Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Assessing the Reliability of ANN-Assisted Multiobjective Optimization for Urban Low Impact Development Design

Abraão Jhonny da Costa Brazão

abraao.brazao@ifrn.edu.br

Federal University of Paraíba

Gustavo Barbosa Lima da Silva

Federal University of Paraíba

Research Article

Keywords: Green infrastructure, Urban stormwater management, Low impact development, Artificial neural networks, Multiobjective optimization, Decision support

Posted Date: November 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11446856/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Assessing the Reliability of ANN-Assisted Multiobjective Optimization for Urban Low Impact Development Design

Abraão Jhonny da Costa Brazão [a, *], Gustavo Barbosa Lima da Silva[b]

[a] Technology Center, Federal University of Paraíba, 58051-900 João Pessoa, Paraíba, Brazil; abraao.brazao@ifrn.edu.br

[b] Department of Civil and Environmental Engineering, Federal University of Paraíba, 58051-900 João Pessoa, PB, Brazil; gustavo.silva@academico.ufpb.br

* Corresponding author

Abstract: The design of low impact development (LID) systems requires balancing stormwater control performance, implementation costs, and numerous spatial and design configurations. Although simulation-based multiobjective optimization can support these decisions, its high computational demand may limit the exploration of complex LID design spaces. In this context, surrogate models, such as artificial neural networks (ANNs), provide a promising alternative, although their prediction errors may affect Pareto fronts and subsequent design decisions. Therefore, this study evaluated the reliability of an ANN-assisted multiobjective optimization framework for LID design by comparing it with direct optimization using the Storm Water Management Model (SWMM). Multilayer perceptron models were developed to predict peak outflow (POF) and total outfall volume (TOV) from rainfall characteristics and LID design variables. The ANNs were coupled with NSGA-II to minimize POF, TOV, and LID implementation and maintenance costs for 5- and 25-year design storms. Nondominated surrogate-assisted solutions were subsequently re-evaluated using SWMM, and their dominance relationships were reassessed. The ANN models achieved high overall predictive performance, with NSE values of 0.996 and 0.998 for POF and TOV, respectively, using 50,000 training samples. However, prediction accuracy decreased within the high-performance region explored during optimization, particularly for POF. Despite these

local errors, the ANNs successfully guided NSGA-II toward promising regions. High-fidelity re-evaluation improved agreement with the direct SWMM-based Pareto fronts, reducing the Inverted Generational Distance by 38.9% and 28.0% for the 5- and 25-year scenarios, respectively. The surrogate-assisted approach also preserved the main LID design patterns. The complete workflow reduced computational time by 46%, while the optimization stage alone achieved an 86% reduction. These findings demonstrate that ANN surrogates combined with targeted high-fidelity re-evaluation provide an efficient and reliable decision-support framework for exploring hydrological and economic trade-offs in urban LID planning.

Keywords: Green infrastructure; Urban stormwater management; Low impact development; Artificial neural networks; Multiobjective optimization; Decision support.

1. Introduction

The increase in impervious areas resulting from urbanization, combined with extreme rainfall events and population growth, has significantly altered the urban hydrological cycle in many cities worldwide, leading to increased runoff volumes, higher peak flows, reduced time to peak, and elevated pollution levels in urban streams (Bibi et al. 2023; Ionno et al. 2024; Seemuangngam et al. 2024). In response to these challenges, Low Impact Development (LID) practices have been increasingly adopted as sustainable stormwater management strategies. Techniques such as green roofs (GR), permeable pavements (PP), and bioretention cells (BT) aim to restore pre-development hydrological conditions by enhancing infiltration, retention, and evapotranspiration processes at the source (Eckart et al. 2018; Ekmekcioglu et al. 2021; Li et al. 2019). Their effectiveness in reducing urban runoff and improving water quality has been widely demonstrated (Sabbagh et al. 2025; Sui and Ven 2023; Yang et al. 2020).

Despite these benefits, the design of LID systems remains challenging because of the large number of possible combinations of LID types, locations, dimensions, and implementation areas. Multiobjective optimization has therefore been increasingly used to explore these large decision spaces and identify trade-offs among conflicting objectives, such as implementation cost, peak flow reduction, and runoff volume reduction (Eckart et al. 2018; Yang et al. 2023). Evolutionary and metaheuristic algorithms, including the Non-dominated Sorting Genetic Algorithm II (NSGA-II), particle swarm optimization, simulated annealing, and pattern search, have been successfully applied to LID allocation and design problems (Eckart et al. 2018; Yang et al. 2023).

In this context, optimization frameworks frequently incorporate physically based hydrological models to simulate system responses under different LID implementation scenarios, enabling the evaluation of multiple design alternatives and thereby characterizing the simulation-based optimization framework (Farina et al. 2023; Garzon et al. 2022). The Storm Water Management Model (SWMM), developed by the United States Environmental Protection Agency, is widely used for this purpose because it explicitly represents urban rainfall-runoff processes, drainage-system hydraulics, and LID controls (Chiu et al. 2025; Lopes and Silva 2021). However, physically based optimization approaches are often time consuming and computationally intensive, as they rely on hydraulic principles to solve the Saint-Venant equations using numerical methods and require a large number of objective function evaluations (Li et al. 2023; Yang et al. 2023; Zhang et al. 2024). For example, Lopes and Silva (2021) coupled SWMM with an adapted NSGA-II for LID optimization and performed 60 independent optimization runs under three rainfall return periods. Individual runs required approximately 6 to 8 h, resulting in a total computational time of approximately 420 h.

In light of these limitations, surrogate models provide an alternative for reducing this computational burden by approximating the input-output relationships of computationally expensive hydrological models. Once trained using high-fidelity simulation data, surrogate models can replace repeated numerical simulations during optimization and other computationally intensive analyses, substantially reducing computational time (Brazão et al. 2026; Garzon et al. 2022; Yi et al. 2025). Approaches used for this purpose include polynomial response surfaces, radial basis functions, multiple linear regression, Gaussian processes, and artificial neural networks (ANNs) (Feng et al. 2023; Yang et al. 2023). Among these approaches, ANNs are particularly attractive because of their ability to approximate complex and nonlinear relationships involving large numbers of input variables.

Recent studies have demonstrated the potential of ANN-based surrogates for urban drainage and LID optimization. Seyedashraf et al. (2021) developed a disaggregation-emulation framework in which ANN models represented hydraulic interactions within drainage networks, substantially reducing the computational burden of many-objective optimization, with MLP models outperforming generalized regression neural networks. Yang et al. (2023) coupled machine-learning surrogates with NSGA-II for LID optimization and found that a backpropagation neural network provided the best predictive performance, reducing optimization time by approximately 96% relative to SWMM-based optimization when training time was excluded. Similarly, Feng et al. (2023) coupled an MLP surrogate with a multiobjective genetic algorithm, reducing computational time to 0.37% of that required by conventional SWMM-based optimization. More recently, Lee and Park (2026) coupled a machine-learning surrogate with a genetic algorithm to optimize land-use-specific LID combinations across multiple LID types and rainfall conditions.

Despite these advances, an important question concerning the reliability of surrogate-assisted optimization for LID systems remains unexplored. Surrogate models are approximations of the underlying high-fidelity model and inevitably introduce prediction errors. Although conventional validation metrics can demonstrate high overall predictive accuracy, they do not necessarily reveal how predictive accuracy varies across the decision space, particularly within regions preferentially explored by the optimization algorithm. This issue is particularly relevant in multiobjective optimization, where even relatively small local prediction errors may modify the relative performance and dominance relationships among candidate solutions, potentially altering the resulting Pareto front. Indeed, previous studies have recognized that surrogate errors may bias the optimization search and lead to suboptimal decisions, with this problem becoming increasingly relevant when multiple objectives are simultaneously evaluated (Brazão et al. 2026; Macasieb et al. 2025; Seyedashraf et al. 2021).

Moreover, agreement in the objective space does not necessarily imply consistency in the underlying designs, as different combinations of decision variables may produce similar objective-function values. However, systematic comparisons of surrogate-assisted and direct high-fidelity optimization in terms of both objective functions and decision variables remain limited for LID systems. Thus, questions remain regarding whether surrogate-assisted optimization preserves the main design characteristics identified by high-fidelity optimization and whether surrogate-induced errors can be efficiently corrected without compromising computational efficiency.

To address these gaps, this study systematically evaluates the reliability of an ANN-assisted multiobjective optimization framework for LID design by comparing it with direct SWMM-based optimization. ANN surrogate models are developed to predict peak outflow (POF) and total outfall volume (TOV) under different rainfall and LID configurations and are coupled with NSGA-II to simultaneously minimize POF, TOV,

and LID implementation and maintenance costs. The analysis evaluates not only the predictive performance of the surrogate models but also their ability to guide the optimization toward Pareto-optimal regions comparable to those identified by the high-fidelity framework. The nondominated solutions identified through surrogate-assisted optimization are subsequently re-evaluated using SWMM, providing a computationally efficient high-fidelity verification step in which surrogate-predicted objective values are replaced by SWMM-computed values and dominance relationships are reassessed. Finally, the two optimization approaches are compared in terms of both objective-space performance and the resulting LID design characteristics, including LID selection, spatial allocation, implementation area, and layer thicknesses, while their overall computational requirements are quantified.

2. Methodology

2.1. Study area

The study was conducted in a predominantly residential area encompassing blocks 305 and 306 in Asa Norte, Brasília, Federal District, Brazil (Fig. 1a). The area covers approximately 14.7 ha and is located within the Paranoá Lake basin. This relatively small urban area allows detailed representation of land-use characteristics and the spatial distribution of surfaces potentially suitable for implementing LID practices.



