Journal of Hydrology 662 (2025) 133821

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

Optimizing hydrological modelling on real urban catchment: impact of calibration data selection and objective function

Mohammed N. Assaf [a,*], Nicolò Salis [a], Sauro Manenti [a,b], Lorenzo Tamellini [c], Enrico Creaco [a,b], Sara Todeschini [a,b]

[a] Department of Civil Engineering and Architecture (DICAr), University of Pavia, Pavia, Italy
[b] Interdepartmental Centre for Water Research (CRA), University of Pavia, Pavia, Italy
[c] CNR-IMATI, National Research Council - Institute for Applied Mathematics and Information Technologies, Pavia, Italy

A R T I C L E I N F O

This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Zhenxing Zhang, Associate Editor

Keywords:
Hydrological modelling
Model calibration
Objective function
Global sensitivity analysis
Rainfall–runoff model
Stormwater management
Urban hydrology

A B S T R A C T

Hydrological models, like the Stormwater Management Model (SWMM), play a crucial role in water resource management. However, their effectiveness often depends on calibration based on historical rainfall-runoff events. Typically, calibration of urban drainage models is based on a select few observed rainfall-runoff events, which can be chosen from a broader dataset using various selection methods. This study explores the interaction between event selection criteria and objective functions in hydrological modelling. Four event selection approaches (Rainfall Depth (RD), Maximum Rainfall Intensity over a Five-Minute Interval (H5), Mean Intensity (MI), and Hyetograph Centre of Mass (HCM)) and six objective functions, were tested using high-resolution data from 42 rainfall-runoff events in an urban catchment area. Additionally, this research conducts a Global Sensitivity Analysis (GSA) of the drainage model parameters using the Morris method through an innovative integration of Mat-SWMM and the Sensitivity Analysis for Everybody (SAFE) Toolbox, creating a robust framework for assessing how various uncertain input parameters affect the hydrological model outputs. This integration facilitates the GSA process and allows researchers to conduct sensitivity studies on SWMM models using widely available tools. It provides a reproducible, scalable, and computationally efficient framework suitable for handling hydrological datasets of various sizes. Results reveal that sensitivity to model parameters varies significantly depending on the characteristics of the rainfall events. Only five influential parameters, out of the ten tested parameters, were identified and subsequently optimised using a genetic algorithm.

The study highlights the strengths of different event selection criteria for specific outputs: the RD criterion provided accurate estimates for total runoff volume, whereas the HCM and H5 criteria performed better in estimating peak flow rates. The results show that traditional objective functions, such as the squared difference, are particularly sensitive to the choice of the calibration event selection approach.

1. Introduction

Modelling hydrological processes serves as a critical tool for simulating watershed processes and supporting water resource management. However, their reliability is constrained by challenges in model identification, which stem from parameter uncertainty inherently associated with data limitations, model structural inadequacies, and observational errors (Manovam and Todini, 2006). Model identification refers to the process of defining a model structure and selecting appropriate parameter values that best represent the physical system under study. The inherent complexity of hydrological systems, compounded by varied land-use patterns, slopes, and coverage in the urbanized area, leads to complex interactions between different urban drainage structures and systems for wastewater and stormwater. These complexities are magnified by the varying temporal and spatial scales at which these interactions occur, which in turn escalate data requirements and adds to the model's complexity (Hamel and Tim, 2014). These factors, combined with the often-limited availability of detailed data, make it difficult to obtain a physically meaningful set of model parameters (Beven, 2006).

* Corresponding author.
E-mail addresses: mohammed.assaf@unipv.it, ma.mohammed.assaf@gmail.com (M.N. Assaf).

https://doi.org/10.1016/j.jhydrol.2025.133821
Received 31 March 2025; Received in revised form 5 June 2025; Accepted 2 July 2025
Available online 8 July 2025

Hydrological models are simplified representations of natural systems, so model parameters lose clear physical meaning and instead act as conceptual tools to simulate real-world behaviours. Because of this, there is no straightforward link between observed data and model parameters, which means parameter estimation process should be done indirectly. This model estimation process includes adjusting model parameters using location-specific historical data to improve the alignment between simulated and observed hydrological responses, which known as model calibration. The effectiveness of these calibration processes can be attributed to three main aspects: the approach for identifying optimal parameter sets, refinement of model optimisation, and the choice of data for calibration and validation (Beven, 2012).

One key challenge in model calibration is the difficulty of achieving parameter identifiability. This issue stems from three main sources of uncertainty: (i) data limitations, (ii) model uncertainty, and (iii) uncertainties in observations (Beven and Binley, 1992). Data limitations occur due to the limited availability of hydrological measurements across both space and time, often leading to overparameterization and equifinality, where different parameter combinations produce similarly accurate results (Beven, 2001). Model uncertainty arises from the simplifications made in representing real-world hydrological processes, as models use mathematical formulations that may not fully account for the intricate interactions between water, energy, and vegetation (Sorooshian and Gupta, 1983; Sivapalan et al., 2003). Lastly, uncertainties in observations result from measurement errors, equipment limitations, and the challenge of capturing hydrological variables at an appropriate scale (Blöschl and Sivapalan, 1995; Kuczera and Mroczkowski, 1998).

One key factor influencing parameter identifiability is the number of observations used for calibration. While increasing the number of calibration events can contribute to a broader representation of system behaviour, it only partially mitigates parameter equifinality. This is because the model outputs, such as total basin outflow, may be insensitive to certain parameters, meaning that variations in those parameters do not result in distinguishable changes in the output. When adjusting a large number of parameters with limited observational data, multiple parameter sets can produce similar model performance, leading to equifinality, non-uniqueness, and non-identifiability (Beven, 2001; Jakeman and Hornberger, 1993; Blöschl and Sivapalan, 1995; Moradkhani and Sorooshian, 2008; Beven and Freer, 2001). Studies have shown that increasing the number of observations can improve parameter identifiability and reduce equifinality (Kuczera and Mroczkowski, 1998, her and Chaubey, 2015). Recently, Wu et al. (2024) demonstrated that using 3–5 events in the calibration process can significantly decrease the model uncertainty and equifinality while providing proper model performance. Moreover, the resolution of land use cover data of the catchment area was found to have a high influence on model performance and uncertainty (Rahimi and Ebrahimian, 2024). Low-resolution land cover data (like data derived from coarse-resolution satellite imagery) may lead to overestimating or underestimating hydrological responses and increase model uncertainty (Hatt et al., 2004; Walsh et al., 2005; Balha et al., 2023). To mitigate the impact of data limitations and uncertainties in observations, this study draws on a rich 20 years dataset from the Cascina Scala experimental area. It includes long-term hydrological records and detailed land use data, capturing diverse hydrological responses. The large pool of high-resolution data strengthens model calibration and validation, reducing model uncertainty and improving parameter identifiability, making this site an ideal case study for various hydrological modelling research (Barco et al., 2008a; Wang and Altunkaynak, 2012; Granata et al., 2016; Azar et al., 2025).

Reduce the number of model parameters that need to be estimated (calibrated) is considered an effective strategy for addressing equifinality and enhancing the identifiability of parameter values (Hornberger et al., 1985; Jakeman & Hornberger, 1993; Wagener et al., 2001; Moreda et al., 2006). Parameter reduction also helps improve computational efficiency and faster the calibration process (Muleta and Nicklow, 2005; Knighton et al., 2016).

Sensitivity analysis is a fundamental procedure in hydrological modelling to determine the impact of various input parameters on the model outputs. This analysis is generally categorized into two main approaches: local and global. Local sensitivity analysis changes one input parameter at a time around a predefined baseline to see how it affects outputs (Song et al., 2015). However, this approach has significant limitations, particularly in complex models, where numerous nonlinear relationships exist between input and output variables (Shin et al., 2013; Zhan et al., 2013). The current work used global sensitivity analysis (GSA), which assesses the influence of all parameters simultaneously, thereby providing a more comprehensive understanding of their collective impact on output uncertainty (Baker et al., 2022; Baker et al., 2023).

Event-based calibration, which focuses on distinct rainfall-runoff events, is favoured over continuous method for its flexibility and ability to target informative events, leading to more robust and generalizable model calibration (Hossain et al., 2019; Singh and Bárdossy, 2012). A main challenge in event-based calibration is choosing which events to use for calibration and which for validation. The goal is to select events that reflect the full range of expected conditions, so the model remains accurate and reliable.

The impact of calibration event selection on urban drainage model performance has been subject of previous research (Tscheikner-Gratl et al. 2016; Kleidorfer et al. 2009; Schütze et al. 2002; and Mourad et al. 2005). Tscheikner-Gratl et al. (2016) calibrated water levels in a catchment's outflow pipe using 10 different rain events. In the results of this work, while some events were successfully replicated in calibration, the calibrated models varied in their ability to predict other events and produced different flooding volumes when applied to design storms. Kleidorfer et al. (2009) compared calibration results from the five longest duration events against the five highest peak flow events in combined sewer overflow (CSO) volumes. They found that using the longest duration events could reduce the number of required measurement sites for successful calibration. Schütze et al. (2002) noted that calibrating based on discrete events was time-efficient compared to full time series calibration, but it introduced additional uncertainties. Mourad et al. (2005) demonstrated the sensitivity of stormwater quality model calibration to the selection and number of events used.

Even though the studies cited above bring to the fore the effects of event selection on model calibration, it can also be noted that they have specific limitations in relation to the research methodologies employed. First, it can be noted that they consider only a limited range of generally available options when selecting calibration events. Secondly, they mainly employ the traditional event-based approach. Finally, it is common for these studies to depend on a limited number of events, which may not fully represent the range of scenarios that may be encountered by a model.

In hydrological modelling, manual calibration uses a trial-and-error to adjust parameters by comparing observed and simulated data, through metrics like NSE, and RMSE. While potentially accurate, it is time-consuming and subjective. To overcome these limitations, calibration is shifted to automated, framed as an optimisation problem aiming to determine parameter set that best aligns with specific optimisation objectives. These objective functions quantify how well the model replicates observed catchment behaviours.

Numerous research have explored hydrological models performance under numerous objective functions. Wu and Liu (2014) found that using sum of absolute errors is a viable option that delivers enhanced performance within an automated calibration framework. Fowler et al. (2018) examined eight objective functions across five conceptual model structures, focusing on multi-year droughts in the east and south of Australia. Houshmand Kouchi et al. (2017) conducted a study where they applied three different optimisation algorithms together with five different objective functions within a SWAT model to calibrate monthly discharges in two watersheds in Iran. Moussa and Chahinian (2009) evaluated single and multi-objective functions for a lumped rainfall-runoff model in the Gardon catchment, Southern France, highlighting how parameter calibration heavily depends on the selected criteria. Garcia et al. (2017) assessed the sensitivity and robustness within rainfall-runoff model of an array of objective functions, each one focusing on varying discharge transformations or low-flow indices to amalgamations of such functions.

Even with the availability of the wide-ranging studies cited above, it can still be noted that there is a gap in research specifically paying attention to the estimated parameter values resulting from varying calibration event selection approaches and objective functions. Based on our best knowledge, no previous studies have examined how the interaction between event selection criteria and objective function choice influences both the calibration process and the resulting model behaviour. This gap extends to the influence of these parameters on the interpretation of hydrological processes in the watersheds under study. Understanding how different calibration event selection approaches and objective functions affect the model's calibrated parameters is crucial for ensuring model accuracy and applicability. Researchers have analysed event selection and chosen the objective function separately, but their effects on model calibration, sensitivity and performance together is not fully understood. This understanding not only impacts the model's performance but also informs our comprehension of the hydrological processes within the studied catchments. These parameters can yield insights into the watershed's behaviour, influencing decisions on water management and conservation strategies. This research gap indicates a potential area for further investigation. Delving into how these optimised parameters are derived and their implications for hydrological process interpretation could significantly enhance the effectiveness and applicability of hydrological models.

The primary aim of this research is to understand how different calibration event selection approaches and objective functions influence the calibrated parameters of hydrological models, which is essential for ensuring model accuracy and applicability. To the best of our knowledge, no prior studies in the literature have explored the interaction between event selection criteria and objective functions in hydrological modelling. This study examines various calibration event selection approaches and objective functions in an experimental urban catchment located in the city of Pavia (Northern Italy), using 42 high-resolution rainfall-runoff events with the Stormwater Management Model (SWMM). Additionally, a Global Sensitivity Analysis (GSA) of the drainage model parameters was conducted using the Morris method through a novel integration of Mat-SWMM and the Sensitivity Analysis for Everybody (SAFE) Toolbox. This framework provides a comprehensive approach to assess the impact of different uncertain input parameters on the model's outputs.

2. Case study and data

The Cascina Scala urban catchment, described more deeply in the studies by Barco et al. (2008b) and Todeschini et al. (2019), is located in Pavia, northern Italy (Fig. 1). This catchment, which is primarily used for residential purposes with approcimatly 1500 inhabitants, covers an area of 12.7 ha and is characterized by a slope from northwest to southeast, averaging 0.15 %. The land is connected to the sewer system, with around 62 % of the area being impervious, comprising 22.4 % roofs and 39.6 % streets and paved surfaces. The remaining 38 % of the area is pervious. The sewer system in this catchment is combined and constructed from standard concrete pipes, with a Manning's roughness coefficient n = 0.0147 s $m^{-1/3}$, which is within the typical range for concrete pipes reported in the literature (Davis, 1952). The catchment is divided into 42 subbasins, each drained by a pipe. The catchment design, featuring egg-shaped sewer sections and an average slope of 0.042 m/m, helps maintain high flow velocities and minimizes sediment deposition. The detailed characteristics of each sub-catchment, including slope, area, and ratio of pervious to impervious surfaces, can be found in Giudicianni et al. (2023).

Rainfall events at Cascina Scala area were monitored using two tipping-bucket rain gauges. These gauges possess a sensitivity of 0.2 mm and a funnel area of 1,000 cm^2. Additionally, runoff data at the terminal reach of the sewer network were captured using an ultrasonic depth meter strategically positioned upstream of a Venturi flume until June 2000 and after that, bubble flow meter was used instead of ultrasonic depth meter.

For the rainfall-runoff simulation in this study, a representative subset of rainfall events was meticulously chosen based on specific criteria to ensure the reliability and relevance of the data. These criteria include:

- Reliable Instrumentation: ensuring regular operation of rain gauges and the flow meter was crucial. This means that the selected events must report neither recorded failures in the instrumentation nor occurrences of pressurized flow conditions that could skew the data.
- Minimum Total Rainfall Depth: each event in the subset needed to have a total rainfall depth of at least 5 mm. This criterion ensures that the selected events have enough rainfall to impact the runoff, making them relevant for the study.
- Minimum Rainfall Intensity: the intensity of the rainfall during these events should be equal to or greater than 0.1 mm/min. This

Fig. 1. Cascina Scala experimental catchment, Pavia, Italy.

minimum intensity threshold is important for filtering out lighter rainfall events that might not affect the runoff dynamics in the urban watershed.
- Minimum Rainfall Depth over a Short Duration: The events should exhibit a maximum rainfall depth of at least 2 mm over a 15-minute period.

