., Villez, K., and Wegner, J.D.and Leitão, J. (2018). Calibrating urban flood models with qualitative probabilistic flooding information extracted from CCTV footage. *Proceedings of the 11th Int. Conference on Urban Drainage Modelling, 23-26 Sep. 2018, Palermo, Roly*.

Pan, A., Hou, A., Tian, F., Ni, G., Hu, H., 2012. Hydrologically Enhanced Distributed Urban Drainage Model and Its Application in Beijing City. J. Hydrol. Eng. 17 (6), 667–678.

Rojas Arques, S., Rubinato, M., Nichols, A., Shucksmith, J.D., 2018. Cost effective measuring technique to simultaneously quantify 2D velocity fields and dept h-averaged solute concentrations in shallow water flows. Flow Meas. Instrum. 64, 213–223.

Rubinato, M., Lee, S., Kesserwani, G., and Shucksmith, J. (2016). Experimental and numerical investigation of water depths around a manhole under drainage conditions. *12th International Conference on Hydroinformatics HIC 2016*.

Rubinato, M., Nichols, A., Peng, Y., Zhang, J.-m., Lashford, C., Cai, Y.- p., Lin, P.-z., and Tait, S. (2019). Urban and river flooding: Comparison of flood risk management approaches in the UK and China and an assessment of future knowledge needs. *Water Science and Engineering*, 12(4):274-283.

Rubinato, M., Martins, R., Kesserwani, G., Leandro, J., Djordjevic, S., Shucksmith, J., 2017. Experimental calibration and validation of sewer/surface flow exchange equations in steady and unsteady flow conditions. J. Hydrol. 552, 421–432.

Rubinato, M., Lee, S., Martins, R., Shucksmith, J.D., 2018a. Surface to sewer flow exchange through circular inlets during urban flood conditions. J. Hydroinf. 20 (3), 564–576.

Rubinato, M., Martins, R., Shucksmith, J., 2018b. Quantification of energy losses at a surcharging manhole. Urban Water J. 15 (3), 234–241.

Rubinato, M., Lashford, C., Goerke, M., 2021. Advances in experimental modelling of urban flooding. In: Wang, X.C., Fu, G. (Eds.), Water- Wise Cities and Sustainable Water Systems: Concepts, Technologies, and Applications. IWA Publishing, pp. 235–257.

Rubinato, M., Helms, L., Vanderlinden, M., Hart, J., Martins, R., 2022. Flow exchange, energy losses and pollutant transport in a surcharging manhole linked to street profiles. J. Hydrol. 604.

Shrestha, A., Mascaro, G., Garcia, M., 2022. Effects of stormwater infrastructure data completeness and model resolution on urban flood modeling. J. Hydrol. 607, 127498.

Soares-Frazão, S., Zech, Y., 2008. Dam-break flow through an idealised city. J. Hydraul. Res. 46 (5), 648–658.

Stewart, I.T., Ficklin, D.L., Carrillo, C.A., McIntosh, R., 2015. 21st century increases in the likelihood of extreme hydrologic conditions for the mountainous basins of the Southwestern United States. J. Hydrol. 529, 340–353.

Svetlana, D., Radovan, D., Ján, D., 2015. The economic impact of floods and their importance in different regions of the world with Emphasis on Europe. Proc. Econ. Finance 34, 649–655.

Van Emelen, S., Soares-Frazão, S., Riahi-Nezhad, C.K., Hanif Chaudhry, M., Imran, J., Zech, Y., 2012. Simulations of the New Orleans 17th Street Canal breach flood. J. Hydraul. Res. 50 (1), 70–81.

Vojinovic, Z., Tutiilic, D., 2009. On the use of ID and coupled 1D–2D modelling approaches for assessment of flood damage in urban areas. Urban Water J. 6 (3), 183 199.

Weitbrecht, V., Kühn, G., Jirka, G., 2002. Large scale PIV- measurements at the surface of shallow water flows. Flow Meas. Instrum. 13 (5–6), 237–245.

Willis, T., Wright, N., Sleigh, A., 2019. Systematic analysis of uncertainty in 2D flood inundation models. Environ. Model. Softw. 122, 104520.

Wirahadikusumah, R., Abraham, D.M., Iseley, T., Prasanth, R.K., 1998. Assessment technologies for sewer system rehabilitation. Autom. Constr. 7 (4), 259–270.