153 **Fig. 1** Study area and model representation: (a) urban drainage network and feasible areas
154 identified for the implementation of BTs, PPs, and GRs; and (b) PCSWMM model of the
155 study area

156

157 Approximately 49% of the catchment consists of impervious surfaces, including building
158 rooftops (19%), roads and parking areas (18%), and other paved surfaces, such as
159 sidewalks and sports courts (12%). The terrain slopes predominantly toward the northeast,
160 with an average gradient of 2.5% and elevations ranging from 1,092 to 1,108 m. The
161 regional climate is classified as tropical savanna, with an average annual precipitation of
162 approximately 1,540 mm, concentrated mainly between October and April. Red Latosol
163 is the predominant soil type, with a reported infiltration capacity of 24.4 mm h^{-1} and a

164 classification corresponding to Hydrologic Soil Group A (Costa 2013; NRCS 2009; Souza
165 2014).

166 2.2. Model setup

167 The hydrological–hydraulic representation of the study area (Fig. 1b) was based on the
168 SWMM model originally developed by Lopes and Silva (2021), which was adapted for
169 the purposes of the present investigation. The original model was constructed using
170 PCSWMM, supported by land-use information derived from aerial imagery and
171 cartographic data for the local drainage system. Its spatial discretization comprised 57
172 subcatchments, including 20 building rooftops, 13 paved areas associated with roads and
173 parking facilities, and 24 subcatchments characterized by mixed land uses. For the present
174 study, one subcatchment was removed, resulting in 56 hydrological units, of which 33
175 represent homogeneous rooftop or paved surfaces and 23 comprise heterogeneous areas,
176 including gardens, green spaces, sports courts, and public squares.
177 The original drainage network representation was retained. Lopes and Silva (2021)
178 developed a simplified configuration of the existing drainage system, focusing on its main
179 conduits and excluding upstream connections that conveyed runoff from outside the
180 modeled catchment. The network geometry was established using drainage information
181 provided by the local regulatory agency, while manhole elevations were estimated from
182 topographic data.
183 The hydrological parameters were also largely preserved from the original model.
184 Subcatchment areas were obtained through PCSWMM, while characteristic flow widths
185 were estimated using Eq. (1), following the approach adopted by Lopes and Silva (2021):

$$W = k\sqrt{A} \tag{1}$$

186

where W represents the characteristic flow width, A is the subcatchment area, and k is a dimensionless coefficient. A value of 0.7 was adopted for k, consistent with the original model and supported by the calibration results reported by Krebs et al. (2014). This formulation provides a simplified representation of overland flow geometry without requiring individual flow-path lengths for each subcatchment.

Surface slopes and impervious fractions were defined according to the spatial characteristics represented in the original model. Rooftop slopes were assigned a value of 1%, whereas slopes for other surfaces were estimated from topographic contours and prevailing runoff directions. Imperviousness was characterized using aerial imagery, and infiltration was simulated using the SCS Curve Number method, with parameter values corresponding to Hydrologic Soil Group A. Depression storage and surface roughness parameters were retained from the original model, as summarized in Table 1. During the generation of the simulation datasets, subcatchment flow widths were automatically adjusted to account for changes in the areas occupied by LID controls, following the procedure described by Rossman and Simon (2022). This adjustment ensured that the hydrological representation remained consistent with the different LID implementation areas considered in the simulations.

213 **Table 1** Subcatchment parameters (Source: Lopes and Silva 2021)

Parameter		Adopted Value
k – dimensionless coefficient		0.7
Manning's roughness coefficient	Pervious	0.15
	Impervious	0.011
Depression storage capacity (mm)	Pervious	3.81
	Impervious	1.9
CN	Pervious	39
	Impervious	98

214

215 2.3. LID controls

216 Three LID practices were considered in the optimization framework: green roofs (GRs),
217 permeable pavements (PPs), and bioretention cells (BTs). These practices were selected
218 because they provide complementary mechanisms for managing stormwater runoff,
219 including interception, infiltration, temporary storage, and flow attenuation. Their
220 combined implementation also allows the investigation of alternative spatial
221 arrangements and design configurations, building upon the LID modeling approach
222 previously established by Lopes and Silva (2021).

223 The distribution of potential LID locations was defined according to the characteristics
224 of the existing urban surfaces. The original model included 20 potential GR locations on
225 building rooftops and 13 PP locations associated with roads and parking areas, both of
226 which were retained in the present study. In contrast, the number of potential BT locations
227 was increased from three to ten. The additional locations were identified through the
228 analysis of satellite imagery and local land-use conditions, prioritizing subcatchments
229 containing impervious surfaces from which runoff could be conveyed to bioretention

facilities. The resulting spatial configuration comprises 43 potential LID locations, with the maximum implementation areas determined according to the available surfaces.

In addition to expanding the number of potential BT locations, the runoff routing configuration was modified based on the findings of Lopes and Silva (2021). In the original model, rooftop runoff was directed to either PPs or BTs, depending on their proximity. In the present study, rooftop runoff, including contributions from both GR and non-GR portions, was directed exclusively to downstream subcatchments containing potential PPs, as Lopes and Silva (2021) reported that PPs exhibited greater hydrological efficiency and lower costs than the other LID practices. Within these subcatchments, rooftop runoff and runoff generated from impervious pavement surfaces were routed through the PP units before entering the stormwater drainage network.

BTs were configured to receive direct precipitation and runoff generated from the impervious portions of their respective contributing subcatchments. GRs were modeled to receive only direct precipitation. The revised runoff routing pathways are illustrated in Fig. 2, with runoff ultimately conveyed to the drainage network through the corresponding manholes according to the local topography.

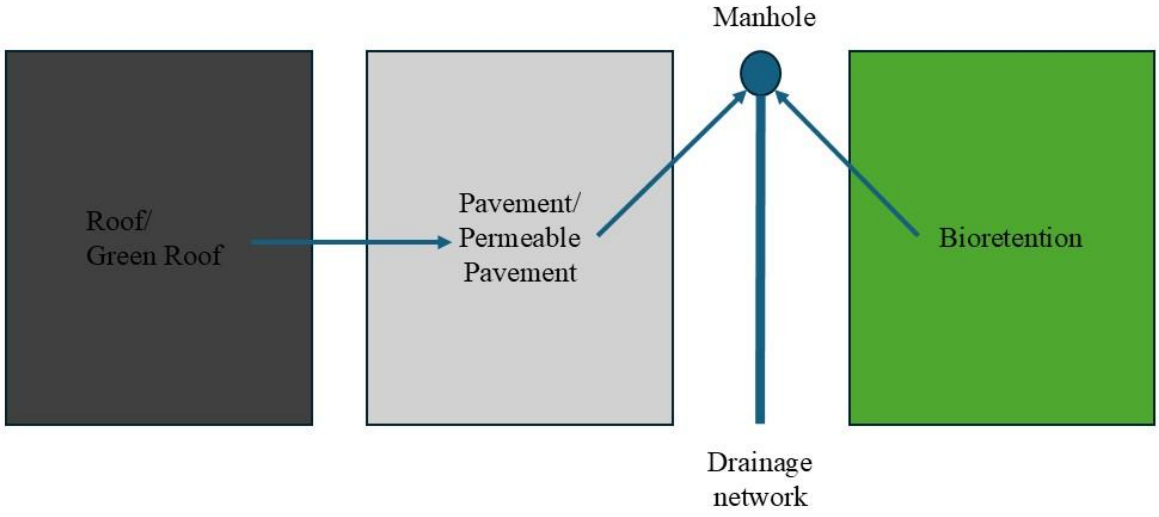


Fig. 2 Schematic representation of the revised runoff routing pathways in the study area

The physical parameters used to represent the LID controls were primarily inherited from Lopes and Silva (2021), whose values were established based on previous studies rather than site-specific calibration. The main modification concerned the surface berm height of BTs, which was increased from 150 to 300 mm to provide additional surface storage capacity. Other parameters, including soil properties, pavement characteristics, storage-layer properties, and drainage-layer characteristics, were retained from the original model. The adopted values are summarized in Table 2, with the parameters expressed as ranges corresponding to the design variables subsequently considered in the optimization procedure.

Table 2 LID parameters (Adapted from Lopes and Silva 2021)

Layer	Parameter	Bioretention	Permeable Pavement	Green Roof
Surface	Berm Height (mm)	300	0	25
	Vegetation volume fraction	0.1	0	0.1
	Roughness (manning's n)	0.1	0.012	0.1
	Surface slope (%)	1	1	1
Soil	Thickness (mm)	600-1500	-	50
	Porosity (volume fraction)	0.5	-	0.5
	Field capacity (volume fraction)	0.2	-	0.3
	Wilting point (volume fraction)	0.1	-	0.1
	Conductivity K (mm/h)	119	-	120
	Conductivity slope (%)	10	-	10
	Suction head (mm)	45	-	60
Pavement	Thickness (mm)	-	100	-
	Void Ratio (voids/solids)	-	0.1	-

	Impervious surface fraction	-	0	-
	Permeability (mm/h)	-	500	-
	Clogging factor	-	0	-
Storage	Thickness (mm)	150-500	150-900	-
	Void Ratio (voids/solids)	0.75	0.5	-
	Seepage rate (mm/h)	36.1	36.1	-
	Clogging factor	0	0	-
Drainage mat	Thickness (mm)	-	-	50
	Void fraction	-	-	0.5
	Roughness (manning's n)	-	-	0.1

2.4. Design storms

The ANN training process incorporated different rainfall scenarios to ensure that the resulting models could represent a wide range of precipitation conditions. These scenarios comprised design storms with return periods of 2, 5, 10, and 25 years, and durations of 30, 140, and 360 minutes. The duration of 140 minutes corresponds to the time of concentration of the study area, representing the critical condition for peak flow generation. The additional durations of 30 and 360 minutes were included to evaluate the predictive capability of the neural networks under shorter and longer rainfall events, respectively, thereby exposing the models to a wider range of hydrological responses. These scenarios were selected to represent a broad spectrum of rainfall intensities, while avoiding extreme events (e.g., 50- and 100-year return periods), for which previous studies have reported reduced effectiveness in peak runoff mitigation, particularly for bioretention systems (Ekmekcioglu et al. 2021).