After reviewing a list of potential events, forty-two were selected as matching the criteria mentioned above. These events are described in detail in Table 1. The steps for selecting and splitting events for calibration and validation are explained in Section 3.3. In a related study, Todeschini et al. (2018) studied the placement strategies and cumulative effects of wet-weather control practices for intermunicipal sewerage systems, adopting for the analysis a continuous series of rainfall events acquired in Pavia from August 2006 to July 2007. The values of precipitation depth and duration for that year are consistent with the corresponding ranges for events reported in the current study. Moreover, in Todeschini et al. (2019), some of the events used in the current study were used for an analysis aimed at testing an innovative first flush identification methodology. The selected subset of storms proved effective in characterizing the dynamics of different types of pollutants in wet-weather runoff. These findings suggest that the rainfall events selected for this study are appropriate for urban catchments and can be used in wider hydrological and water quality studies.

3. Methodology

The methodological framework adopted in this study is depicted in Fig. 2. The first stage is the development of the Hydrology Module and the Hydraulics Module for the case study using the EPA Storm Water Management Model (SWMM). These modules are key parts of the SWMM and are designed to simulate hydrological processes and water movement through the system. The hydrograph is the output from these modules, displaying how wet weather flow or runoff changes over time in response to precipitation events. The methodology's next phase involves determining the influential parameters that have a significant effect on the hydrograph. To achieve this goal, the Morris method of sensitivity analysis methodically changes each parameter to determine its impact on the output of the model. This process is achieved using Mat-SWMM, together with the Sensitivity Analysis for Everybody (SAFE) Toolbox. Mat-SWMM plays the role of the interface linking MATLAB and SWMM to enable an exchange and analysis of data that is simple. Consequently, the SAFE Toolbox is employed to perform the Morris sensitivity analysis. This is an important phase for narrowing

Table 1
Main characteristic of the 42 rainfall events used in the study. Parameters include rainfall duration (minutes), rainfall depth (mm), maximum 5-minute rainfall intensity (mm), peak flow rate (m^3/s), antecedent dry period (days), and hyetograph centre of mass.

Event No.	Rainfall duration (min)	Rainfall depth (mm)	Maximum intensity rainfall over a five-minute (mm)	Peak flow rate (m^3/s)	Antecedent dry period (d)	Centre of Mass
1*	11	12.8	8.8	0.96	4.2	0.68
2	197	11.8	2.7	0.25	0.5	0.27
3**	539	22.6	1.16	0.18	36.2	0.55
4**	108	16.4	4.5	0.55	10.8	0.53
5*	23	15.6	7.5	0.88	4.8	0.46
6*	50	7	1.8	0.33	10	0.28
7*	64	11	3.4	0.38	1.8	0.46
8*	215	10.6	1.3	0.16	0.9	0.53
9*	247	3.8	0.4	0.03	0.3	0.65
10**	443	18.6	1.7	0.26	11	0.55
11*	111	8.4	1	0.16	3.3	0.63
12*	380	15.8	2.3	0.26	0	0.55
13*	100	4.8	0.8	0.11	0.6	0.58
14*	862	18	0.4	0.03	5.7	0.52
15*	964	23.4	0.8	0.13	1.2	0.40
16*	248	12.6	2	0.23	25.9	0.75
17*	231	16.2	1.5	0.21	6.8	0.50
18*	286	7	5.1	0.06	14.2	0.40
19*	666	10.4	0.3	0.03	1.6	0.57
20*	487	19.2	1.64	0.22	0.7	0.41
21**	14	12.8	8.8	0.97	14.1	0.97
22**	308	12.8	1.6	0.13	3.8	0.46
23**	237	16.2	1.5	0.21	6.3	0.55
24	161	11	3.9	0.39	11.7	0.28
25	93	7.8	4	0.27	3.5	0.19
26**	17	5.8	4	0.30	10.0	0.43
27	28	7.4	2.7	0.24	1.0	0.61
28	110	8.4	1	0.16	3.3	0.67
29**	445	8.6	1.8	0.06	14.3	0.40
30	51	10.6	2.3	0.26	2.1	0.64
31	155	16.2	2.8	0.29	2.1	0.43
32	18	25	3.9	0.25	3.8	0.50
33	22	7.4	4.4	0.30	14.8	0.44
34	102	14.4	6.7	0.43	5.9	0.33
35	62	5.8	3.2	0.30	0.2	0.43
36	38	15.2	7.9	0.34	2.5	0.40
37	70	15	6	0.45	7.0	0.23
38	115	7.2	1.4	0.16	30.7	0.34
39	32	8.4	3.1	0.28	18.7	0.56
40	36	21.2	7.4	0.53	0.8	0.58
41	22	14.2	8.4	0.44	10.0	0.72
42	47	14.4	5.8	0.21	4.1	0.49

* event used for calibration stage.
** event used for validation stage.

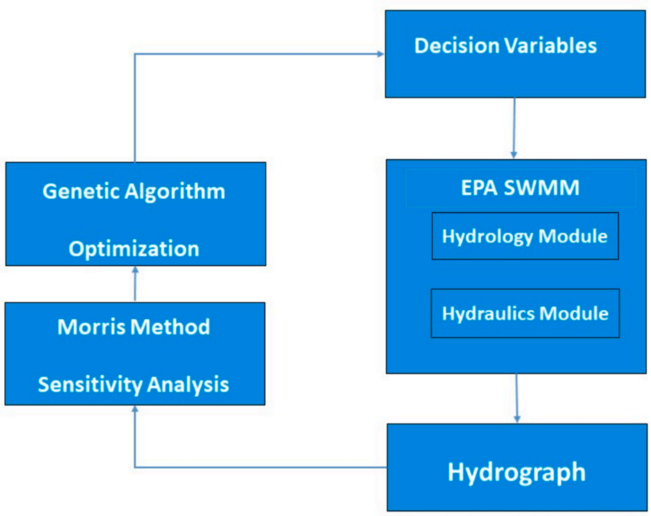

Fig. 2. Research methodology flowchart.

down the parameters that will be calibrated. After the sensitivity analysis, the influential parameters are fine-tuned through an optimisation process employing a Genetic Algorithm (GA), where the GA and SWMM integration is enabled by Mat-SWMM. These parameters are iteratively adjusted by the GA to optimise the performance of the model, with the objective of either maximising or minimising an objective function reflecting the hydrograph prediction's accuracy. Through the process of optimisation, critical parameters are calibrated precisely, which enhances the accuracy and reliability of the model. In this study, the model calibration process is framed as a hydrological model identification problem, in which the goal is to determine a set of parameter values that best reproduces observed hydrological responses under uncertainty. The problem is ill-posed due to equifinality, non-uniqueness, and uncertainties. To address these challenges, we adopt a structured methodology that integrates sensitivity analysis, parameter reduction, and optimisation using high-resolution, long-term data from an experimental catchment.

3.1. EPA-SWMM model description

The EPA-SWMM (Environmental Protection Agency – Storm Water Management Model) is a conceptual Runoff-Route model designed for both rural and urban settings. It is suitable for event-based modelling. It accommodates spatial variability by categorising catchment areas into homogeneous and smaller sub-catchments. The model simulates hydrological processes using three main algorithm groups: surface runoff algorithms, loss algorithms, and conveyance algorithms (Rossman, 2010; Rossman and Huber, 2016).

Surface runoff algorithms generate hydrographs from sub-catchments using nonlinear reservoir routing methods (Huber and Dickinson, 1988). Loss algorithms estimate stormwater losses due to depression storage, infiltration, and evaporation. Depression storage represents the initial losses that accumulate in surface depressions and do not directly contribute to runoff. In this study, evaporation is not taken into account, considering the limited temporal interval characteristic of the event-based modelling. The Horton technique was employed to calculate infiltration losses, where infiltration rate starts high and decreases over time until it reaches a minimal constant rate when the soil gets saturated. Overland flow calculations apply the law of conservation of mass, guaranteeing accurate accounting of water entering and leaving sub-catchments (Rossman and Huber, 2016). Additionally, the conveyance algorithms use dynamic wave routing, which consider flow's continuity and momentum, making a detailed analysis of the stormwater drainage system's flow characteristics available.

In SWMM, the role of the parameters is pivotal in defining the characteristics of the area under study and refining the rainfall-runoff simulations. They are generally divided into two types: inferred and measured parameters. The measured parameters denote the real measurements on the ground, like conduit dimensions and junction elevations. These parameters make quantifiable explicit data about the area being studied available. On the other hand, inferred parameters such as the Manning coefficient and depression storage are not measured directly in the field and introduce uncertainty in the model. Such parameters need careful adjustment to reduce inconsistencies in the rainfall-runoff simulations. Regarding the present study, ten inferred parameters are chosen for parameter estimation process, as shown in Table 2, along with their variability ranges, which are based on Rossman (2010). The sensitivity of these parameters is initially assessed using the Morris method to determine the parameters that are the most influential. Subsequently, a genetic algorithm is used to estimate the parameters.

As detailed by Riaño-Briceño et al. (2016), Mat-SWMM is an open-source, adaptable software package designed to enhance simulation results analysis and system editing. As an interface tool, Mat-SWMM integrated MATLAB functionalities into the SWMM code. It enables users to take advantage of the visualisation and computational prowess of MATLAB together with the simulation features of SWMM. In this study, Mat-SWMM was utilized to combine SWMM with MATLAB's genetic algorithm toolbox.

Mostly used in event-based calibration, the event-by-event approach usually generates unique calibrated parameters for each rainfall event. Nevertheless, this technique has numerous limitations because of its characteristic variability and possible inconsistencies. In most cases, this approach needs post-calibration synthesis, such as averaging, which may fail to accurately represent precise event nuances. Recent developments in modelling tools and automation, such as R package for SWMM, PySWMM, and Mat-SWMM, have led to the adoption of an integrated event calibration approach (Assaf et al. 2024; Giudicianni et al. 2023). This new method, which is the one used in this work, marks a significant advancement in stormwater modelling. It calibrates all events collectively, addressing the fragmentation and inconsistency of the traditional approach and offering a more holistic representation of stormwater dynamics. It also eliminates the need for post-calibration synthesis, enhancing model generalizability and applicability.

3.2. Morris method sensitivity analysis

In this study, sensitivity analysis was employed using the Morris method (Morris, 1991) to evaluate the influence of uncertainties in the input parameters on the rainfall-runoff model's performance. This method is effective for identifying key influential parameters, especially when the number of uncertain parameters is high and/or the model

Table 2
Model parameters selected for calibration process, including their notations and corresponding ranges of variation. The ranges are extracted from Rossman (2010).

Parameter	Notation	Range of variation
Manning coefficient for roof ($s/m^{1/3}$)	n_{roof}	0.008–0.030
Manning coefficient for street ($s/m^{1/3}$)	n_{street}	0.008–0.055
Manning coefficient for the pervious area ($s/m^{1/3}$)	n_{per}	0.10–0.35
Depression storage on impervious area (mm)	S_{imp}	1.00–3.54
Depression storage on pervious area (mm)	S_{per}	2.54–5.08
Maximum infiltration rate (mm/hr)	I_{max}	58–170
Minimum infiltration rate (mm/hr)	I_{min}	1–57
Decay coefficient (1/hr)	decay	2–7
Manning roughness for conduit ($s/m^{1/3}$)	n_{cond}	0.008–0.02
Width coefficient (–)	W_c	1–2

computational time is expensive (Cappato et al. 2022). The Morris method evaluates how individual parameter variations affect model outputs by computing gradients at multiple randomly selected points in the parameter space. This is achieved by altering the value of each parameter (p_i) by a constant quantity (Δ_i) at time along a specific path (trajectory j) in the model's parameter space. After each parameter is altered, a new initial point is selected for further adjustments along a different trajectory. when the model is run with these varying parameter combinations, the impact, known as the elementary effect (*EE*), of each parameter on the selected model output *f* (like total volume or peak flow rate) is computed using the following formula:

$$EE_{pi}(j) = \frac{f(p_1, \cdots, p_i + \Delta, \cdots, p_M) - f(\mathbf{p})}{\Delta} \tag{1}$$

The mean elementary effect (μ) for each parameter is then calculated by averaging these effects, following the methodologies outlined by Morris (1991), Campolongo et al. (2007), Ruano et al. (2012), and Janetti et al. (2019). A larger μ value for a parameter indicates a higher sensitivity of the model output to that parameter. Furthermore, to account for the effects with opposing signs, which could reduce the magnitude of the elementary effect, the absolute mean elementary effect (μ^*) is computed as follows:

$$\mu^*_{pi} = \frac{1}{r}\sum\nolimits_{j=1}^{r} \left|EE_{pi}(j)\right| \tag{2}$$

where r denotes the aggregate number of trajectories. While the Morris method's standard approach uses multiple trajectory sizes, we examined various magnitudes to find out their impact. There is a need for the number of trajectories to strike a balance: such a number should be adequately large enough to produce significant results, lessening variance in μ^*, while still being small enough to ensure efficient time for calculation, given that the model has evaluated $m_e = r(M + 1)$ times, with M representing the total number of input parameters considered in the sensitivity analysis. Our findings show that the employment of $r = 1500$ archives a desirable balance between efficiency and accuracy. This number of trajectories makes it possible to conduct a comparative analysis of the relative impact of each parameter on the output of the model (Reinecke et al., 2019; Campolongo et al., 2007). In addition, the elementary effects (σ) standard deviation is determined with the aim of assessing whether the impact of the parameter on the output of the model is linear or involves interactions with other parameters. When the σ is large compared to the μ^*, it implies that there is a nonlinear interaction or relationship with other parameters (Reinecke et al., 2019; Feng et al., 2019; Morris (1991). Sensitive parameters are identified by computing the absolute elementary effects (μ) of each of the 10 parameters across all 42 rainfall events to ensure a comprehensive assessment of parameter behaviour across a wider range of conditions.