Rainfall hyetographs were generated using the alternating block method, considering three temporal peak positions located at 1/7, 1/2, and 7/8 of the total storm duration. For the 360-minute rainfall events with a 25-year return period, two additional hyetographs were generated by positioning the maximum rainfall intensity at 45 and 315 minutes. The inclusion of these temporal arrangements was intended to expose the ANN models to a broader range of rainfall patterns during training, thereby improving the representation of different hydrological responses. A temporal discretization of 1 minute was adopted for all simulations. Equation (2) presents the local Intensity–Duration–Frequency (IDF) relationship used to compute rainfall intensities.

$$i = \frac{1{,}574.7\ T^{0{,}207}}{(D + 11)^{0{,}884}} \quad (2)$$

where i denotes the design rainfall intensity (mm/h), T represents the return period (years), and D corresponds to the rainfall duration (minutes).

2.5.ANN models for POF and TOV prediction

This study adopted a Multilayer Perceptron (MLP) neural network architecture, characterized by feedforward information propagation and trained using supervised learning. The selection of an ANN-based surrogate was supported by previous studies demonstrating high predictive accuracy in reproducing SWMM-derived hydraulic responses, including superior performance compared with alternative machine-learning approaches (Seyedashraf et al. 2021; Yang et al. 2023), as well as substantial reductions in the computational burden associated with LID optimization (Feng et al. 2023). Moreover, the MLP architecture was considered appropriate because the relationship

between the LID design parameters and the target SWMM outputs (POF and TOV) is static and does not require modeling temporal dependencies. The networks were trained using backpropagation, with weights and biases optimized using the Levenberg–Marquardt algorithm.

The model inputs were defined to represent both rainfall characteristics and key design and allocation parameters associated with the selected LIDs, ensuring direct correspondence with variables relevant to subsequent optimization problems. Rainfall was characterized by three descriptors: return period (RP), duration (D), and peak time (PT). The inclusion of these variables allowed the ANN to simulate system responses under different rainfall conditions using a single model, rather than relying on separate surrogate models tailored to individual scenarios. The LID design inputs included the areas of BTs, PPs, and GRs, as well as the soil and storage layer thicknesses of BTs and the storage layer thickness of PPs. Although layer thicknesses are typically prescribed based on design guidelines or previous studies, their hydrological and hydraulic relevance justified their inclusion as ANN inputs and, consequently, as decision variables in the optimization framework. Based on the findings of Lopes and Silva (2021), GR soil layer thickness was not considered as a variable because its influence on hydrological performance was found to be limited under the conditions of the study area.

To represent the spatial allocation of LIDs, a binary variable (I) was assigned to each predefined LID location. When $I = 0$ for a specific LID unit, that unit was considered inactive within its corresponding subcatchment, whereas $I = 1$ indicated its activation. These binary variables enable the optimization algorithm to determine whether a LID should be implemented at each candidate location, which is particularly relevant when searching for cost-efficient configurations. Furthermore, the inclusion of I allows a single ANN to represent multiple LID implementation configurations, avoiding the need to generate separate datasets for specific combinations of LID types (e.g., BT only, BT

combined with GR, or configurations including PP). This formulation therefore increases the flexibility and computational efficiency of the surrogate modeling framework.

For the study area, this formulation resulted in a total of 122 input variables for each ANN model, comprising 119 LID-related variables and three rainfall-related variables (RP, D, and PT). The 119 LID-related variables were subsequently used as decision variables in the multiobjective optimization, whereas the rainfall descriptors were fixed according to the rainfall scenario being optimized. Each ANN model had a single output, corresponding to either POF or TOV.

The hyperbolic tangent (*tanh*) activation function was adopted because of its suitability for modeling nonlinear input–output relationships and its widespread use in feedforward neural networks for nonlinear regression problems. Accordingly, the input and output data were normalized to the range [−1, 1], consistent with the output range of the *tanh* function. The ANN outputs were subsequently denormalized to recover their values in the original physical domain.

2.6. Dataset Generation

The datasets were constructed using a MATLAB-based computational routine that automated the execution of SWMM simulations and the collection of the corresponding results. For each simulation, input parameters were randomly sampled within predefined bounds, according to the characteristics of each decision variable. Continuous values were assigned to LID areas and layer thicknesses, whereas binary values were used to indicate whether each LID unit was implemented.

The minimum areas adopted were 100 m² for GRs, 20 m² for BTs, and 100 m² for PPs, in order to avoid very small installations with limited effectiveness. For permeable pavements and green roofs, the maximum allowable area corresponded to the total area

of each respective element. For bioretention cells, the maximum area was set to 20% of the impervious area of the subcatchment in which they are located.

After the input data were generated, SWMM was executed successively via command line calls. In each iteration, the input text file (*.inp*) was modified to reflect the values corresponding to the new sample to be processed. Upon completion of each simulation, the information related to POF and TOV was extracted from the output file (.rpt) and saved, thereby forming the dataset outputs used for training and validating the neural networks.

To improve the representation of high-performance configurations in the training domain, 100 samples with all LID units activated (i.e., $I = 1$ for all LIDs) were incorporated into the dataset. This strategy improves the representation of configurations with widespread LID implementation, which is particularly relevant in optimization contexts where search algorithms may preferentially explore regions associated with high hydrological performance.

For inactive LID units (i.e., I = 0), their associated area and layer thicknesses were assigned zero values. This procedure prevented physically irrelevant design parameters from introducing unnecessary variability into the ANN input space, promoting a more consistent representation of LID configurations and their corresponding SWMM responses.

2.7. ANN model configuration

The development of the ANN models involved evaluating the influence of training dataset size and hidden-layer configuration on predictive performance. To identify an appropriate number of training samples, model accuracy was examined alongside the computational effort associated with data generation and network training. Four dataset sizes were investigated (5,000, 10,000, 20,000, and 50,000 samples), following the ranges explored

in previous studies (Garzón et al. 2022; Yang et al. 2023). This analysis aimed to identify a dataset size capable of adequately representing the relationships between model inputs and outputs without imposing excessive computational demands. To ensure comparability, the ANN hyperparameters were kept constant throughout these experiments, as specified in Table 3.

Table 3 ANN parameters for preliminary tests

Parameter	Value/Function
Sample division	80% training and validation – 20% test
Minimum training error	1×10^{-8}
Maximum number of epochs	500
Activation function (all layers)	Hyperbolic Tangent (tanh)
Number of hidden layers	1
Neurons per hidden layer	20
Optimization algorithm	Levenberg–Marquardt (LM)
Loss function	MSE
Normalization	Min–max scaling
Batch size	Full batch

Bayesian optimization was adopted to determine suitable ANN architectures by exploring different combinations of hidden layers and neurons, using the root mean square error (RMSE) as the minimization criterion. The search space was defined based on preliminary simulations and the findings of Yang et al. (2023), who successfully applied a single-hidden-layer ANN with 21 neurons to estimate TOV and POF. Accordingly, the optimization considered architectures comprising either one or two hidden layers, with 10 to 30 neurons in each layer. To maintain a manageable computational cost, the Bayesian optimization procedure was restricted to a maximum of 20 iterations.

2.8. Multiobjective optimization

An adapted version of NSGA-II implemented using the MATLAB Global Optimization Toolbox (Wang et al. 2018) was employed to solve the multiobjective optimization problem. This implementation incorporates a controlled elitism strategy that limits the number of individuals from the Pareto front retained for the subsequent generation, thereby preserving population diversity and reducing the likelihood of premature convergence.

The multiobjective optimization problem was formulated based on the objective functions adopted by Lopes and Silva (2021), considering the simultaneous minimization of POF, TOV, and the economic costs associated with LID implementation and maintenance, as expressed in Eqs. (3)–(5).

$$\textit{Objective function (1)} = \min \text{POF} \tag{3}$$

$$\textit{Objective function (2)} = \min \text{TOV} \tag{4}$$

$$\textit{Objective function (3)} = \min \sum_{i=1}^{43} [C_C + C_{OM}]_i \tag{5}$$

where *POF* is peak flow at the outfall of the catchment, *TOV* is total volume at the outfall of the catchment, i is the candidate LID location index, C_C is construction cost of each LID, and C_{OM} is its operation and maintenance (O&M) cost.

2.8.1. Decision variables

The selection of decision variables is a key step in formulating the multiobjective optimization problem. For BTs, the decision variables comprised the implementation area and the thicknesses of the soil and storage layers. PPs were represented by their implementation area and storage-layer thickness, whereas only the implementation area was optimized for GRs. In addition to these continuous design variables, the presence or absence of each LID at the predefined candidate locations was represented by binary decision variables. A value of 0 indicated that the corresponding LID was not

implemented, whereas a value of 1 indicated its implementation. Altogether, the optimization problem comprised 119 decision variables.

2.8.2. Economic objective

The economic objective represented by Eq. (5) considered both the implementation and O&M costs of the LID systems, following the cost-estimation methodology adopted by Lopes and Silva (2021). The unit costs of the materials and construction activities required for BTs, PPs, and GRs are summarized in Table 4. The original cost estimates were updated to 2026 values using SINAPI (National System for Research on Civil Construction Costs and Indices) and SICRO (Referential System of Construction Costs), widely used Brazilian databases for construction cost estimation, and were supplemented with market prices from local suppliers when necessary. Further details on the cost-estimation procedure, including the calculation of O&M costs, can be found in Lopes and Silva (2021).

Table 4 Unit Costs for each LID system (Adapted from Lopes and Silva 2021)

LID	Data source[1]	Item	Cost (R$)[2]
Bioretention	SICRO	Excavation (m^3)	46.56
	SINAPI	Gravel (m^3)	175.45
	SINAPI	Geotextile (m^2)	10.34
	SINAPI	Soil (m^3)	203.57
	Online research	Vegetation (unit)	20.00
Permeable Pavement	SICRO	Excavation (m^3)	46.56
	SICRO	Porous asphalt (m^3)	341.46
	SINAPI	Gravel (m^3)	175.45
	SINAPI	Geotextile (m^2)	10.34
Green Roof	SINAPI	Grass (m^2)	16.00
	SINAPI	Soil (m^3)	203.57
	SINAPI	Geotextile (m^2)	10.34
	SINAPI	Expanded clay (m^3)	775.28
	SINAPI	Waterproof layer (m^2)	117.36

Notes: September 2026, $ 1 = R$ 5.09

[1] All data were collected in 2026 for Brasília (Federal District)

[2] Local currency - Real

2.8.3. NSGA-II configuration

The parameters governing the NSGA-II search and their values are summarized in Table 5. The algorithm configuration was established based on parameter settings reported in previous studies (Lopes and Silva 2021; Wang et al. 2018), with adaptations to accommodate the continuous and binary decision variables considered in this study. For

each independent optimization run, the initial population was randomly generated within the search space defined by the lower and upper bounds of each decision variable. Continuous variables, including LID implementation areas and layer thicknesses, were randomly sampled within their respective bounds, whereas the variables representing the presence or absence of each LID unit were generated as binary values.

Table 5 NSGA-II parameters adopted in the multiobjective optimization

Parameter	Type	Size/rate/value
Population	Random	500
Selection	Tournament	2
Crossover	Arithmetic/uniform inheritance	0.80
Mutation	Gaussian/ bit-flip	0.10/0.05
Pareto Fraction	Controlled elitism	0.35
Stopping criteria	Maximum generation	200
	Function tolerance	10^{-4}

Note: For crossover and mutation, customized operators were employed to handle continuous and binary decision variables. Continuous variables underwent arithmetic crossover and Gaussian mutation, whereas binary variables were inherited from either parent with equal probability and mutated using a 5% bit-flip probability.

The multiobjective optimization was performed separately for the 5- and 25-year return-period scenarios, both considering a storm duration of 140 min and a peak time of 70 min. For each scenario, the rainfall descriptors were fixed throughout the optimization process, while the 119 LID-related variables were optimized by NSGA-II.

2.9. Surrogate-assisted optimization and high-fidelity validation

To evaluate the reliability of the surrogate-assisted optimization framework, NSGA-II was independently executed ten times using the ANN models as surrogates for the SWMM-derived hydrological objective functions. For each candidate solution, POF and

TOV were estimated by the corresponding ANN models, whereas the economic objective was calculated directly from the decision variables.

The nondominated solutions obtained from the ten independent runs were subsequently pooled, and a global Pareto front was identified using the ANN-predicted POF and TOV values together with the corresponding LID costs. All solutions belonging to this global front were then re-evaluated using the SWMM model. The ANN-predicted POF and TOV values were replaced by their corresponding SWMM-simulated values, while the economic objective remained unchanged because it depends exclusively on the decision variables.

Following this high-fidelity re-evaluation, the dominance relationships among the candidate solutions were reassessed using the SWMM-derived objective values. Solutions that became dominated after replacement of the surrogate predictions were removed, resulting in a SWMM-validated Pareto front derived from the surrogate-assisted optimization.

For benchmarking purposes, this validated front was subsequently compared with a reference Pareto front obtained by directly coupling NSGA-II with SWMM, in which the hydrological objective functions were evaluated using SWMM throughout the evolutionary search. The surrogate-assisted multiobjective optimization framework adopted in this study is illustrated in Fig. 3.

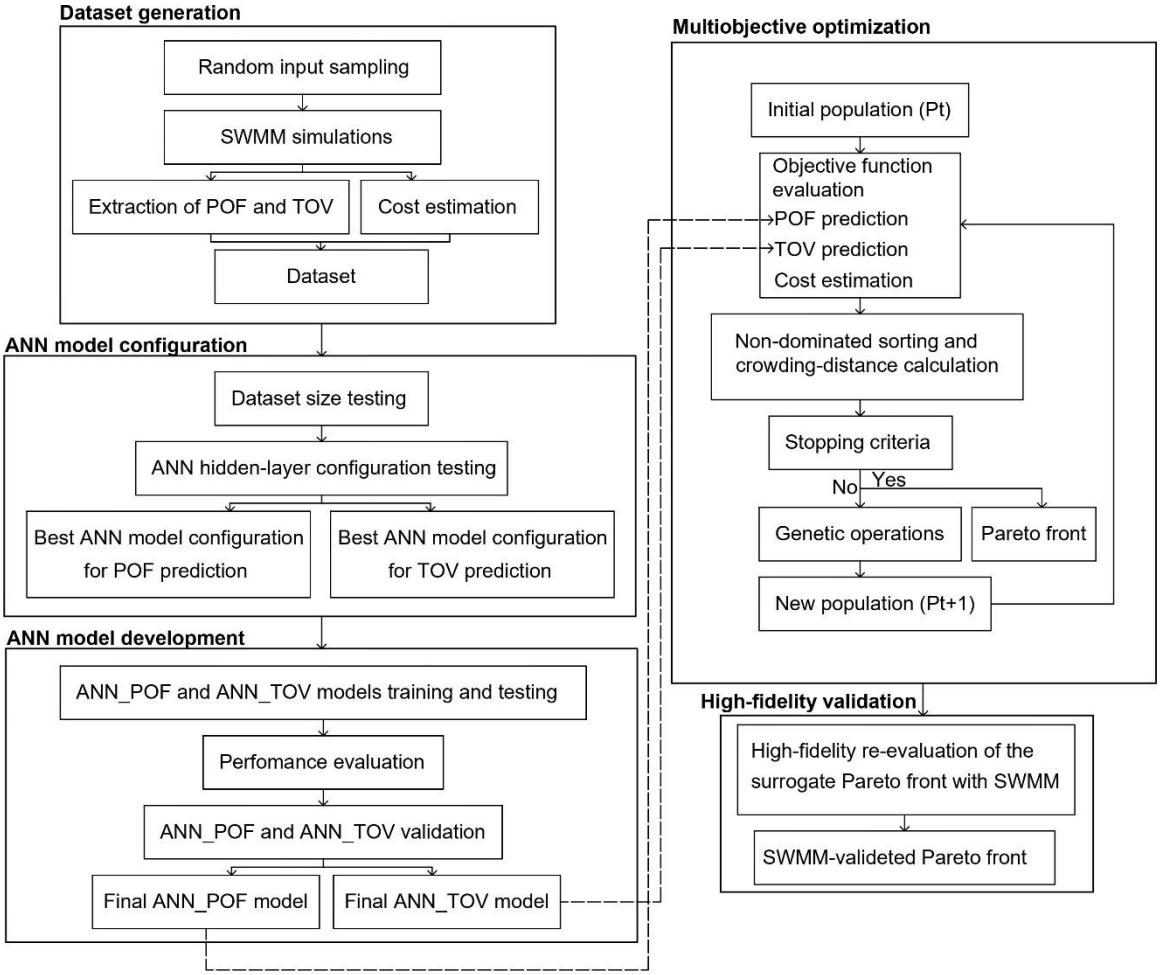


Fig. 3 Surrogate-assisted multiobjective optimization framework adopted in this study

2.10. Evaluation metrics

2.10.1. ANN predictive performance

Five statistical metrics were used to assess the predictive accuracy of the ANN models: mean absolute percentage error (MAPE), mean squared error (MSE), root mean squared error (RMSE), percent bias (PBIAS), and Nash–Sutcliffe efficiency (NSE). Their

mathematical formulations are provided in Eqs. (6) – (10). Model performance was considered more satisfactory as MAPE, MSE, and RMSE approached zero, while PBIAS was used to quantify systematic overestimation or underestimation. NSE values closer to unity indicated stronger agreement between predicted and reference values.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i}\ x\ 100 \tag{6}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{8}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{9}$$

$$PBIAS = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{\sum_{i=1}^{n}(y_i)} \tag{10}$$

where n represents the total number of samples, y_i and $\hat{y}_i$ denote the observed and predicted values for sample i, respectively, and $\bar{y}$ corresponds to the mean of the observed values.

2.10.2. Pareto-front comparison

The agreement between the Pareto fronts was quantitatively assessed using the Inverted Generational Distance (IGD), which measures the average distance from each solution in a reference Pareto front to its nearest solution in the evaluated front. The NSGA-II + SWMM front was adopted as the reference set. Prior to calculation, the objective functions were normalized to avoid scale-dependent effects. The IGD was calculated as shown in Eq. (11).

513

$$IGD(A,R) = \frac{1}{|R|}\sum_{r\epsilon R} min_{a\epsilon A}\,||r-a||_2 \quad (11)$$

514

515 where R is the reference Pareto front, A is the evaluated Pareto front, |R| is the number of
516 solutions in the reference set, and ||r-a||2 denotes the Euclidean distance in the normalized
517 objective space. Lower IGD values indicate greater proximity between the evaluated and
518 reference fronts, with values approaching zero representing higher agreement.