In this study, the sensitivity analysis is performed using an innovative integration of Mat-SWMM and the Sensitivity Analysis for Everybody (SAFE) Toolbox, creating a robust framework for assessing how various uncertain input parameters affect the hydrological model's outputs. Mat-SWMM acts as a bridge between SWMM and MATLAB, enabling smooth data exchange between the two. This integration allows the use of powerful analysis tools within MATLAB, including the SAFE Toolbox. The SAFE Toolbox, developed by Pianosi et al. (2015), is a flexible, standalone tool designed for global sensitivity analysis across different computational environments, including MATLAB. In this study, the SAFE Toolbox is used to systematically assess how changes in various input parameters affect the SWMM model's outputs, specifically runoff volume, peak flow rate, and time to peak flow rate. The innovative integration of Mat-SWMM with the SAFE Toolbox makes this process efficient and effective, enabling a thorough global sensitivity analysis within the SWMM framework. The integration of global sensitivity analysis into SWMM through this method is a novel contribution of the present work, providing an easy and efficient way to manage the complex interactions between model parameters. By combining the strengths of Mat-SWMM and the SAFE Toolbox, this study introduces a novel and powerful method for enhancing the performance of hydrological models. This integration marks a significant step forward in applying global sensitivity analysis to SWMM, offering a more streamlined and effective approach to improving model accuracy and reliability in hydrological predictions.

3.3. Calibration validation process

From the initial pool of 42 rainfall events (Table 1), sixteen events were selected for the model calibration phase based on hydrological and statistical criteria, ensuring representation of the overall rainfall events characteristics, including total rainfall depth, peak flow rate, varying seasons, dry periods and event durations. Four calibration selection strategies were employed, each selecting eight events from 16 events allocated for calibration. For the validation, eight events were chosen from the remaining 26 events and were selected to exhibit similar statistical characteristics to the sixteen events used for calibration. This careful selection ensured consistency in the characteristics of the dataset, allowing for valid and reliable comparison across different calibration strategies. This segregation of events into calibration and validation sets ensures that each calibration strategy was uniformly evaluated against an identical dataset, which ensures that any advantage that may have been gained from the overlapping events in both the calibration and validation sets was excluded.

In the present study, four distinct criteria are employed in calibration event selection: Rainfall Depth (RD), Maximum Rainfall Intensity over a Five-Minute Interval (H5), Mean Intensity (MI), and Hyetograph Centre of Mass (HCM). For each criterion, events were first ranked accordingly, for example, in the RD approach, events were ranked by rainfall depth, whereas in the H5 approach the rank was based on maximum rainfall intensity within a period of five minutes. Events were then alternately selected (1st, 3rd, 5th, and so on) to form the calibration set. This technique results in a dataset that preserves the characteristics defined by the selection criteria defined, which is selected to represent the calibration set for that specific criterion.

The first selection criterion is Rainfall Depth (RD), which ranks the calibration events based on the total accumulated rainfall. This approach ensures that the calibration dataset encompasses a wide range of precipitation events, from low to high rainfall depths, thereby offering a comprehensive representation of the different rainfall scenarios that could influence rainfall-runoff process.

The second selection criterion is Mean Intensity (MI), which ranks calibration events according to the average rainfall intensity over the duration of each event. The incorporation of a range of mean intensity scenarios enables the model to capture a wider spectrum of rainfall patterns, spanning from mild, extended events to more intense, short-duration occurrences. This variability in rainfall scenarios improves the model's capability to simulate and predict diverse hydrological responses, thereby enhancing its overall robustness and accuracy.

The third selection criterion is the Maximum Rainfall Intensity over a Five-Minute Interval (H5), which ranks calibration events according to the highest rainfall intensity observed within any five-minute segment of the event. Including various H5 scenarios ensures the model encounters to a spectrum of peak intensity events, ranging from moderate increases to extreme downpours. Incorporating these scenarios allows the model to effectively account for short-duration, high-intensity rainfall events, which, despite their short occurrence, can significantly influence hydrological responses.

The centre of mass of the hyetograph is the fourth criterion of event selection, which orders events based on coordinates identified as the mean precipitation and the time that equally divides the precipitation volume (ensuring equal volumes on the left and right). Values were standardised to introduce fairness to comparisons among events of various durations and levels of precipitation. This involves categorising all heights based on maximum height and each time by the event's duration. Consequently, each event achieves a variation between 0 and 1 for both precipitation depth and duration. Such standardisation guarantees that both the duration of the event and precipitation depth are assigned equal importance. For instance, an event with a long duration but low precipitation will be comparable to an event with high intensity but short duration.

3.4. Objective functions

Based on an extensive review of the literature (Table 3), the objective functions applicable to studying deviations between simulated and observed values can be methodically categorised into four different groups, each of which is characterised by its distinct methodological approach to quantifying and minimising these deviations:

- Group I – In this group, the objective functions are designed to decrease the squared differences between the simulated and squared values. Squaring the deviations places greater weight on larger errors, making these criteria particularly useful when it is important to emphasize larger discrepancies more than smaller ones.
- Group II – Minimization of Absolute Deviations: In this category, the criteria do not square the deviations, reducing their sensitivity to larger discrepancies. Usually, this approach is employed in instances where the main objective is to reduce the general magnitude of errors without over-emphasising the larger errors.
- Group III: In this category, the objective functions use numerous transformations, including taking the square root or applying logarithmic transformations to both simulated and observed runoff values before the least squares computations are applied. Such a transformation can assist in stabilising the variance and rendering the data more normally distributed, which can give benefits to certain kinds of statistical analysis.
- Group IV – Aggregated Multi-Objective Approaches: multi-objective techniques in modelling offer a significant advantage by enabling the combination of various objective functions within a custom-built framework. These techniques can be employed through the derivation of aggregated or combined objective functions. This approach involves merging multiple objectives into a single composite metric, simplifying the optimisation process while still capturing the essence of multiple criteria.

In this study, at least one representative function of each group has been selected and tested. The selection was based on recommendations derived from the literature and preliminary analyses, which are not reported in this paper. Table 3 presents the objective functions described and the formulation group to which they belong.

Table 3
Description of each objective function tested for model calibration.

Name	Group	Formula	Reference
Squared Difference (SD)	I	$SD = \sum_{i=1}^{N_r}\sum_{j=1}^{N_{t,i}}(Q_{p,i,j} - Q_{m,i,j})^2$	Diskin and Simon (1977)
NSE	I	$NSE = 1 - \frac{\sum_{i=1}^{N_r}\sum_{j=1}^{N_{t,i}}(Q_{p,i,j} - Q_{m,i,j})^2}{\sum_{i=1}^{N_r}\sum_{j=1}^{N_{t,i}}(Q_{m,i,j} - \overline{Q_{m,i,j}})^2}$	Nash and Sutcliffe (1970)
Absolute Error (Abs)	II	$Abs = \sum_{i=1}^{N_r}\sum_{j=1}^{N_{t,i}}\lvert Q_{p,i,j} - Q_{m,i,j}\rvert$	Legates and McCabe (1999)
Index of Agreement (IoA)	II	$IoA = 1 - \frac{\sum_{i=1}^{N_r}\sum_{j=1}^{N_{t,i}}\lvert Q_{p,i,j} - Q_{m,i,j}\rvert}{2\sum_{i=1}^{N_r}\sum_{j=1}^{N_{t,i}}\lvert Q_{m,i,j} - \overline{Q_{m,i,j}}\rvert}$	Willmott et al. (2012)
NSE- sqrt	III	$NSE-sqrt = 1 - \frac{\sum_{i=1}^{N_r}\sum_{j=1}^{N_{t,i}}(\sqrt{Q_{p,i,j}} - \sqrt{Q_{m,i,j}})^2}{\sum_{i=1}^{N_r}\sum_{j=1}^{N_{t,i}}(\sqrt{Q_{m,i,j}} - \overline{\sqrt{Q_{m,i,j}}})^2}$	Chiew et al. (1995)
Volume and Peak (Q&P)	IV	$Q\&P = \sum_{i=1}^{N_r}\left(\frac{Peak_{p,i} - Peak_{m,i}}{Peak_{m,i}}\right)^2 + \left(\frac{Volume_{p,i} - Volume_{m,i}}{Volume_{m,i}}\right)^2$	Barco et al. (2008a)

$N_{t,i}$ and N_r are the number of time instants in the generic *i*-th rain event and the number of rain events considered for optimisation, respectively. $Q_{m,i,j}$ and $Q_{p,i,j}$ are the measured and predicted flow rate, respectively, at the j-th time and in the i-th rain event. $\overline{Q_{m,i,j}}$ is the mean value of the measured flow rate. $Peak_m$ and $Peak_p$ are the measured and predicted peak flow rate, respectively. $Volume_{m,i}$ and $Volume_{p,i}$ are the measured and predicted volume, respectively.

3.5. Goodness-of-fitness

In the present study, the evaluation of the calibration and validation performance involves the employment of goodness-of-fit indices, as recommended by numerous scholars (Legates and McCabe, 1999). For this purpose, two main efficiency criteria were selected: the Coefficient of Efficiency (CE), also known as the Nash-Sutcliffe Efficiency (NSE), and the Root Mean Square Error (RMSE). For these indices, the formulas are as follows:

- Coefficient of Efficiency (CE) is given by the equation:

$$CE = 1 - \frac{\sum_{k=1}^{N_r}(V_{m,k} - V_{p,k})^2}{\sum_{k=1}^{N_r}(V_{m,k} - \overline{V_m})^2} \quad (3)$$

- Root Mean Square Error (RMSE) is calculated as:

$$RMSE = \sqrt{\frac{1}{N_r}\sum_{k=1}^{N_r}(V_{m,k} - V_{p,k})^2} \quad (4)$$

Where $V_{m,k}$ represents the measured variable, $V_{p,k}$ is the predicted variable, $\overline{V_m}$ is the mean of the measured variable and Nr denotes the total number of records.

Equations (3) and (4) (for CE and RMSE, respectively) quantify the model's predictive accuracy by conducting a comparison of the simulated (predicted) data against the observed (measured) data. The CE represents a normalised statistic that determines the comparative magnitude of the residual variance compared to the measured data variance. The RMSE is an available measure of the average magnitude of the error. In this study, CE and RMSE were applied as standard goodness-of-fit measures to evaluate the model's performance after calibration. Although these indices can also serve as optimisation criteria, our focus here is on their role in assessing how well the calibrated model reproduces observed behaviour. To ensure a consistent and comprehensive evaluation, these measures were applied to three relevant hydrological variables: total volume, peak flow rate, and time to peak. This consistent evaluation allows for a fair and comprehensive comparison of model performance across different objective functions and calibration strategies and ensures that all relevant aspects of hydrograph behaviour are assessed.

4. Results and discussion

4.1. Sensitivity analysis results

The Morris method is adopted for conducting sensitivity analysis of SWMM parameters. This method was employed to evaluate the sensitivity of 10 uncertain model parameters: n_{roof}, n_{street}, n_{per}, S_{imp}, S_{per}, I_{max}, I_{min}, decay, n_{cond}, and W_c (as listed in Table 2). The analysis focuses on three hydrologic metrics: the total runoff volume, the peak flow rate and the time to peak flow rate. The focus on both global and instantaneous flow-controlled variables ensures that the analysis captures both the cumulative water volume generated by rainfall events and the maximum instantaneous flow rate. The sensitive parameters are identified by computing the absolute elementary effects μ^* of each of the 10 parameters first and then taking their sum. A threshold is then established at 5 % of this total sum. Parameters with μ^* values exceeding this threshold are classified as sensitive. The analysis is applied event by event across the entire dataset, which includes 42 events.

Without losing generality, we present the full results of sensitivity analysis for a selected event (Figs. 3 and 4) as representative examples. The complete set of sensitivity results across all events is available in the Supplementary Materials (S1–S3), allowing readers to assess the robustness and consistency of sensitivity analysis under various event conditions. Fig. 3 illustrates the convergence behaviour of the absolute value of the mean elementary effect (μ^*) which has been computed for each model parameter across an increasing number of model evaluations (parameter trajectories) for a specific event (event 8). As detailed in Section 3.2, analyses were conducted using 1500 parameter trajectories to ensure the stability of parameter rankings. The plot shows that the μ^* values for all parameters tend to stabilise as the number of model evaluations increases for the three metrics, indicating robust parameter rankings. This convergence analysis underlines the stability of parameter rankings across the model evaluations, thereby validating the reliability of the sensitivity assessment conducted using the Morris method. This finding is also observed when analysing all events in the dataset.

Fig. 4 presents the absolute means and standard deviations of the EEs for the various model parameters concerning the total volume (Fig. 4.a), peak flow rate (Fig. 4.b), and time to peak flow rate (Fig. 4.c). For the total volume, S_{imp} stands out with the highest μ^*, approaching 1200, and a relatively small standard deviation, indicating its substantial influence and consistent impact on the total volume (Fig. 3.a; Fig. 4.a). n_{street} and W_c also show relatively high μ^* values with moderate standard deviations, suggesting their notable effects on total volume. In contrast, the other seven parameters (n_{roof}, n_{cond}, n_{per}, S_{per}, I_{max}, I_{min}, and decay) show low μ^* values, falling below the sensitivity threshold, signifying their minimal impact on this metric. The peak flow rate is most sensitive to n_{street}, followed by W_c, n_{cond}, and n_{roof}, respectively (Fig. 3.b; Fig. 4.b). The peak flow rate metric is insensitive to the S_{imp}, n_{per}, S_{per}, I_{max}, I_{min}, and decay in case of event 9. For the time to peak flow rate, n_{street}, W_c, n_{cond}, n_{roof}, and S_{imp} exceed the sensitivity threshold, highlighting their noteworthy impact on peak flow timing.

To develop a comprehensive evaluation of model parameter sensitivity, the Morris sensitivity analysis method is applied to all events in the dataset. This approach allows the evaluation of the influence of each parameter on the three target metrics across various hydrological scenarios. The results of this sensitivity analysis are summarised in Table S1 for total volume, Table S2 for peak flow rate and Table S3 for time to peak flow rate. These tables present the μ^* values for each parameter across the entire dataset event by event. Parameters that exceed the sensitivity threshold are highlighted in bold, indicating that they are crucial factors in the modelling process.