519 2.10.3. LID design-pattern comparison

520 In addition to the comparison of the Pareto fronts in the objective space, Spearman's rank
521 correlation coefficient (ρs) was used to assess the consistency of the LID design patterns
522 identified by the two optimization approaches (NSGA-II+ANN+SWMM validation and
523 NSGA-II+SWMM). This nonparametric measure was selected because it evaluates
524 monotonic associations without requiring assumptions of linearity or normality.
525 Spearman's correlation was calculated for all candidate LID locations jointly and
526 separately for each LID type (BT, PP, and GR), allowing the agreement between the two
527 approaches to be evaluated both globally and for individual LID practices. Since the
528 analyzed design variables may contain repeated values, particularly the binary variables
529 describing the presence or absence of LIDs, tied observations were assigned average
530 ranks. Accordingly, Spearman's coefficient was calculated as the Pearson correlation
531 coefficient between the ranked variables, as shown in Eq. (12).

532

$$\rho_s = \frac{cov\,(R_x,R_y)}{\sigma_{Rx}\sigma_{Ry}} \quad (12)$$

533

where (R_x) and (R_y) represent the ranks of the paired observations obtained from the two optimization approaches, and σ_{Rx} and σ_{Ry} are their respective standard deviations. The coefficient ranges from −1 to 1, with values approaching 1 indicating strong agreement between the design patterns identified by the two approaches, values near zero indicating weak or no monotonic association, and values approaching −1 indicating opposite ranking patterns. Statistical significance was assessed using the corresponding p-value, adopting a significance level of $p < 0.05$.

2.11. Software and technical implementation

The computational framework combined PCSWMM, EPA SWMM 5.2, and MATLAB R2025b to support hydrological–hydraulic modeling, ANN development, and multiobjective optimization. PCSWMM served as the interface for constructing the drainage system model, whereas EPA SWMM 5.2 was used to execute the simulations required for dataset generation. The simulation workflow was automated in MATLAB, with multiple SWMM runs distributed across four processing cores using the parfor function to reduce computational time.

The surrogate models were constructed and trained using the fitnet and train functions of the MATLAB Deep Learning Toolbox. Network architecture selection was supported by Bayesian optimization through the bayesopt function. For the multiobjective optimization stage, the NSGA-II algorithm was implemented using gamultiobj from the Global Optimization Toolbox. Custom crossover and mutation operators were incorporated to handle the combination of continuous and binary decision variables. Parallel execution was also applied to objective-function evaluations during optimization. All computational experiments were performed on a computer equipped with an 11th-generation Intel® Core™ i7 processor (3.40 GHz) and 16 GB of RAM.

559 ## 3. Results and discussion

560 ### 3.1. ANN model configuration

561 #### 3.1.1. Training dataset size analysis

562 The computational time associated with dataset generation for the ANN_POF and
563 ANN_TOV models is summarized in Table 6. Larger datasets required progressively
564 longer processing times due to the increased number of SWMM simulations. Across the
565 evaluated dataset sizes, the average simulation time was approximately 0.16 s per sample.

566 **Table 6** Computational time for dataset generation

Dataset size	Time (h)
5,000 samples	0.19
10,000 samples	0.46
20,000 samples	0.96
50,000 samples	1.93

567

568 Fig. 4 presents the training and validation performance of the ANN_POF and ANN_TOV
569 models, for the four investigated training dataset sizes. For the models trained with 20,000
570 and 50,000 samples, the training and validation curves followed similar trends and
571 training extended over a larger number of epochs, suggesting more stable training and
572 validation behavior compared with the models trained using 5,000 and 10,000 samples.
573 Table 7 presents the test-phase performance of ANN_POF and ANN_TOV models, for
574 the four investigated dataset sizes. For both models, increasing the dataset size from
575 10,000 to 20,000 samples resulted in only marginal performance improvements, whereas
576 training with 50,000 samples yielded a more substantial improvement. In particular, the
577 50,000-sample dataset produced the best overall test performance for both models, with
578 the lowest MAPE and RMSE values, the highest NSE values, and PBIAS values close to

zero. Considering these performance gains and the low computational cost required to generate the datasets (less than 2 h), the 50,000-sample dataset was selected for training both models in the subsequent stages of this study.

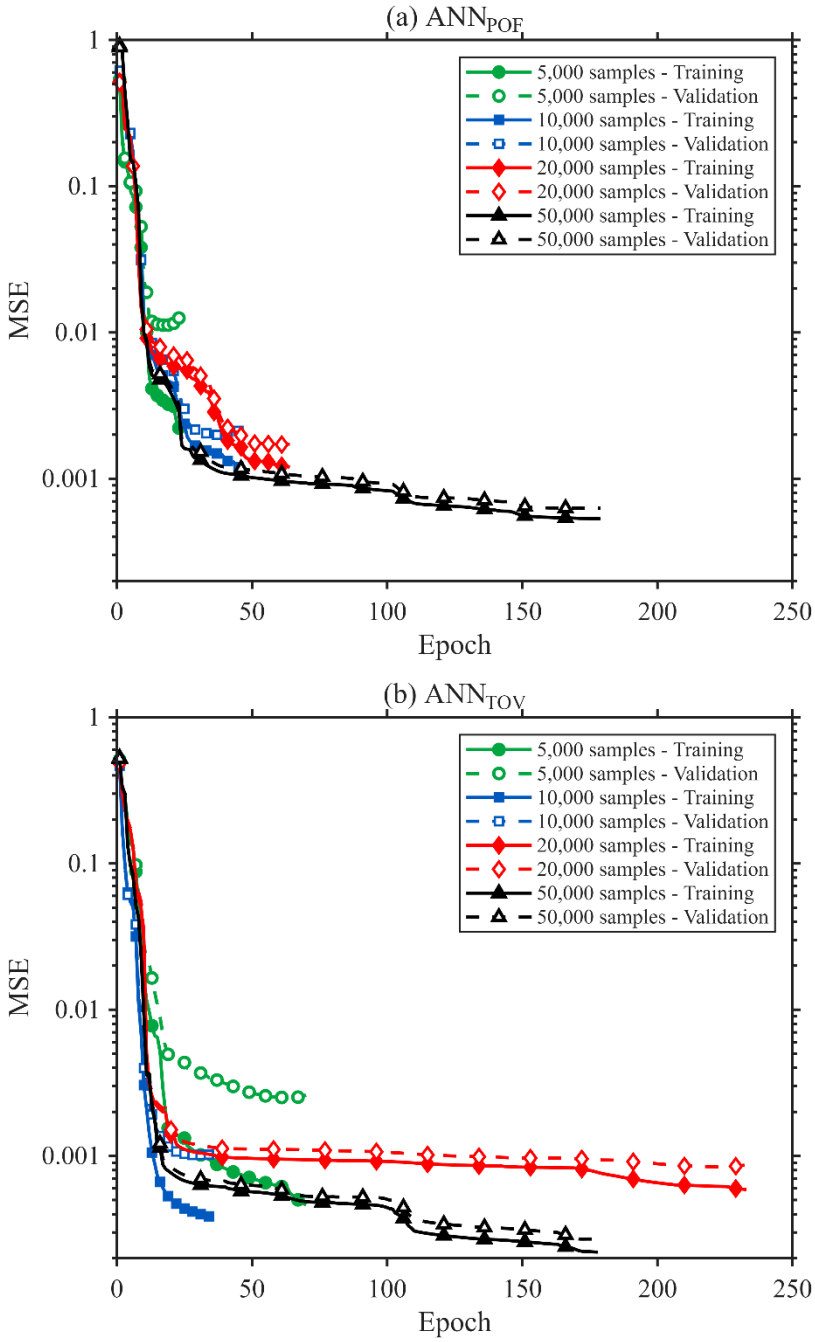


Fig. 4 Training and validation MSE evolution for the ANN models trained using different dataset sizes: (a) ANN_POF and (b) ANN_TOV

585 **Table 7** Test-phase performance of the ANN_POF and ANN_TOV models
586 trained using different dataset sizes

Dataset size	MAPE (%)	RMSE	NSE	PBIAS
		ANN_POF		
5,000 samples	11.647	0.258	0.9254	1.5692
10,000 samples	5.3749	0.112	0.9874	0.3667
20,000 samples	5.044	0.086	0.9925	0.1567
50,000 samples	3.006	0.062	0.9961	0.0461
		ANN_TOV		
5,000 samples	5.772	162.088	0.9845	-0.0556
10,000 samples	3.982	110.782	0.9930	-0.0073
20,000 samples	3.739	98.903	0.9945	-0.0181
50,000 samples	2.311	59.282	0.9980	0.0258

587 Note: RMSE is expressed in m^3/s for ANN_POF and in m^3 for ANN_TOV.

588

589 3.1.2. ANN architecture analysis

590 The effect of ANN architecture on test performance is shown in Fig. 5 for the ANN_POF
591 and ANN_TOV models. For ANN_POF, configurations with two hidden layers exhibited
592 similar performance over the range of 15 to 30 neurons per hidden layer. A comparable
593 behavior was observed for the models with a single hidden layer, with RMSE values
594 remaining relatively stable across most of the investigated configurations. The main
595 exceptions were the configuration with 10 neurons and two hidden layers and that with
596 28 neurons and a single hidden layer, both of which resulted in considerably higher RMSE
597 values. Overall, these results indicate that ANN_POF performance was relatively
598 insensitive to the number of neurons within the investigated range. Among the evaluated

architectures, the lowest RMSE was obtained with two hidden layers and 20 neurons per layer, which was therefore selected for the subsequent stages of this study.

A different behavior was observed for ANN_TOV. As shown in Fig. 5(b), the use of two hidden layers generally resulted in lower RMSE values than the use of a single hidden layer. Moreover, for the two-hidden-layer configurations, RMSE tended to decrease as the number of neurons per layer increased, although some variability was observed. The lowest RMSE was achieved with 30 neurons per layer. Accordingly, the architecture comprising two hidden layers with 30 neurons each was selected for ANN_TOV in the subsequent analyses.

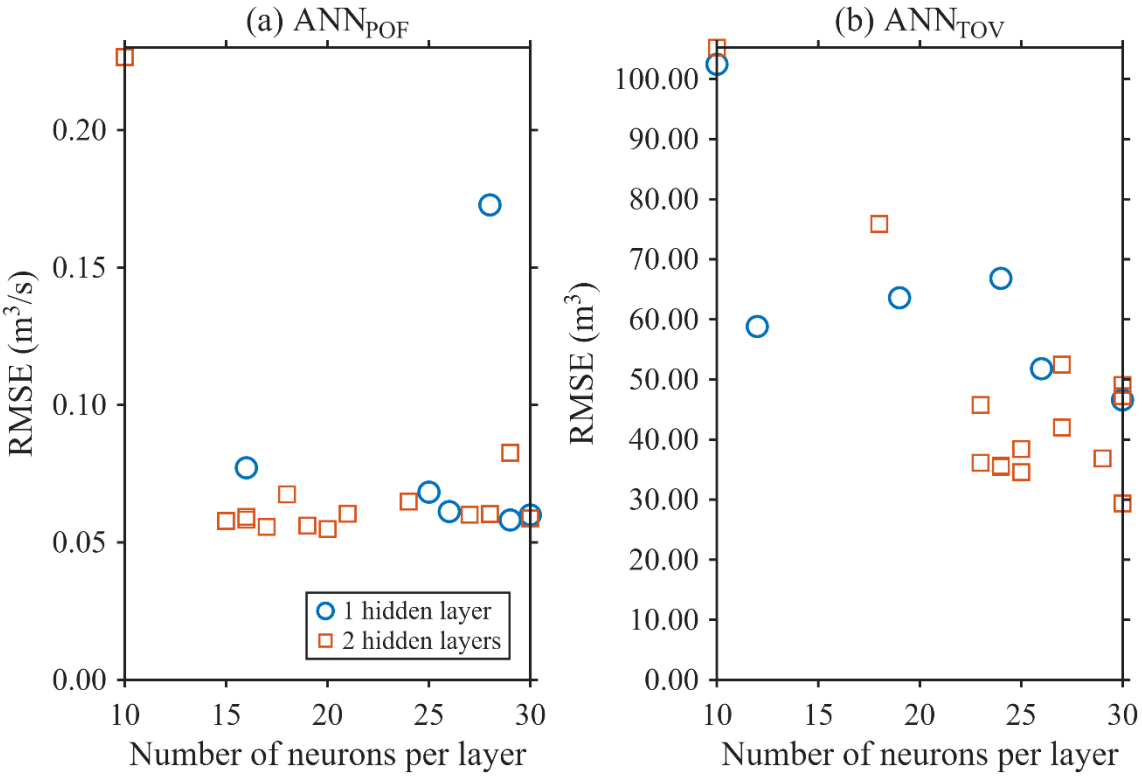


Fig. 5 Test-phase RMSE for different ANN architectures: (a) ANN_POF and (b) ANN_TOV

3.2. ANN model Performance Evaluation

Tables 8 and 9 present the performance of the ANN_POF and ANN_TOV models, respectively, during the testing phase considering the different rainfall scenarios and LID

614 configurations included in the training dataset. Overall, both models exhibited good
615 predictive performance across the investigated conditions, with MAPE values below
616 10%, NSE values close to unity, and PBIAS values close to zero in most cases. The
617 relatively small variation in these performance metrics indicates that differences in
618 rainfall characteristics and LID configurations did not substantially affect model
619 accuracy, highlighting the robustness of both ANN models across the range of conditions
620 considered in the training dataset.

621 **Table 8** Predictive accuracy of the ANN_POF model across different rainfall conditions

Rainfall scenario	MAPE (%)	RMSE ($m^3 s^{-1}$)	NSE	PBIAS
RP=2 y, D=30 min, PT=15 min	3.110	0.058	0.9558	0.8289
RP=2 y, D=140 min, PT=70 min	3.212	0.063	0.9505	-0.7705
RP=2 y, D=360 min, PT=180 min	3.119	0.065	0.9469	0.6287
RP=5 y, D=30 min, PT=15 min	2.430	0.060	0.9658	0.5950
RP=5 y, D=140 min, PT=70 min	2.733	0.070	0.9594	-0.3522
RP=5 y, D=360 min, PT=180 min	2.592	0.070	0.9570	0.4764
RP=10 y, D=30 min, PT=15 min	2.064	0.062	0.9746	-0.1016
RP=10 y, D=140 min, PT=70 min	2.286	0.075	0.9620	-0.2605
RP=10 y, D=360 min, PT=180 min	2.399	0.080	0.9590	0.4244
RP=25 y, D=30 min, PT=15 min	1.630	0.065	0.9764	-0.3858
RP=25 y, D=140 min, PT=70 min	1.919	0.085	0.9601	0.0120
RP=25 y, D=360 min, PT=180 min	2.001	0.089	0.9583	-0.1762
RP=25 y, D=360 min, PT=45 min	1.856	0.079	0.9690	0.1780
RP=25 y, D=360 min, PT=315 min	8.146	0.027	0.7696	0.1094

622 Note: RP = return period; D = rainfall duration; PT = peak time.

623

624

625

626 **Table 9** Predictive accuracy of the ANN_TOV model across different rainfall
627 conditions

Rainfall scenario	MAPE (%)	RMSE (m^3)	NSE	PBIAS
RP=2 y, D=30 min, PT=15 min	1.944	32.523	0.9818	0.3048
RP=2 y, D=140 min, PT=70 min	1.824	47.382	0.9833	-0.0155
RP=2 y, D=360 min, PT=180 min	1.734	52.957	0.9819	0.1238
RP=5 y, D=30 min, PT=15 min	1.798	37.810	0.9819	0.3645
RP=5 y, D=140 min, PT=70 min	1.525	53.707	0.9845	-0.0280
RP=5 y, D=360 min, PT=180 min	1.443	57.512	0.9848	0.0338
RP=10 y, D=30 min, PT=15 min	1.580	40.536	0.9851	0.3254
RP=10 y, D=140 min, PT=70 min	1.383	60.691	0.9839	-0.0304
RP=10 y, D=360 min, PT=180 min	1.330	64.768	0.9861	-0.0683
RP=25 y, D=30 min, PT=15 min	1.481	49.504	0.9822	0.4265
RP=25 y, D=140 min, PT=70 min	1.202	69.148	0.9840	0.1494
RP=25 y, D=360 min, PT=180 min	1.200	78.954	0.9833	0.0370
RP=25 y, D=360 min, PT=45 min	1.209	80.637	0.9840	-0.0823
RP=25 y, D=360 min, PT=315 min	2.897	30.779	0.9717	0.0138