For the total volume (Table S1), S_{imp} consistently exhibits the highest μ^* across all events, highlighting that it is the most sensitive parameter, significantly influencing the total volume of discharge in the model. The parameters n_{street} and W_C also demonstrate significant sensitivity in various events, frequently exceeding the sensitivity threshold. Conversely, parameters such as n_{per}, I_{max}, I_{min}, and decay generally show low μ^* values, often near zero or less than the sensitivity threshold, indicating their relatively minor impact on the total volume of discharge. The n_{cond} parameter displays moderate sensitivity in some

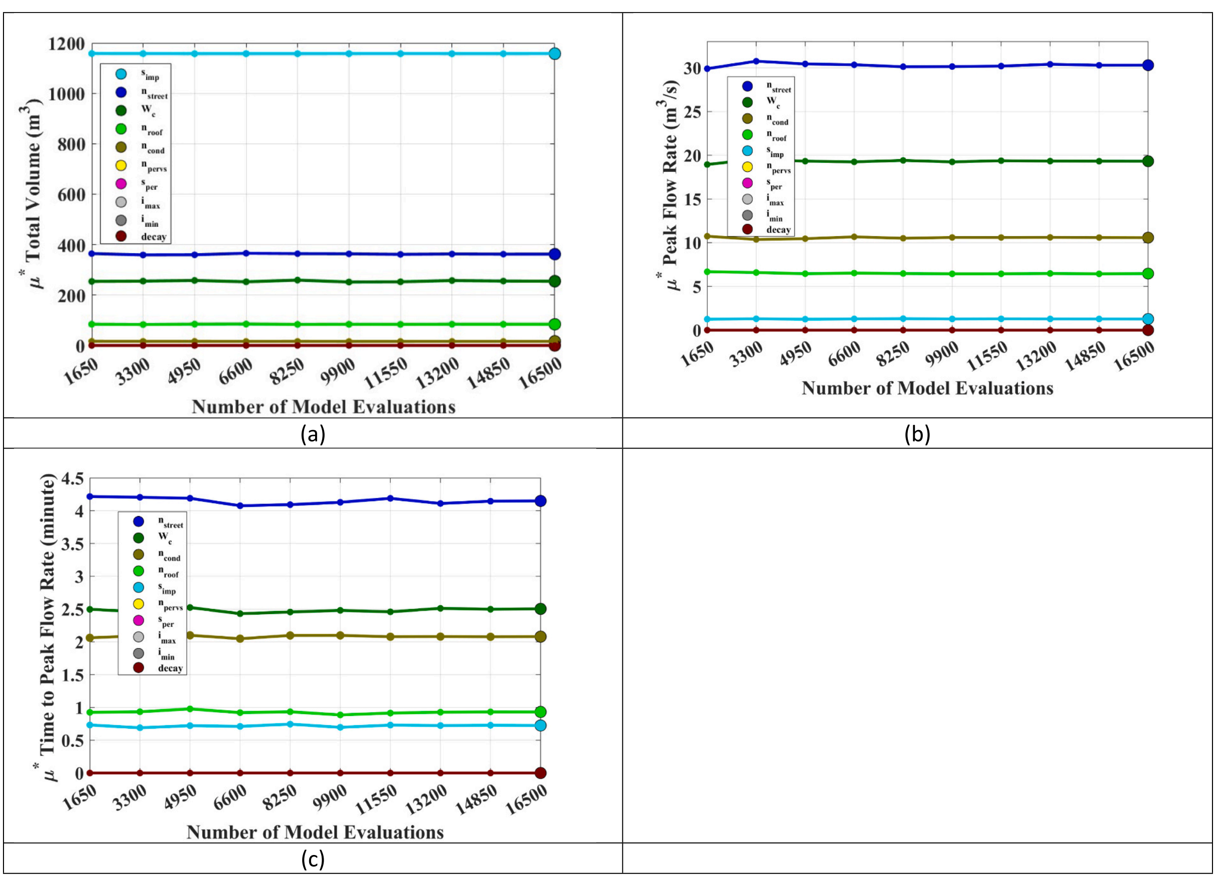

Fig. 3. Convergence of the mean elementary effect absolute value (μ^*) of the 10 model parameters across increasing numbers of model evaluations for event 8. (a) μ^* of the total volume, (b) μ^* of the peak flow rate and (c) μ^* of the time to peak flow rate. The parameters analysed include S_{imp}, n_{street}, W_c, n_{roof}, n_{cond}, n_{per}, S_{per}, I_{max}, I_{min}, and decay, as represented by different coloured markers.

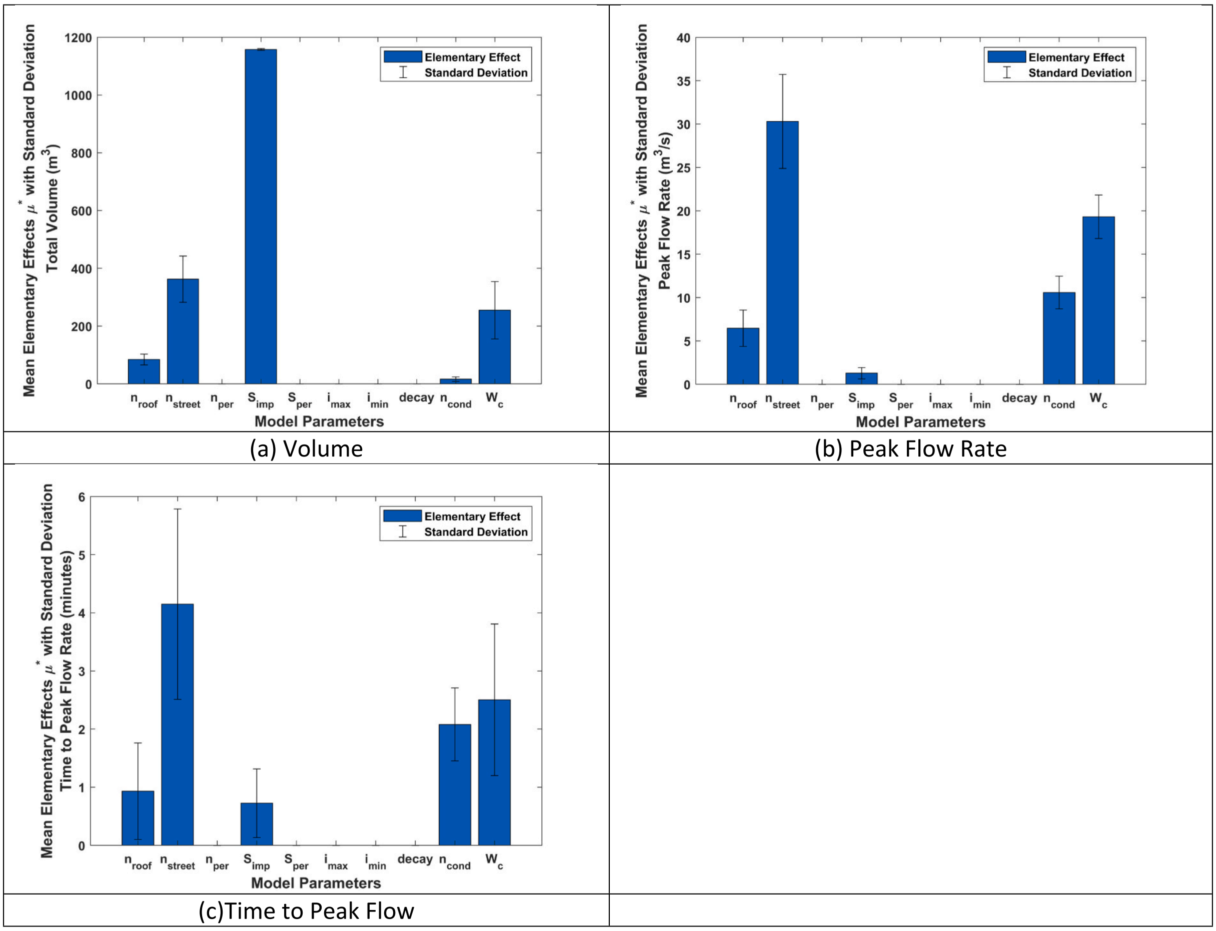

Fig. 4. Mean elementary effects (μ*) with standard deviations for model parameters evaluated for three key hydrological outputs: (a) total runoff volume (m^3), (b) peak flow rate (m^3/s), and (c) time to peak flow (minutes). The parameters include S_{imp}, n_{street}, W_c, n_{roof}, n_{cond}, n_{per}, S_{per}, I_{max}, I_{min}, and decay.

events. For instance, n_{cond} has a μ* value of 714.7 at event 36 and 641.9 at event 27, indicating a notable impact on these events, though it is not as consistently influential as S_{imp} or n_{street}. However, n_{cond} show small μ*, under sensitivity threshold, in several events. These findings indicate that the sensitivity of certain parameters can vary across different events, highlighting the importance of conducting sensitivity analysis for all events in the dataset.

Table S2 presents the mean elementary effect absolute values (μ*) of the ten model parameters for peak flow rate across 42 events in the dataset. The analysis reveals that n_{street} and W_c consistently surpass the sensitivity threshold across all events. Additionally, n_{cond} shows significant sensitivity in many events, although it falls below the sensitivity threshold in a few events. The parameters n_{roof} and S_{imp} also show significant sensitivity in certain events, while they remain insensitive in others. This variability in parameter sensitivity across different events underscores the complex nature of urban hydrological processes. Conversely, parameters like the n_{per}, S_{per}, I_{max}, I_{min} and decay generally exhibit lower μ* values, suggesting they have a lesser impact on peak flow rate.

The sensitivity analysis for time to peak flow rate, as shown in Table S3, shows a similar pattern to the sensitivity results for the peak flow rate. n_{street}, n_{cond} and W_c consistently show significant sensitivity across all events. The parameters S_{imp} and n_{roof} demonstrate variable sensitivity from event to event; while they show moderate sensitivity in several events, they are also insensitive in some instances. Also, the n_{per}, S_{per}, I_{max}, I_{min} and decay show low μ* values, falling below the sensitivity threshold for all events.

Table 4 summarizes the sensitivity of ten model parameters to three target metrics—total volume, peak flow rate, and time to peak flow—across all 42 events.

Table 4
Summary of parameter sensitivity for total volume, peak flow rate, and time to peak flow across 42 events. "Always" indicates consistent significant sensitivity across all events, "Sometime" denotes variable sensitivity from event to event, and "Never" reflects parameters that remain below the sensitivity threshold across all events.

	Total Volume	Peak Flow	Time to Peak Flow
n_{roof}	Never	Sometime	Sometime
n_{street}	Always	Always	Always
n_{per}	Never	Never	Never
S_{imp}	Always	Sometime	Sometime
S_{per}	Never	Never	Never
I_{max}	Never	Never	Never
I_{min}	Never	Never	Never
decay	Never	Never	Never
n_{cond}	Sometime	Sometime	Always
W_c	Always	Always	Always

The sensitivity analysis results for the three metrics provide critical insights into the varying influences of model parameters on urban hydrological modelling. For total volume, S_{imp} emerges as the dominant parameter, with W_c and n_{street} consistently showing significant sensitivity across all events. In contrast, for peak flow rate, n_{street} and W_c are more critical. Some parameters exhibit variability in sensitivity between events, showing μ* values above the sensitivity threshold in some cases and below it in others (e.g., n_{cond} for the total volume metric, and n_{roof}, and S_{imp} for peak flow rate). This variability in parameter sensitivity is more pronounced in the peak flow rate analysis, suggesting that parameters affecting peak flow rate are more influenced by specific event characteristics such as rainfall intensity.

To better illustrate the influence of rainfall characteristics on parameter sensitivity, a summary table has been provided in the supplementary material (Table S4). This table outlines how the sensitivity of SWMM parameters varies with RD, H5, MI, and HCM, based on the results from 42 rainfall-runoff events.

The sensitivity analysis results reveal that model parameter sensitivity varies depending on the target metrics. This finding underscores the importance of conducting sensitivity analyses for different hydrological metrics to gain a comprehensive understanding of the influences of various parameters. Moreover, the analysis indicates that parameter sensitivity is not static but can change with varying external conditions, such as rainfall depth. For example, the sensitivity of infiltration parameters (I_{max}, I_{min} and decay) increases with rainfall depth when total volume and peak flow rate are the metrics. These parameters exhibit zero μ* values during low rainfall depth events, as the rainfall is easily absorbed by the soil in pervious areas, so that the infiltration capacity is not exceeded, and changes in I_{max}, I_{min} and decay don't significantly alter the total volume or peak flow rate. However, when the rainfall depth increases, more water reaches the pervious surfaces, saturating the soil and activating infiltration limitations, causing their EEs to rise as the rainfall depth increases. Additionally, the sensitivity of S_{imp} to peak flow rate also varies with changes in rainfall depth. During low rainfall depth events, a significant portion of rainfall is stored in surface depressions on impervious areas before contributing to runoff making small changes in S_{imp} have a great effect on when runoff begins and the magnitude of peak discharge. Therefore, S_{imp} shows the highest EE among all parameters, as it is the dominant parameter influencing peak flow rate in low depth events. However, as rainfall depth increases, the depression storage reaches capacity quickly, leading to immediate transfer of excess rainfall to runoff, which reduces the S_{imp} relative importance. Instead, W_c and n_{street} take over as the primary controls of peak flow rate at high rainfall depths as these parameters determine how water flows across surfaces and through the drainage system.

Moreover, during events with high H5 (like events 1, 5 and 41), the parameters related to flow conveyance, such as n_{street}, W_c and n_{cond}, increase in sensitivity due to critical role that roughness of street surfaces and conduits plays in determining how quickly and efficiently water is transported.

Overall, the Morris method was used to identify the most influential input parameters affecting various model outputs within their uncertainty ranges. This sensitivity analysis provides a robust framework for evaluating how input parameters contribute to model uncertainty, allowing parameters with minimal impact on model outputs to be assigned standard values, thereby reducing model uncertainty and computational complexity. Previous research has established the Morris method as an optimal sensitivity analysis technique due to its ability to easily remove or substitute failed simulations and its lower computational cost relative to other methods (Campolongo et al., 2007). A comparative study on the computational expense of sensitivity analyses for an energy balance model with similar simulation times to those used in the present paper (1–2 min per run) demonstrated that the Morris method is significantly more cost-effective, requiring only a few tens of runs compared to the hundreds or thousands needed by other methods (Nguyen and Sigrid, 2015).

The findings from the Morris analysis are instrumental in guiding model parameter estimation by identifying parameters to which the model is sensitive and insensitive (Herman et al., 2013). The analysis described in this work indicates that the five parameters (n_{per}, S_{per}, I_{max}, I_{min} and decay) do not significantly affect model outcomes; thus, during calibration their values are fixed in accordance with default values from the SWMM manual (Rossman, 2015) which are respectively 0.15 s/m$^{1/3}$, 2.6 mm, 117 mm/hr, 17 mm/hr, and 5 hr^{-1}, allowing the optimisation efforts to focus on determining the values of the sensitive parameters. It worth noting that these default values are within the bounds covered with the Morris analysis, therefore being eligible as "reference values". Five parameters (S_{imp}, W_c, n_{street}, n_{roof} and n_{cond}) are found to be sensitive based on the sensitivity analysis results for the three metrics across all events. Consequently, these five parameters are considered in the parameter estimation process. As discussed in previous studies, reducing the number of calibration parameters is a well-established strategy to mitigate equifinality (Beven, 2006; Beven and Binley, 1992). By limiting the calibration to the most sensitive parameters, the model becomes better constrained, reducing non-uniqueness and improving the likelihood of identifying parameter sets that are both statistically valid and hydrologically meaningful. This is particularly critical in the context of ill-posed inverse problems, where the number of observations is often insufficient to support the calibration of a large parameter set (Jakeman and Hornberger, 1993; Kuczera and Mroczkowski, 1998). Moreover, focusing on influential parameters helps enhance the efficiency and interpretability of the calibration process, while improving the model's predictive reliability across different hydrological conditions. In line with the broader concept of model identification, this strategy strengthens the correspondence between observed system behaviour and model representation, thus contributing to a more robust and defensible modelling framework under uncertainty (Wagener and Hoshin, 2005; Her and Chaubey, 2015).

A second sensitivity analysis was conducted for each of the six objective functions used in the study. The Morris method was applied to assess how changes in each parameter influence the objective function's value and to evaluate how parameter sensitivity varies across different functions. This analysis aims to ensure that the objective functions are sensitive to the parameters being optimised during calibration process, so that the differences observed in the calibrated parameter values are a direct result of the objective function characteristics, rather than being influenced by insensitive or irrelevant parameters. The analysis is applied on an event by event across the entire dataset, which includes 42 events, and a summary of results is presented in Table 5.

The results confirm that five parameters (S_{imp}, n_{street}, Wc, n_{cond}, and n_{roof}) were consistently identified as sensitive across all six objective functions. This result serves as a validation of the parameter selection process. Since the same set of influential parameters was used to calibrate the model under each objective function, any differences observed in the resulting parameter values or model outputs can be more likely attributed to the intrinsic characteristics of the objective functions.

4.2. Parameter optimisation results

The parameter estimation process focuses on the five identified sensitive parameters (S_{imp}, W_c, n_{street}, n_{roof} and n_{cond}). This section presents the results of the optimisation process applied to the calibration sets, highlighting how different calibration event selection criteria and objective functions impact the estimation of these key parameters. Fig. 5 provides a comprehensive comparison of 24 different optimised values for each parameter, resulting from the combination of four calibration event selection criteria (RD, H5, MI, and HCM) and six objective functions (SD, Abs, NSE, NSE-sqrt, Q&P, and IoA).

The results for n_{roof} (Fig. 5.a) highlight the variation in calibrated values across different combinations of event selection approaches and objective functions. The n_{roof} values range from 0.0080 to 0.029 s/m$^{1/3}$, with the RD approach tending to produce higher values across most objective functions. Fig. 5.b shows the results for n_{street}, which indicate variation depending on the calibration event selection approach. The RD approach consistently produces higher n_{street} values across most objective functions. In contrast, the H5, MI, and HCM approaches yield lower and more consistent n_{street} values.

Fig. 5.c presents the optimisation results for S_{imp}, which indicate substantial variation in S_{imp} values depending on the combination of calibration approach and objective function. The S_{imp} values range from 1.1 to 3.5 mm, highlighting the critical and combined role of calibration

Table 5
Sensitivity analysis of the six objective functions to 10 SWMM parameters, across 42 events. The numbers represent the count of events in which each parameter exceeds the sensitivity threshold, with the corresponding percentage of total events shown in parentheses.

	n_{roof}	n_{street}	n_{per}	S_{imp}	S_{per}	I_{max}	I_{min}	decay	n_{cond}	W_c
SD	38 (90.4 %)	42 (100 %)	0	42 (100 %)	0	0	0	0	42 (100 %)	39 (92.9 %)
Abs	35 (83.3 %)	42 (100 %)	0	42 (100 %)	0	0	0	0	40 (95.2 %)	42 (100 %)
NSE	37 (88.1 %)	42 (100 %)	0	42 (100 %)	0	0	0	0	42 (100 %)	41 (97.6 %)
NSE-sqrt	36 (85.7 %)	41 (97.6 %)	0	42 (100 %)	0	0	0	0	41 (97.6 %)	40 (95.2 %)
Q&P	29 (70.7 %)	42 (100 %)	0	42 (100 %)	0	0	0	0	40 (95.2 %)	42 (100 %)
IoA	31 (73.8 %)	42 (100 %)	0	42 (100 %)	0	0	0	0	40 (95.2 %)	41 (97.6 %)

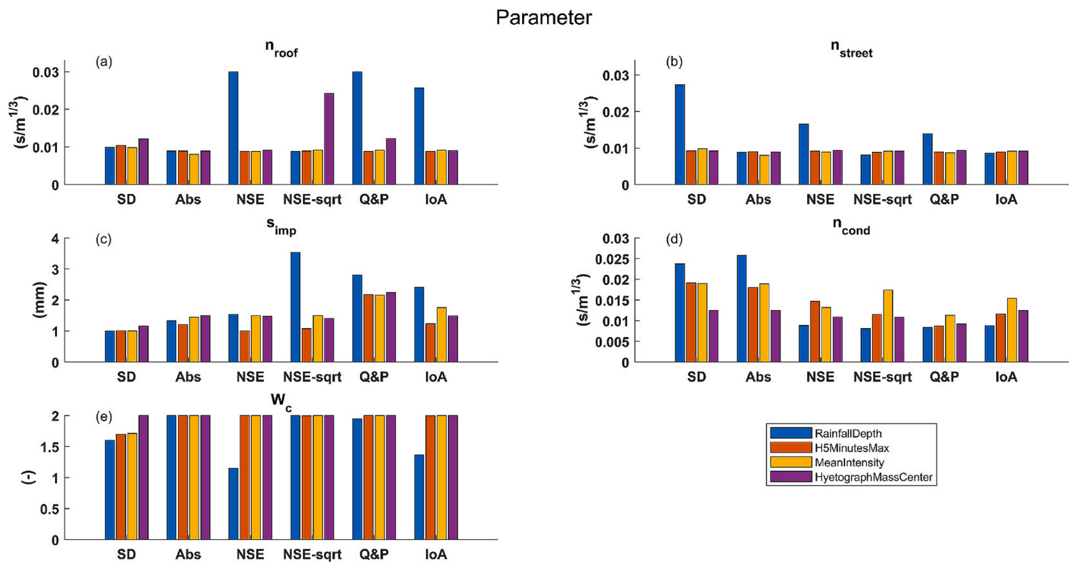

Fig. 5. Parameter estimation results for the five sensitive parameters identified in the study: (a) n_{roof}, (b) n_{street}, (c) S_{imp}, (d) n_{cond} and (e) W_c. These results are based on the optimisation of the calibration set across different combinations of calibration event selection approaches: RD (blue), H5 (red), MI (yellow) and HCM (purple) and objective functions (SD, Abs, NSE, NSE-sqrt, Q&P, and IoA).

event selection approaches and objective functions in influencing the estimation of this key parameter, which has been identified as having the highest influence on total volume. Fig. 5.d underscores the significant influence of both calibration event selection approaches and objective functions on the estimation of n_{cond}. The SD and Abs objective functions tend to predict higher roughness values; however, the results from these objective functions do not show consistency over the different selection criteria.

The W_c results (Fig. 5.e) show that the RD approach generally produces lower W_c values, while the MI, H5, and HCM approaches often push the W_c value to its upper limit of 2.0. Notably, it appears that the optimiser pushes the W_c values toward this upper limit in many cases. However, this upper limit cannot be increased due to the physical meaning of W_c, which is constrained to a range between 1.0 and 2.0 (refer to Section 3.1). These results underscore the significant impact of both the calibration event selection approaches and the objective functions on the estimation of these key parameters, which directly influence the hydrological predictions of the SWMM model.

Fig. 6 presents the normalised parameter values across different calibration event selection approaches and objective functions. These normalised values are obtained by dividing each parameter's value by the upper limit of its corresponding range, resulting in values that range from 0 to 1. This normalisation facilitates the comparison of different parameters on a common scale. The 32 different parameter sets displayed in Fig. 6 demonstrate the sensitivity of SWMM model parameters to the calibration event selection approach and objective functions used.

The results show that the choice of the calibration event selection method greatly impacts the consistency of the calibrated parameter sets within the same objective function. Specifically, when using the H5, MI, and HCM approaches, the resulting parameter sets tend to exhibit less variability across the same objective functions (except for n_{roof} and, to some extent, for n_{street}). In contrast, the RD approach leads to significantly different parameter sets for the same objective functions compared to the H5, MI, and HCM approaches. In other words, while the H5, MI, and HCM approaches generally produce similar parameter sets within a given objective function, the RD approach introduces more variability, resulting in greater differences in the calibrated parameter sets. This difference can likely be attributed to the nature of the RD criterion, which can include both long, low-intensity storms and short, intense storms, introducing greater hydrological heterogeneity. This variability may activate different runoff generation mechanisms, leading to more diverse sensitivity patterns in model parameters, especially in urban drainage systems where storage, and surface runoff respond non-linearly to rainfall intensity and temporal distribution.

However, it should be noted that the relationships between the parameters are complex and non-linear, meaning that even small changes in the parameter set can lead to significant changes in model performance (Fatone et al., 2021; Peng et al., 2023). These nonlinear interactions can cause similar parameter sets to produce different model performance metrics when applied under different calibration approaches. A comprehensive comparison of model performance is presented in Section 4.3.

A cross-performance analysis between objective functions was conducted to evaluate how a parameter set optimised for one objective function performs when evaluated under alternative objective functions. The aim of this analysis is to determine if the differences in the parameter calibrated values obtained in this study come from the intrinsic characteristics of each objective function, associated with the

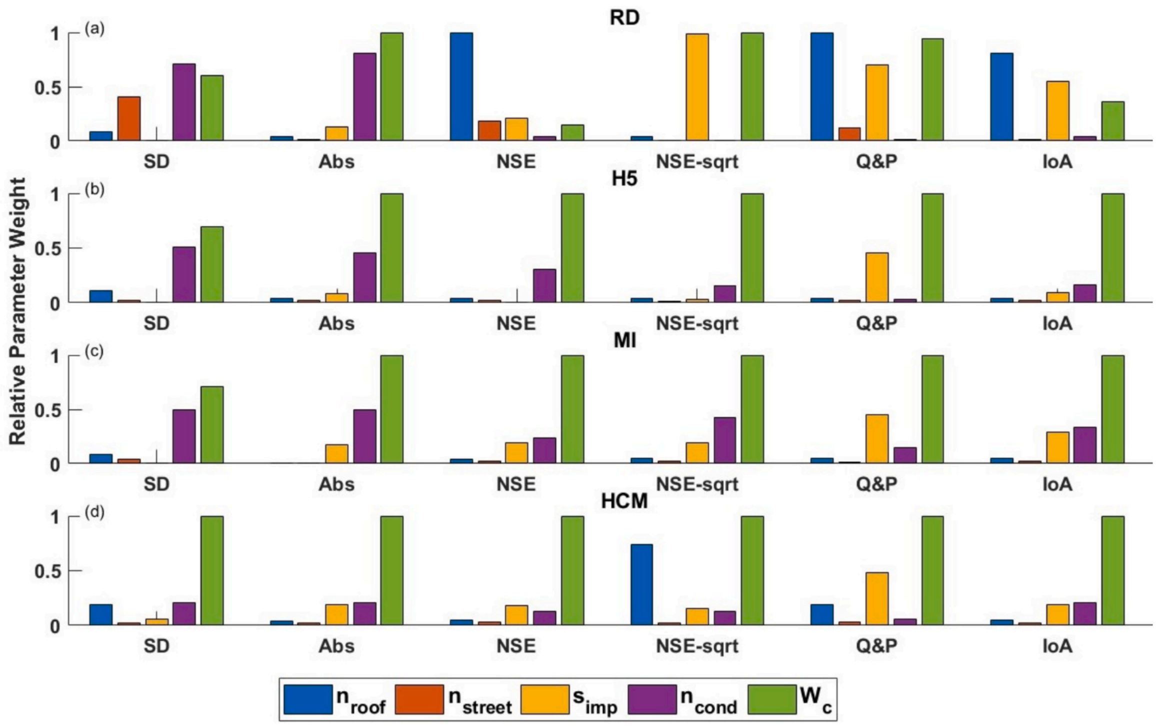

Fig. 6. Normalised parameter values across different calibration event selection approaches and objective functions. The figure is divided into four panels (a, b, c, d), each corresponding to one of the calibration event selection approaches: Rainfall Depth (panel a), H5MinutesMax (panel b), Mean Intensity (panel c), and Hyetograph Mass Center (panel d). Within each panel, the normalised parameter values of the five sensitive parameters—n_{roof} (blue), n_{street} (orange), S_{imp} (yellow), n_{cond} (purple), and W_c (green)—are shown across eight objective functions: SD, Abs, NSE, NSE-sqrt, Q&P, and IoA.

random patterns in the optimisation, or weak identifiability in the optimisation. In this analysis, parameter sets optimised for one objective function (e.g., SD, Abs, etc.) are evaluated on all other objective functions. This analysis was applied to the four selection criteria (RD, H5, MI, and HCM). The performance of each optimised parameter set was normalised with respect to its own self-evaluation. For a parameter set optimised using objective function f_i, its performance when evaluated under another objective function f_j was normalised using the following expression:

$$N_{ij} = \frac{f_j(p_i)}{f_i(p_i)}$$

where $f_j(p_i)$ is the performance value of the parameter set p_i (optimised

RD

	SD	Abs	IoA	NSE	NSE_sqrt	Q&P
SD	**1.00**	1.21	2.16	1.76	3.28	2.16
Abs	1.18	**1.00**	1.19	1.16	1.34	1.18
IoA	1.31	1.20	**1.00**	1.20	1.23	1.19
NSE	1.32	1.22	1.21	**1.00**	1.51	1.27
NSE_sqrt	1.50	1.32	1.46	1.34	**1.00**	1.63
Q&P	1.78	1.68	1.17	1.38	1.29	**1.00**

MI

	SD	Abs	IoA	NSE	NSE_sqrt	Q&P
SD	**1.00**	1.17	1.16	1.15	1.35	1.43
Abs	1.16	**1.00**	1.16	1.15	1.19	1.20
IoA	1.18	1.26	**1.00**	1.17	1.16	1.16
NSE	1.25	1.35	1.20	**1.00**	1.21	1.24
NSE_sqrt	1.21	1.25	1.15	1.15	**1.00**	1.18
Q&P	1.40	2.10	1.13	1.22	1.19	**1.00**

HMC

	SD	Abs	IoA	NSE	NSE_sqrt	Q&P
SD	**1.00**	1.19	1.18	1.18	1.23	1.32
Abs	1.16	**1.00**	1.20	1.20	1.16	1.18
IoA	1.20	1.19	**1.00**	1.22	1.19	1.29
NSE	1.14	1.27	1.16	**1.00**	1.20	1.75
NSE_sqrt	1.18	1.19	1.16	1.21	**1.00**	1.54
Q&P	1.45	1.20	1.42	1.45	1.50	**1.00**

H5

	SD	Abs	IoA	NSE	NSE_sqrt	Q&P
SD	**1.00**	1.29	1.62	1.34	1.79	2.09
Abs	1.17	**1.00**	1.23	1.17	1.25	1.32
IoA	1.35	1.27	**1.00**	1.21	1.20	1.31
NSE	1.23	1.13	1.43	**1.00**	1.19	1.55
NSE_sqrt	1.30	1.19	1.38	1.16	**1.00**	1.46
Q&P	1.73	1.53	1.25	1.58	1.46	**1.00**

Fig. 7. Normalised cross-objective function performance matrices for parameter sets optimised under six different objective functions (SD, Abs, IoA, NSE, NSE_sqrt, and Q&P), evaluated using four event selection criteria: RD, MI, HCM, and H5. Each cell shows the normalised performance value, with diagonal values set to 1.00. Off-diagonal values indicate the relative performance of a parameter set optimised for one objective function when evaluated under another.

for f_i) evaluated under objective function f_j and $f_i(p_i)$ is the original performance of the same parameter set under its own objective function.