628 Note: RP = return period; D = rainfall duration; PT = peak time.

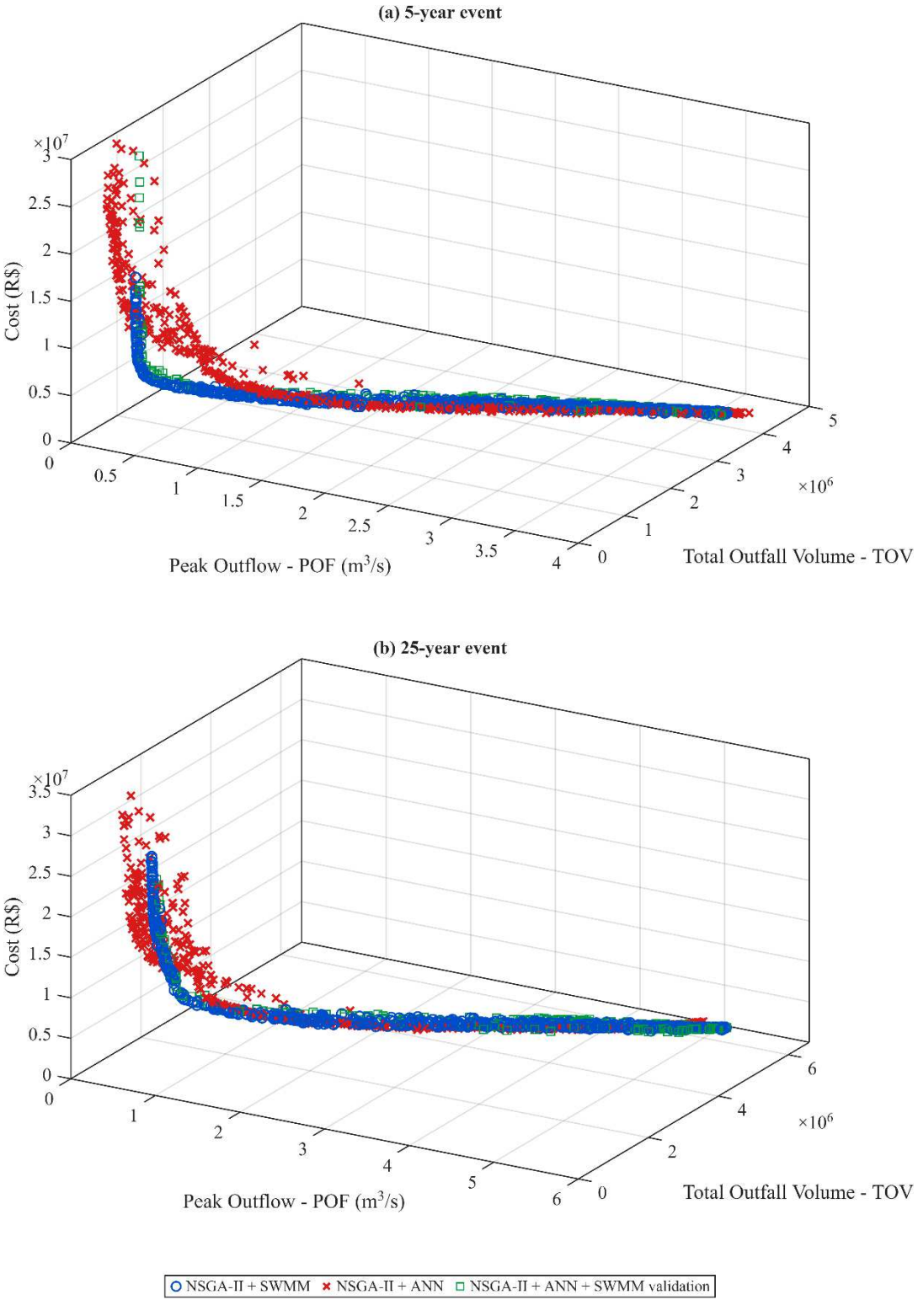
629

630 3.3 Pareto front and Cost-Efficiency Analysis

631 Fig. 6 presents the Pareto fronts obtained for the 5- and 25-year rainfall scenarios using
632 three optimization approaches. The proposed surrogate-assisted framework (NSGA-II +
633 ANN + SWMM validation) was compared with direct high-fidelity optimization (NSGA-
634 II + SWMM), adopted as the reference approach, and with surrogate-based optimization
635 without high-fidelity re-evaluation (NSGA-II + ANN), included to assess the effect of the
636 final SWMM re-evaluation on the resulting Pareto front. For the 5-year rainfall scenario,
637 651 non-dominated solutions were obtained using NSGA-II + SWMM, 532 using NSGA-

638 II + ANN, and 258 using NSGA-II + ANN + SWMM validation. For the 25-year rainfall

639 scenario, the corresponding numbers were 258, 484, and 248, respectively.



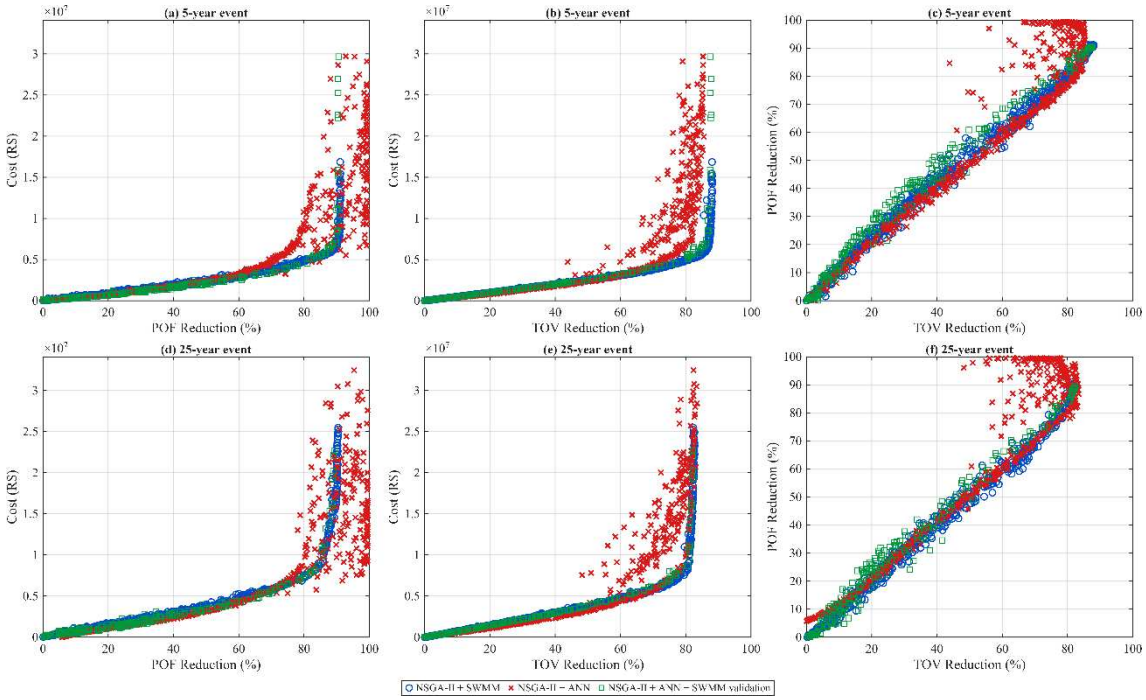
640

641 **Fig. 6** Pareto fronts obtained using the three optimization approaches for the (a) 5-year

642 and (b) 25-year rainfall scenarios

Regardless of the optimization approach, well-distributed Pareto fronts were obtained for both rainfall scenarios, providing a diverse set of solutions representing different trade-offs among the competing objectives. Moreover, the three approaches produced Pareto fronts occupying similar regions of the objective space, indicating that the ANN-assisted optimization was able to identify trade-off solutions comparable to those obtained through the direct coupling of NSGA-II and SWMM. The fronts also exhibit smooth trade-offs between cost and hydraulic performance up to a clearly identifiable change in slope, beyond which substantial increases in cost result in comparatively small additional improvements in hydraulic performance. This behavior indicates diminishing marginal benefits as the solutions approach the highest levels of hydraulic performance.

Fig. 7 presents the cost-efficiency curves for POF and TOV reduction, as well as the trade-off between POF and TOV reduction, for the 5- and 25-year rainfall scenarios. Overall, despite the differences in rainfall magnitude, the cost-efficiency curves exhibited similar patterns for both return periods. In both rainfall scenarios, increasing hydraulic performance was associated with progressively higher costs, with similar trends in the trade-offs between POF and TOV reduction. The three optimization approaches also exhibited similar cost-efficiency patterns across most of the objective space. However, the optimization based on NSGA-II + ANN resulted in greater dispersion of solutions in the region associated with higher hydraulic performance.



669 **Fig. 7** Cost-efficiency and hydraulic performance trade-offs for the 5- and 25-year rainfall
670 scenarios: (a) cost-efficiency curve for POF reduction under the 5-year rainfall scenario;
671 (b) cost-efficiency curve for TOV reduction under the 5-year rainfall scenario; (c) trade-
672 off between POF and TOV reduction under the 5-year rainfall scenario; (d) cost-efficiency
673 curve for POF reduction under the 25-year rainfall scenario; (e) cost-efficiency curve for
674 TOV reduction under the 25-year rainfall scenario; and (f) trade-off between POF and
675 TOV reduction under the 25-year rainfall scenario
676
677 Table 10 presents the predictive performance of the ANN models for the non-dominated
678 solutions identified through NSGA-II + ANN, while Figure 8 shows the corresponding
679 scatter plots comparing the ANN predictions with the SWMM simulations. For TOV, the
680 ANN predictions showed good agreement with the SWMM results, with NSE values
681 above 0.98 for both rainfall scenarios. MAPE values below 12% further indicate that the
682 ANN maintained relatively low average prediction errors for TOV during the
683 optimization process. In contrast, the ANN exhibited greater difficulty in predicting POF,
684 as evidenced by higher MAPE values, on the order of 30%, greater PBIAS values, and

lower NSE values. The scatter plots further indicate that the largest discrepancies between ANN predictions and SWMM simulations were concentrated at lower POF values, suggesting that the predictive accuracy of the surrogate model was not uniform across the range of solutions explored during the optimization.

Table 10 Predictive performance of the ANN models for the non-dominated solutions identified through the NSGA-II + ANN coupling, based on comparison with the corresponding SWMM simulations

Model	MAPE (%)	RMSE	NSE	PBIAS
		5-year Rainfall Scenario		
POF	34.639	0.316	0.919	0.941
TOV	11.794	1.541×10^5	0.987	-0.554
		25-year Rainfall Scenario		
POF	29.736	0.628	0.832	15.976
TOV	6.129	2.519×10^5	0.982	4.377

Note: RMSE is expressed in m^3/s for POF and in m^3 for TOV.