Each matrix in Fig. 7 presents the normalised performance of optimised parameter sets for a specific event selection criterion when evaluated under different objective functions. Each diagonal element (values in bold) is set to 1.00, as it represents the self-evaluated performance for the corresponding function. The off-diagonal values indicate the relative performance of each optimised parameter set when tested against different objective functions.

The findings indicate that the performance of parameter sets optimised for one function degrades when evaluated under an alternative function. This degradation ranged from 13 % to approximately 200 % in some cases. This finding indicates that optimisation is function-specific, and one parameter set does not perform equivalently across different objectives. The observed variability in parameter values is thus more likely due the function-specific nature of the calibration process rather than equifinality or random effects introduced by the optimisation algorithm.

4.3. Model performance

The comparison of performance for the four calibration selection criteria (RD, H5, MI, and HCM) and six objective functions (SD, Abs, NSE, NSE-sqrt, Q&P, and IoA) is conducted using three key hydrological metrics: total volume, peak flow, and time to peak. Fig. 8 and Fig. 9 present the calibration and validation results for total volume, peak flow rate, and time to peak, using the CE and RMSE, respectively.

The calibration results for total volume (Fig. 8.a and 9.a) show that the calibration approach based on RD demonstrates the highest CE values (ranging from 0.77 to 0.88) and lowest RMSE (ranging from 113.7 m^3 to 93.8 m^3) across all objective functions. The RD approach achieves CE values of 0.86 and 0.88 for NSE-sqrt and IoA, respectively. The HCM approach also provides a high performance across all objective functions, especially for NSE-sqrt (CE = 0.82 and RMSE = 118.1 m^3) and IoA (CE = 0.82 and RMSE = 116.2 m^3). The H5 approach provides a moderate performance, with CE values ranging from 0.45 to 0.79 and RMSE values ranging from 185.5 to 109.2 m^3. In contrast, the MI approach typically produces the lowest CE values and the highest RMSE values, especially for the SD (CE = 0.43 and RMSE = 274.6 m^3), Abs (CE = 0.48 and RMSE = 246.8 m^3), and NSE (CE = 0.52 and RMSE = 178.7 m^3). This indicates that MI approach is less effective at accurately modelling the total volume compared to the other methods.

The validation results for total volume are presented in Fig. 8.b and 9.b using CE and RMSE, respectively. The results indicate that the RD approach, which performs well during the calibration stage, achieves the highest CE values and lowest RMSE values across several objective functions, particularly in NSE-sqrt (CE = 0.77 and RMSE = 138.4 m^3) and IoA (CE = 0.77 and RMSE = 136.3 m^3). These findings suggest that the RD approach is the most effective in capturing the total volume of runoff events. The H5 approach shows moderate performance during validation with CE values ranging from 0.43 to 0.72 and RMSE values ranging from 194.6 to 146.4 m^3. A similar trend is observed with the HCM approach, where CE values range from 0.44 to 0.6 and RMSE values from 192.2 to 162 m^3. The MI approach, again, demonstrates lower performance across most objective functions, with CE values as low as 0.31 and RMSE reaching up to 229 m^3 in the Abs objective function. Overall, the validation results are consistent with the calibration performance. The RD approach proves to be the most reliable approach access calibration and validation steps, especially with objective functions like NSE-sqrt and IoA, which show high CE values and low RMSE values.

Further analysis on the objective functions reveals distinct performance patterns across the different calibration event selection approaches. The objective functions of group II (SD and NSE) show variable effectiveness across different approaches, especially during the calibration phase, with CE values ranging from 0.43 for MI approach to 0.82 for RD approach. This finding indicates that SD and NSE are particularly sensitive to the choice of the calibration event selection approach. The Abs objective function follows a similar trend; it shows the highest CE value of 0.79 with RD and the lowest value at a CE value

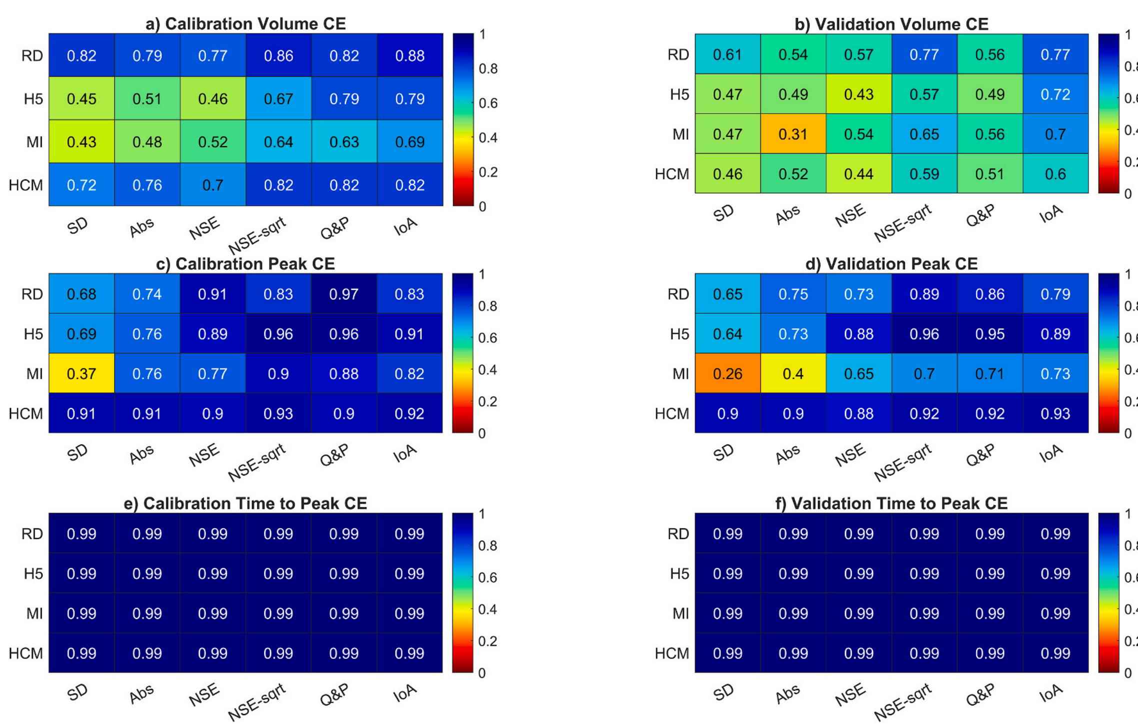

Fig. 8. The heatmap visualizes the performance metrics, where each cell's colour intensity corresponds to the CE value, following a gradient from 0 (red) to 1 (blue), as depicted by the colour bar to the right. The rows of the heatmap represent four distinct calibration event selection approaches: RD, H5, MI, and HCM. The columns signify the six objective functions: SD, Abs, NSE, NSE-sqrt, Q&P,and IoA.

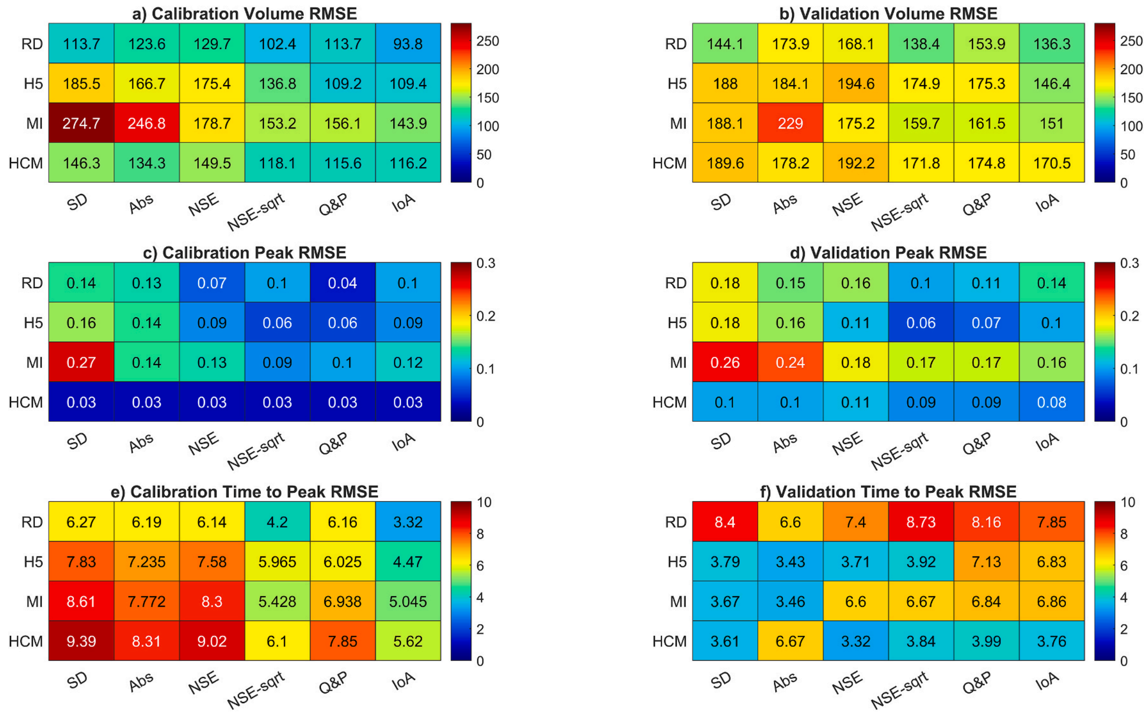

Fig. 9. The heatmap visualizes the performance metrics, where each cell's colour intensity corresponds to the RMSE value, following a gradient from 0 (red) to 1 (blue), as depicted by the colour bar to the right. The rows of the heatmap represent four distinct calibration event selection approaches: RD, H5, MI, and HCM. The columns signify the six objective functions: SD, Abs, NSE, NSE-sqrt, Q&P, and IoA.

of 0.48 with MI during the calibration phase. During validation, Abs shows CE values ranging from 0.31 for MI to 0.54 for RD. The Q&P objective function demonstrates good performance during the calibration and moderate performance during the validation, with stable performance across different calibration selection approaches. The NSE-sqrt and IoA objective functions provide high performance, especially with RD (CE of NSE-sqrt = 0.86 and CE of IoA = 0.89) and HCM (CE = 0.82 for both) during the calibration phase, and with RD (CE = 0.77 for both) during the validation phase. However, NSE-sqrt and IoA demonstrate a significant variable level of performance during calibration (CE ranging from 0.64 to 0.86 for NSE-sqrt, and from 0.69 to 0.88 for IoA) and validation (CE ranging from 0.57 to 0.77 for NSE-sqrt, and from 0.6 to 0.77 for IoA). This outcome signifies that NSE-sqrt and IoA are sensitive to the choice of calibration event selection approach. However, this sensitivity is less pronounced compared to that observed for SD, NSE, and Abs.

Fig. 8.c and 9.c present the calibration results for peak flow rate. The analysis shows that the HCM approach consistently achieves high performance across all objective functions, with CE values ranging from 0.9 to 0.92 and RMSE values of 0.03 m^3/s for all objective functions. This indicates that the HCM approach allows effective prediction of the peak flow rates during calibration. The RD approach also performs well in NSE (with CE = 0.91 and RMSE = 0.07 m^3/s) and Q&P (with CE = 0.97 and RMSE = 0.04 m^3/s).

The H5 approach shows consistent and robust performance, with CE values reaching 0.96 for NSE-sqrt and Q&P, indicating high predictive accuracy. In contrast, MI approach has the lowest performance among the four approaches, particularly in SD (CE = 0.37 and RMSR = 0.27 m^3/s), Abs (CE = 0.76 and RMSE = 0.14 m^3/s) and NSE (CE = 0.77 and RMSE = 0.13 m^3/s). This indicates its relative ineffectiveness in accurately simulating peak flow rates during calibration. Although it shows moderate performance in some functions such as NSE-sqrt (CE = 0.9), overall, the MI approach lags behind the other approaches.

The validation results for peak flow rate (Fig. 8.d and 9.d) indicate that HCM approach, which also provides a well performance during the calibration, exhibits the highest performance across all objective functions with CE values ranging from 0.88 to and 0.93 and RMSE values ranging from 0.11 to 0.08 m^3/s. This suggests that HCM is highly effective in accurately predicting peak flow rates. The H5 approach also demonstrates robust performance, particularly in Q&P, NSE-sqrt, and IoA, where CE values reach 0.92, 0.92, and 0.93, respectively.

Conversely, the MI approach shows lower performance, particularly in SD (CE = 0.26 and RMSE = 0.26 m^3/s) and Abs (CE = 0.40 and RMSE = 0.24 m^3/s), suggesting challenges in accurately predicting peak flows during validation. However, MI performs moderately in functions like NSE-sqrt (CE = 0.7) and IoA (CE = 0.73), indicating some level of predictive capability. The RD approach shows moderate to high performance overall, with CE values ranging from 0.65 to 0.89 and RMSE values ranging from 0.18 to 0.10 m^3/s.

Examining the performance of different objective functions across the different calibration event selection approaches reveals that SD and Abs objective functions show a wide range of CE values from 0.26 for SD and 0.4 for Abs with MI to 0.90 for both functions with HCM. This high variability performance suggests that SD and Abs are highly sensitive to the choice of calibration approach. The similar pattern can be detected with NSE but with less variability in CE values that range from 0.65 to 0.88. The three objective functions (NSE-sqrt, Q&P, and IoA) consistently show superior performance (CE above 0.85) across different approaches, except for MI approach. The IoA demonstrates the highest performance across calibration approaches with CE ranging from 0.73 to 0.93. The high performance of IoA indicates that it is reliable and effective in accurately capturing the peak flow rate magnitude over different calibration selection approaches.

Fig. 8.e and 9.e present the calibration results for time to peak flow rate. The results show uniformly high performance across all approaches and objective functions, each achieving a CE value of 0.99. The results show that the RMSE values for time to peak flow rate are consistently low across all approaches and objective functions, indicating highly accurate prediction of peak flow timing. The highest RMSE value of 9.39 min, observed for the HCM approach with the SD objective function, is

still good given the long events duration. This demonstrates that even the highest RMSE values reflect strong predictive accuracy. This outstanding performance is also maintained in the validation results (Fig. 8.f and 9.f), where all combinations achieve CE value of 0.99 and RMSE values range from 8.73 to 3.32 min. These validation results further confirm the model's robustness and reliability in accurately simulating peak flow rate timing. This consistently good performance suggests that the SWMM model is highly effective at predicting the time to peak flow rate, regardless of the calibration event selection approach or objective function used. Also, this shows that the time to peak flow rate is less sensitive to the choice of calibration approach compared to other metrics like total volume or peak flow rate. The SWMM model robustness in estimating the time to peak can also be detected by the narrow range of change in μ^* values for time to peak for all events (Table S3). This suggests that even with wide changes in parameter values, the model's prediction of the time to peak will remain largely unaffected. This narrow change in μ^* values indicates that the time to peak is not sensitive to changes in the parameters, which indicates that the model is accurate in predicting this metric.

The comparison between the model optimisation results (Figs. 5 and 6) and the model performance (Figs. 8 and 9) reveals that different parameter sets can yield good model performance. This outcome is due to the inherently nonlinear nature of hydrological systems. In a nonlinear system, different parameter sets can interact with model inputs in complex ways, leading to similar outputs (Kouchi et al., (2017); Beven and Binley (1992)). This conclusion aligns with the concept of equifinality stated by Beven and Freer (2001), which suggests that various parameter sets can effectively simulate catchment characteristics. Similar findings have been observed in previous studies, such as Kouchi et al. (2017), who used three different optimisation algorithms and eight different objective functions to create calibrated parameter sets in two watersheds in Iran, Muleta (2012), who evaluated the performance of nine objective functions, and Moussa and Chahinian (2009), who compared the performance of various single and multi-objective functions on a catchment in Southern France.

The results show that RD approach is effective in providing an accurate estimation of total volume. The RD approach guarantees that the calibration dataset encompasses events with a wide range of precipitation magnitudes, which is important as the total volume is related to the amount of rainfall during the event. This diversity of light and heavy rainfall events allows the model to be calibrated under a wide variety of conditions, thereby enhancing its ability to accurately predict total runoff volume across different scenarios. For the peak flow rate estimation, the H5 approach demonstrates an accurate performance. This can be attributed to the fact that peak flow rate is highly sensitive to the intensity of rainfall over short durations, as the high intensity can cause rapid increases in runoff, which lead to peaks in flow rate. The H5 criterion includes a diverse array of peak intensity events in the calibration dataset, making the model better prepared to predict peak flow rates under various conditions. RD proved to be the most reliable selection criteria, offering the best performance for total volume estimation while also delivering reasonable accuracy in peak flow rate predictions.

The findings present a strong case for advancing beyond traditional objective functions like the SD or NSE. As mentioned in section 1, both SD and NSE are highly sensitive to large deviations because they square the differences between observed and predicted values, inherently giving disproportionate weights to larger errors. This makes them particularly vulnerable to outliers or significant errors. If a model prediction deviates significantly from the observed value, this deviation can disproportionately impact the overall score, potentially resulting in lower performance metrics. Also, this characteristic makes SD and NSE more sensitive to the specific characteristics of the calibration events, which are determined by the calibration event selection approach. In contrast, NSE-sqrt and IoA are designed to be less sensitive to extreme deviations between predicted and observed values. NSE-sqrt achieves this through the square root transformation, while IoA uses absolute differences. These approaches cause IoA and NSE-sqrt to reduce the disproportionate impact of large errors and outliers, which provides a balanced evaluation of model performance across the entire data range. As a result, NSE-sqrt and IoA demonstrate more consistent performance across different calibration event selection approaches.

Based on the discussed results, we can detect that those objective functions like NSE-sqrt and IoA, which consider instantaneous measurements at each time step, can give a better estimation for total volume and peak flow rate than objective functions that directly target the total volume and peak flow rate like Q&P objective function. This recommendation may initially seem confusing for some readers, as the objective function is typically selected to align closely with the model's primary goals. Consequently, many modellers might assume, albeit incorrectly, that optimising the same metric during the calibration period will maximize its performance during the validation. However, achieving robust simulation performance under varying conditions, relies on the accurate representation of underlying processes. Therefore, the calibration objective functions should be selected based on their ability to extract process-relevant information from the calibration data. This discussion underscores the importance of differentiating between calibration objective function(s) and modelling objectives.

4.4. Application and limitation

The calibration of urban hydrological models such as SWMM has traditionally relied on selecting a limited number of rainfall-runoff events based on subjective criteria or expert knowledge. These selections are usually made by event magnitude (e.g., largest events, highest peak events), without considering how well they represent key features of hydrological events (Kleidorfer et al. 2009; Mourad et al. 2005; Broekhuizen et al. 2020). The selection criteria for these approaches are often not clear and may not reflect the full range of hydrological variations under different rainfall conditions. In contrast, our work suggests a data-driven and replicable method for choosing events, evaluating four criteria, RD, H5, MI and HCM, to see how they affect parameter estimation. Most studies tend to use traditional SD or NSE metrics when selecting objective functions, without taking into account how sensitive these metrics are to various event characteristics. Our approach uses six objective functions and shows how their performance changes based on which events and calibration targets are chosen, showing that model performance can significantly change depending on the pairing between selection strategy and objective function.

In addition, most existing studies either skip sensitivity analysis, test with a small number of events or analyse only a single hydrological metric (Niazi et al., 2017). These approaches do not show how parameters change in different hydrological situations. Our integration of GSA using the Morris method, executed through an innovative coupling of Mat-SWMM and the SAFE toolbox, shows that parameter sensitivity varies substantially across rainfall types and target metrics.

The proposed methodology, combining event selection strategies, multi-objective calibration, and global sensitivity analysis, is designed to be applicable across a wide range of rainfall conditions, including high-intensity storm events. Its structure is inherently flexible and computationally efficient, making it suitable for broader applications in urban watershed modelling. We encourage future studies to apply and evaluate this framework under more extreme rainfall scenarios to further validate its robustness, particularly for flood-focused analyses and stormwater infrastructure design.

Although this study is based on rainfall-runoff events from a single urban catchment located in a Mediterranean climate, the methodology we propose is broadly applicable and not limited by these contextual factors. The modelling framework introduced in this study (statistical event selection criteria, objective function evaluation, and global sensitivity analysis), is designed to be modular and scalable. Each component is structured in a way that allows independent adaptation, making the overall approach flexible for application across various urban catchment configurations. The event selection strategy is based on generalizable statistical criteria, and the multi-objective calibration approach can accommodate diverse performance goals. Additionally, the integration of open-source Mat-SWMM with SAFE toolbox facilitates efficient, automated sensitivity analysis, supporting large-scale simulation workflows.

The use of a well-characterized, relatively small residential catchment provides a controlled environment to develop and test the methodology under realistic urban hydrological conditions. This type of catchment is common in many cities globally and has been widely used in previous modelling studies (Broekhuizen et al., 2020; Liu et al., 2015; Zhu et al., 2019; Tamm et al..2023; Yang et al., 2023; Broekhuizen et al., 2021). Furthermore, although the selected events represent a diverse range of rainfall characteristics, including moderate and prolonged events, the proposed methodology is applicable and scalable to high-intensity storm events as well. We encourage future research to test the framework under more extreme rainfall conditions to further evaluate its robustness in flood-focused studies.

However, the successful application of this framework to larger and more complex urban catchments requires the availability of monitoring data. In particular, uniform and reliable runoff measurements and rainfall data across multiple locations are crucial to ensure accurate calibration and model interpretation. The acquired rainfall data should accurately describe the temporal and spatial variability of precipitation in the catchment. The available runoff data must have a fine discretisation to allow an exhaustive description of the experimental hydrographs and therefore a reliable evaluation of relevant hydrological variables, i.e. total volume, peak flow rate and time to peak. While the methodology is not inherently limited by catchment size, its effectiveness in more spatially heterogeneous settings will depend on the extent and resolution of the input data available. Despite these challenges, the core methodology remains applicable and offers a systematic, reproducible approach for parameter estimation and sensitivity analysis. Future research is encouraged to explore and refine the framework in larger-scale applications to assess its performance under more complex urban hydrological settings.

5. Conclusion

This study presents a comprehensive analysis of the interaction between calibration event selection criteria and objective functions for modelling rainfall-runoff in urban catchment area. Four event selection criteria (RD, H5, MI and HCM) and six objective functions (SD, Abs, NSE, NSE-sqrt, Q&P and IoA,) were tested.

The research conducts a GSA of SWMM model parameters using Morris method, introducing an innovative integration of Mat-SWMM and the Sensitivity Analysis for Everybody (SAFE) Toolbox. Using a comprehensive dataset of 42 high-resolution rainfall events, this work investigates how various rainfall characteristics impact the sensitivity of 10 uncertain model parameters. The analysis focuses on three hydrologic metrics: the total runoff volume, the peak flow rate and the time to peak flow rate. The results from the GSA showed that the sensitivity of model parameters is highly dependent on the characteristics of each rainfall event, highlighting considerable differences in parameter influence based on event type. Also, the results reveal that model parameter sensitivity varies depending on the target metrics. The analysis identified five influential parameters, which were then optimised using a genetic algorithm.

The results show that the RD approach is effective in providing accurate estimates of total volume. For the peak flow rate estimation, the H5 and HCM approaches demonstrate, instead, an accurate performance. These findings offer practical guidance for urban hydrologists and water managers: the RD method is recommended when accurate total volume estimation is prioritized, while H5 and HCM approaches are more appropriate for capturing peak flow dynamics.

The findings highlight the limitations of traditional objective functions like SD and NSE, which tend to emphasize large errors due to squaring differences, making them sensitive to outliers and specific calibration event characteristics. In contrast, NSE-sqrt and IoA reduce this sensitivity, NSE-sqrt by applying a square root transformation and IoA by using absolute differences, resulting in a more balanced evaluation of model performance across the dataset. Consequently, NSE-sqrt and IoA show more stable performance across various calibration event selection approaches.

CRediT authorship contribution statement

Mohammed N. Assaf: Software, Formal analysis, Writing – original draft, Methodology, Conceptualization. **Nicolò Salis:** Software, Formal analysis, Writing – review & editing, Methodology. **Sauro Manenti:** Supervision, Conceptualization, Writing – review & editing, Methodology. **Lorenzo Tamellini:** Supervision, Funding acquisition, Writing – review & editing, Methodology. **Enrico Creaco:** Writing – review & editing, Conceptualization, Funding acquisition. **Sara Todeschini:** Supervision, Funding acquisition, Writing – review & editing, Methodology, Conceptualization.

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.

Acknowledgements

This research was supported by the Italian Ministry of University and Research (MUR) through the project "PRIN 2020-URCA!, Prot. 20208TAK3H" and through the project "DORIAN of DICAr" - MUR Program "Dipartimento di Eccellenza 23-27". Lorenzo Tamellini, Sara Todeschini, Sauro Manenti, and Mohammed N. Assaf have been supported by the PRIN 2022 PNRR project "Uncertainty Quantification of coupled models for water flow and contaminant transport" (No. P2022LXLYY), financed by the European Union - Next Generation EU.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2025.133821.

Data availability

Data will be made available on request.

References

Assaf, M.N., Manenti, S., Creaco, E., Giudicianni, C., Tamellini, L., Todeschini, S., 2024. New optimization strategies for SWMM modeling of stormwater quality applications in urban area. J. Environ. Manage. 361, 121244. https://doi.org/10.1016/j.jenvman.2024.121244.
Baker, E.A., Cappato, A., Todeschini, S., Tamellini, L., Sangalli, G., Reali, A., Manenti, S., 2022. Combining the Morris method and multiple error metrics to assess aquifer characteristics and recharge in the lower Ticino Basin, in Italy. J. Hydrol. 614, 128536. https://doi.org/10.1016/j.jhydrol.2022.128536.
Baker, E.A., Manenti, S., Reali, A., Sangalli, G., Tamellini, L., Todeschini, S., 2023. Combining noisy well data and expert knowledge in a Bayesian calibration of a flow model under uncertainties: an application to solute transport in the Ticino basin. GEM-International Journal on Geomathematics 14 (1), 8. https://doi.org/10.1007/s13137-023-00219-8.
Balha, A., Singh, A., Pandey, S., Kumar, R., Mallick, J., Singh, C.K., 2023. Assessing the impact of land-use dynamics to predict the changes in hydrological variables using effective impervious area (EIA). Water Resour. Manag. 37 (10), 3999–4014.
Barco, J., Wong, K.M., Stenstrom, M.K., 2008a. Automatic calibration of the US EPA SWMM model for a large urban catchment. J. Hydraul. Eng. 134 (4), 466–474.
Barco, J., Papiri, S., Stenstrom, M.K., 2008b. First flush in a combined sewer system. Chemosphere 71 (5), 827–833.

Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrol. Process. 6 (3), 279–298. https://doi.org/10.1002/hyp.3360060305.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249 (1–4), 11–29. https://doi.org/10.1016/S0022-1694(01)00421-8.

Beven, Keith J. *Rainfall-runoff modelling: the primer*. John Wiley & Sons, 2012.

Beven, K., 2006. A manifesto for the equifinality thesis. J. Hydrol. 320 (1–2), 18–36.

Beven, K., 2001. How far can we go in distributed hydrological modelling? Hydrol. Earth Syst. Sci. 5 (1), 1–12.

Blöschl, G., Sivapalan, M., 1995. Scale issues in hydrological modelling: a review. Hydrol. Process. 9 (3–4), 251–290.

Broekhuizen, I., Leonhardt, G., Viklander, M., 2021. Reducing uncertainties in urban drainage models by explicitly accounting for timing errors in objective functions. Urban Water J. 18 (9), 740–749.

Broekhuizen, I., Leonhardt, G., Marsalek, J., Viklander, M., 2020. Event selection and two-stage approach for calibrating models of green urban drainage systems. Hydrol. Earth Syst. Sci. 24 (2), 869–885.

Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environ. Model. Software 22 (10), 1509–1518. https://doi.org/10.1016/j.envsoft.2006.10.004.

Cappato, A., Baker, E.A., Reali, A., Todeschini, S., Manenti, S., 2022. The role of modeling scheme and model input factors uncertainty in the analysis and mitigation of backwater induced urban flood-risk. J. Hydrol. 614, 128545. https://doi.org/10.1016/j.jhydrol.2022.128545.

Chiew, F.H.S., Whetton, P., McMahon, T., Pittock, A.B., 1995. Simulation of the impacts of climate change on runoff and soil moisture in Australian catchments. J. Hydrol. 167 (1–4), 121–147. https://doi.org/10.1016/0022-1694(94)02649-V.

Davis C.V. 1952. Handbook of Applied Hydraulics, Mcgraw-hill Book Company Inc., New York.

Diskin, M.H., Simon, E., 1977. A procedure for the selection of objective functions for hydrologic simulation models. J. Hydrol. 34 (1–2), 129–149. https://doi.org/10.1016/0022-1694(77)90066-X.

Fatone, F., Szeląg, B., Kiczko, A., Majerek, D., Majewska, M., Drewnowski, J., Łagód, G., 2021. Advanced sensitivity analysis of the impact of the temporal distribution and intensity of rainfall on hydrograph parameters in urban catchments. Hydrol. Earth Syst. Sci. 25 (10), 5493–5516. https://doi.org/10.5194/hess-25-5493-2021.

Feng, K., Lu, Z., Yang, C., 2019. Enhanced Morris method for global sensitivity analysis: good proxy of Sobol'index. Struct. Multidiscip. Optim. 59 (2), 373–387.

Fowler, K., Peel, M., Western, A., Zhang, L., 2018. Improved rainfall-runoff calibration for drying climate: Choice of objective function. Water Resour. Res. 54 (5), 3392–3408.

Garcia, F., Folton, N., Oudin, L., 2017. Which objective function to calibrate rainfall–runoff models for low-flow index simulations? Hydrol. Sci. J. 62 (7), 1149–1166.

Giudicianni, C., Assaf, M.N., Todeschini, S., Creaco, E., 2023. Comparison of nonlinear reservoir and UH algorithms for the hydrological modeling of a real urban catchment with EPASWMM. Hydrology 10 (1), 24. https://doi.org/10.3390/hydrology10010024.

Granata, F., Gargano, R., De Marinis, G., 2016. Support vector regression for rainfall-runoff modeling in urban drainage: a comparison with the EPA's storm water management model. Water 8 (3), 69.

Hamel, P., Tim, D.F., 2014. The impact of stormwater source-control strategies on the (low) flow regime of urban catchments. Water Sci. Technol. 69 (4), 739–745.

Hatt, B.E., Fletcher, T.D., Walsh, C.J., Taylor, S.L., 2004. The influence of urban density and drainage infrastructure on the concentrations and loads of pollutants in small streams. Environ. Manag. 34, 112–124.

Her, Y., Chaubey, I., 2015. Impact of the numbers of observations and calibration parameters on equifinality, model performance, and output and parameter uncertainty. Hydrol. Process. 29 (19), 4220–4237.

Herman, J.D., Kollat, J.B., Reed, P.M., Wagener, T., 2013. Method of Morris effectively reduces the computational demands of global sensitivity analysis for distributed watershed models. Hydrol. Earth Syst. Sci. 17 (7), 2893–2903.

Hornberger, G.M., Beven, K.J., Cosby, B.J., Sappington, D.E., 1985. Shenandoah watershed study: calibration of a topography-based, variable contributing area hydrological model to a small forested catchment. Water Resour. Res. 21 (12), 1841–1850.

Hossain, S., Hewa, G.A., Wella-Hewage, S., 2019. A comparison of continuous and event-based rainfall–runoff (RR) modelling using EPA-SWMM. Water 11 (3), 611.

Kouchi, H., Delaram, K.E., Faridhosseini, A., Sanaeinejad, S.H., Khalili, D., Abbaspour, K. C., 2017. Sensitivity of calibrated parameters and water resource estimates on different objective functions and optimization algorithms. Water 9 (6), 384.

Huber, W.C.; Dickinson, R.E. Storm Water Management Model User's Manual; Version 4.0; U.S. Environmental Protection Agency: Cincinnati, OH, USA, 1988.

Jakeman, A.J., Hornberger, G.M., 1993. How much complexity is warranted in a rainfall-runoff model? Water Resour. Res. 29 (8), 2637–2649.

Janetti, E.B., Guadagnini, L., Riva, M., Guadagnini, A., 2019. Global sensitivity analyses of multiple conceptual models with uncertain parameters driving groundwater flow in a regional-scale sedimentary aquifer. J. Hydrol. 544–556. https://doi.org/10.1016/j.jhydrol.2019.04.035.

Azar, K., Morteza, C.G., Creaco, E., 2025. Sensitivity analysis-aided calibration of urban drainage Modeling. Water 17 (5), 612.

Kleidorfer, M., Möderl, M., Fach, S., Rauch, W., 2009. Optimization of measurement campaigns for calibration of a conceptual sewer model. Water Sci. Technol. 59, 1523–1530. https://doi.org/10.2166/wst.2009.154.

Knighton, J., Lennon, E., Bastidas, L., White, E., 2016. Stormwater detention system parameter sensitivity and uncertainty analysis using SWMM. J. Hydrol. Eng. 21 (8), 05016014.

Kuczera, G., Mroczkowski, M., 1998. Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. Water Resour. Res. 34 (6), 1481–1489.

Legates, D.R., McCabe Jr, G.J., 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Water Resour. Res. 35 (1), 233–241. https://doi.org/10.1029/1998WR900018.

Liu, L., Y. Liu, X. Wang, Dapeng Yu, K. Liu, H. Huang, and G. Hu. "Developing an effective 2-D urban flood inundation model for city emergency management based on cellular automata." Natural hazards and earth system sciences 15, no. 3 (2015): 381-391.

Mantovan, P., Todini, E., 2006. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. J. Hydrol. 330 (1–2), 368–381.

Moradkhani, Hamid, and Soroosh Sorooshian. *General review of rainfall-runoff modeling: model calibration, data assimilation, and uncertainty analysis*. Springer Berlin Heidelberg, 2008.

Moreda, F., Koren, V., Zhang, Z., Reed, S., Smith, M., 2006. Parameterization of distributed hydrological models: learning from the experiences of lumped modeling. J. Hydrol. 320 (1–2), 218–237.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33 (2), 161–174.

Mourad, M., Bertrand-Krajewski, J.-L., Chebbo, G., 2005. Stormwater quality models: sensitivity to calibration data. Water Sci. Technol. 52, 61–68. https://doi.org/10.2166/wst.2005.0110.

Moussa, R., Chahinian, N., 2009. Comparison of different multi-objective calibration criteria using a conceptual rainfall-runoff model of flood events. Hydrol. Earth Syst. Sci. 13 (4), 519–535.

Muleta, M.K., 2012. Model performance sensitivity to objective function during automated calibrations. J. Hydrol. Eng. 17 (6), 756–767. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000497.

Muleta, M.K., Nicklow, J.W., 2005. Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. J. Hydrol. 306 (1-4), 127–145.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. J. Hydrol. 10 (3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Nguyen, A.-T., Sigrid, R., 2015. A performance comparison of sensitivity analysis methods for building energy models. In: *Build. Simul.*, 8 Tsinghua University Press, pp. 651–664.

Niazi, M., Nietch, C., Maghrebi, M., Jackson, N., Bennett, B.R., Tryby, M., Massoudieh, A., 2017. Storm water management model: performance review and gap analysis. J. Sustain. Water Built Environ. 3 (2), 04017002. https://doi.org/10.1061/JSWBAY.0000817.

Peng, J., Zhao, H., Luo, Z., Ouyang, J., Yu, L., 2023. A new approach for sensitivity analysis of the stormwater management model applied in an airport. Water Science & Technology 88 (9), 2453–2464. https://doi.org/10.2166/wst.2023.335.

Pianosi, F., Sarrazin, F., Wagener, T., 2015. A matlab toolbox for global sensitivity analysis. Environ. Model. Software 70, 80–85.

Rahimi, A.Y., Ebrahimian, A., 2024. Urban impervious cover characterization for rainfall-runoff analysis: Impacts of data availability and resolution. J. Hydrol. 641, 131807.

Reinecke, R., Foglia, L., Mehl, S., Herman, J.D., Wachholz, A., Trautmann, T., Döll, P., 2019. Spatially distributed sensitivity of simulated global groundwater heads and flows to hydraulic conductivity, groundwater recharge, and surface water body parameterization. Hydrol. Earth Syst. Sci. 23 (11), 4561–4582.

Riaño-Briceño, G., Barreiro-Gomez, J., Ramirez-Jaime, A., Quijano, N., Ocampo-Martínez, C., 2016. MatSWMM–an open-source toolbox for designing real-time control of urban drainage systems. Environ. Modell. Software 83, 143–154.

Rossman, L.A. Storm Water Management Model User's Manual; Version 5.0; U.S. Environmental Protection Agency: Cincinnati, OH, USA, 2010.

Rossman, L.A. Storm Water Management Model User's Manual; Version 5.1; U.S. Environmental Protection Agency: Cincinnati, OH, USA, 2015.

Rossman, L.A., Huber, W., 2016. Storm water management model reference manual volume III–water quality. US EPA Office of Research and Development, Washington, DC. EPA/600/R-16/093.

Ruano, M.V., Ribes, J., Seco, A., Ferrer, J., 2012. An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors. Environ. Model. Software 37 (2012), 103–109. https://doi.org/10.1016/j.envsoft.2012.03.008.

Schütze, M., Willems, P., and Vaes, G.: Integrated Simulation of Urban Wastewater Systems – How Many Rainfall Data Do We Need?, in: Global Solutions for Urban Drainage, American Society of Civil Engineers, Lloyd Center Doubletree Hotel, Portland, Oregon, United States, 1–11, Doi: 10.1061/40644(2002)244, 2002.

Shin, M.-J., Guillaume, J.HA., Croke, B.FW., Jakeman, A.J., 2013. Addressing ten questions about conceptual rainfall–runoff models with global sensitivity analyses in R. J. Hydrol. 503, 135–152.

Singh, S.K., Bárdossy, A., 2012. Calibration of hydrological models on hydrologically unusual events. Adv. Water Resour. 38, 81–91.

Sivapalan, Murugesu, K. Takeuchi, S. W. Franks, V. K. Gupta, H. Karambiri, V. Lakshmi, X. Liang et al. "IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences." *Hydrological sciences journal* 48, no. 6 (2003): 857-880.

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., Chonggang, Xu., 2015. Global sensitivity analysis in hydrological modeling: review of concepts, methods, theoretical framework, and applications. J. Hydrol. 523, 739–757.

Sorooshian, S., Gupta, V.K., 1983. Automatic calibration of conceptual rainfall-runoff models: the question of parameter observability and uniqueness. Water Resour. Res. 19 (1), 260–268.
Tamm, O., Kokkonen, T., Warsta, L., Dubovik, M., Koivusalo, H., 2023. Modelling urban stormwater management changes using SWMM and convection-permitting climate simulations in cold areas. J. Hydrol. 622, 129656.
Todeschini, S., Manenti, S., Creaco, E., 2019. Testing an innovative first flush identification methodology against field data from an Italian catchment. J. Environ. Manage. 246, 418–425. https://doi.org/10.1016/j.jenvman.2019.06.007.
Todeschini, S., Papiri, S., Ciaponi, C., 2018. Placement strategies and cumulative effects of wet-weather control practices for intermunicipal sewerage systems. Water Resour. Manag. 32 (8), 2885–2900. https://doi.org/10.1007/s11269-018-1964-y.
Tscheikner-Gratl, F., Zeisl, P., Kinzel, C., Leimgruber, J., Ertl, T., Rauch, W., Kleidorfer, M., 2016. Lost in calibration: why people still do not calibrate their models, and why they still should – a case study from urban drainage modelling. Water Sci. Technol. 74, 2337–2348. https://doi.org/10.2166/wst.2016.395.
Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H.V., Sorooshian, S., 2001. A framework for development and application of hydrological models. Hydrol. Earth Syst. Sci. 5 (1), 13–26.
Wagener, T., Hoshin, V.G., 2005. Model identification for hydrological forecasting under uncertainty. Stochastic Environ. Res. Risk Assess. 19, 378–387.
Walsh, C.J., Roy, A.H., Feminella, J.W., Cottingham, P.D., Groffman, P.M., Morgan, R.P., 2005. The urban stream syndrome: current knowledge and the search for a cure. J. N. Am. Benthol. Soc. 24 (3), 706–723.
Wang, K.-H., Altunkaynak, A., 2012. Comparative case study of rainfall-runoff modeling between SWMM and fuzzy logic approach. J. Hydrol. Eng. 17 (2), 283–291.
Willmott, C.J., Robeson, S.M., Matsuura, K., 2012. A refined index of model performance. Int. J. Climatol. 32 (13), 2088–2094.
Wu, Y., Dunxian She, Xia, J., Zhang, Y., Lei, Z., 2024. Evaluation of the number of events' influence on model performance and uncertainty in urban data-scarce areas based on behavioral parameter ranking method. J. Hydrol. 636, 131298.
Wu, Y., Liu, S., 2014. A suggestion for computing objective function in model calibration. Eco. Inform. 24, 107–111.
Yang, B., Zhang, T., Li, J., Feng, P., Miao, Y., 2023. Optimal designs of LID based on LID experiments and SWMM for a small-scale community in Tianjin, north China. J. Environ. Manage. 334, 117442.
Zhan, C.-S., Song, X.-M., Xia, J., Tong, C., 2013. An efficient integrated approach for global sensitivity analysis of hydrological model parameters. Environ. Model. Software 41, 39–52.
Zhu, Z., Chen, Z., Chen, X., Guo, Yu., 2019. An assessment of the hydrologic effectiveness of low impact development (LID) practices for managing runoff with different objectives. J. Environ. Manage. 231, 504–514.