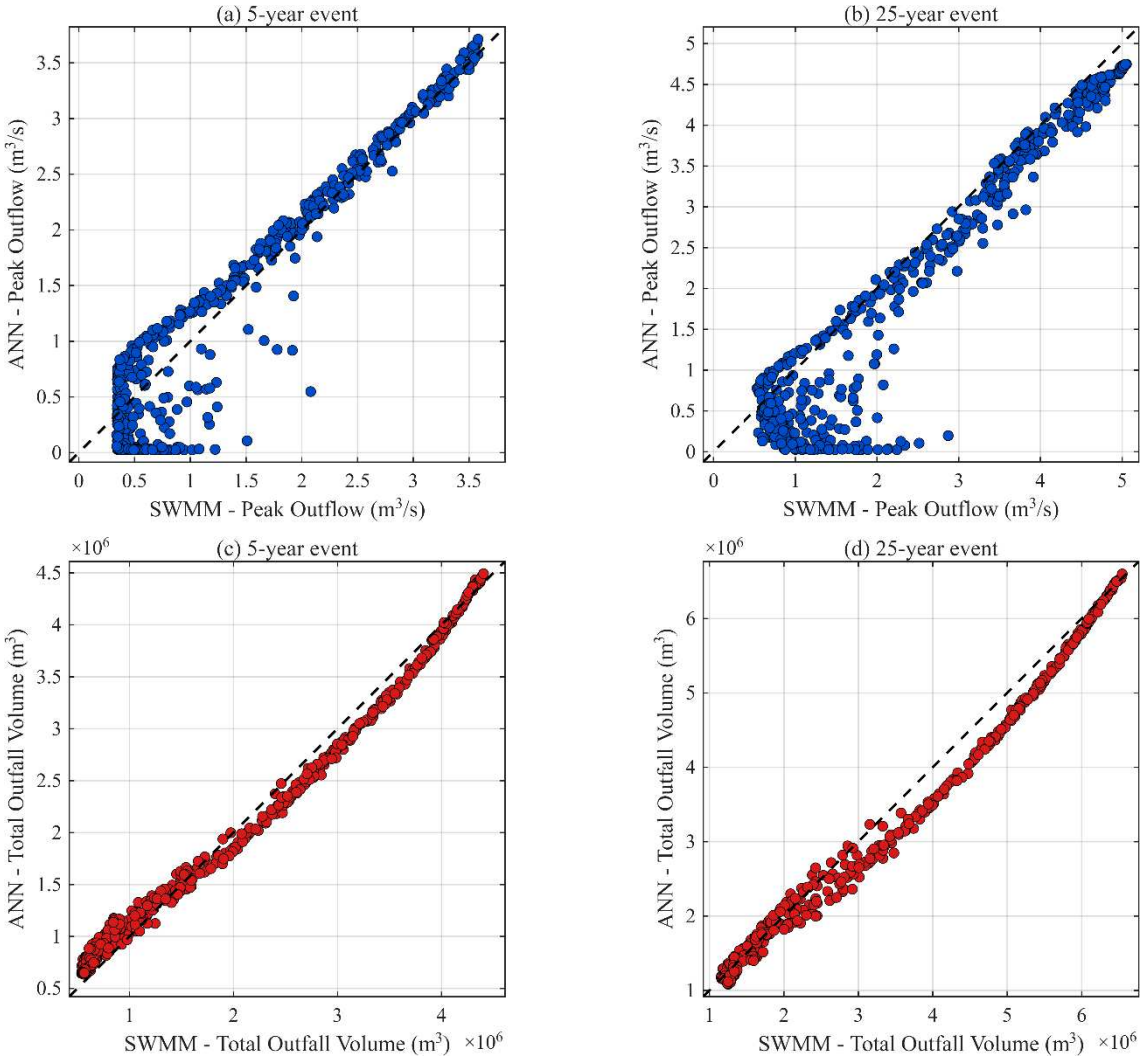


Fig. 8 Scatter plots comparing ANN predictions and SWMM simulations for the non-dominated solutions identified through the NSGA-II + ANN coupling: (a) 5-year rainfall scenario and (b) 25-year rainfall scenario

The results suggest that the surrogate models experienced greater difficulty in accurately estimating the objective functions in the high-performance region of the search space, particularly for POF. A possible explanation is related to the random sampling strategy adopted to generate the training dataset. Although this strategy provides broad coverage of the decision space, Pareto-optimal or near-Pareto-optimal solutions may represent only a small fraction of randomly generated samples. This limited representation was evident in the training dataset. For the 5-year scenario (D = 140 min and PT = 70 min), only 3 of

1,321 samples (0.23%) achieved reductions of at least 70% in POF, and the same number and proportion were observed for TOV. For the 25-year scenario, none of the 1,418 samples achieved reductions of at least 70% in either POF or TOV. Consequently, the ANNs can achieve high overall predictive performance while exhibiting comparatively larger local errors in regions preferentially explored by the optimization algorithm. This effect becomes particularly relevant during the evolutionary search, as prediction errors may alter the relative performance and dominance relationships among candidate solutions, thereby contributing to the greater dispersion observed in the ANN-based Pareto fronts.

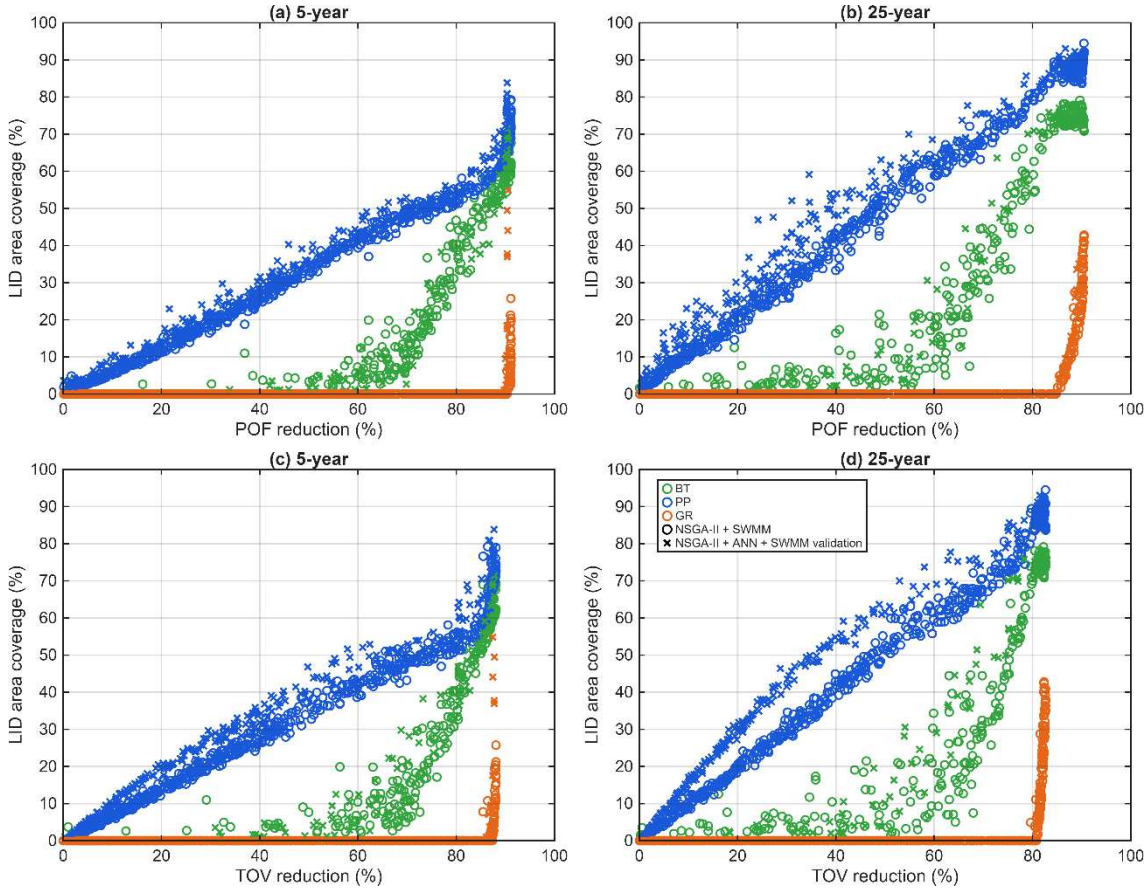
The greater prediction errors observed for POF compared with TOV may also be related to the different hydrological nature of these response variables. POF represents an instantaneous extreme of the outflow hydrograph and is therefore particularly sensitive to changes in runoff timing, flow routing, and the synchronization of contributions from different subcatchments. Consequently, relatively small changes in LID configuration may produce comparatively large changes in peak discharge. In contrast, TOV represents an integrated response over the entire rainfall–runoff event, which tends to smooth short-term variations in the hydrograph and may result in a less complex input–output relationship for the surrogate model.

Despite these local prediction errors, the ANN models were sufficiently accurate to guide NSGA-II toward promising regions of the search space. After the candidate non-dominated solutions identified through the NSGA-II + ANN coupling were re-evaluated using SWMM and the dominance relationships were reassessed based on the SWMM-computed objective values, the resulting high-fidelity Pareto fronts showed greater agreement with those obtained through the direct NSGA-II + SWMM coupling (see Figs. 6 and 7). This improvement was quantitatively assessed using the inverted generational distance (IGD), considering the Pareto fronts obtained through the direct NSGA-II +

SWMM coupling as the reference. For the 5-year rainfall scenario, the IGD decreased from 0.0288 for NSGA-II + ANN to 0.0176 after SWMM re-evaluation and reconstruction of the non-dominated set, corresponding to a 38.9% reduction. Similarly, for the 25-year rainfall scenario, the IGD decreased from 0.0240 to 0.0173, representing a 28.0% reduction. Thus, SWMM re-evaluation consistently improved the agreement between the surrogate-assisted and high-fidelity Pareto fronts for both rainfall scenarios. Although the high-fidelity SWMM re-evaluation caused some ANN-guided solutions to become dominated, reducing the final nondominated set from 532 to 258 solutions for the 5-year event and from 484 to 248 solutions for the 25-year event, the reconstructed Pareto fronts remained well distributed across the objective space. Thus, the removal of solutions affected by surrogate prediction errors did not substantially compromise the diversity of the final solution set, which continued to provide a broad range of trade-offs between implementation cost and hydraulic performance.

3.4. Decision-Variable Analysis

For the decision-variable analysis, the comparison focuses on the solutions obtained using the direct NSGA-II + SWMM approach and the surrogate-guided solutions that remained nondominated after their objective values were re-evaluated using SWMM. Accordingly, Fig. 9 compares the relationship between LID area coverage and hydraulic performance for the nondominated solutions obtained using these two approaches.



764 **Fig. 9** Relationship between LID area coverage and hydraulic performance for the
765 nondominated solutions obtained using the direct and surrogate-assisted optimization
766 approaches: (a) POF reduction, 5-year event; (b) POF reduction, 25-year event; (c) TOV
767 reduction, 5-year event; and (d) TOV reduction, 25-year event

768

769 Overall, the two approaches exhibited similar area–performance relationships for both
770 POF and TOV reductions, indicating that the ANN-guided optimization reproduced the
771 main LID allocation patterns identified by the direct NSGA-II + SWMM coupling. This
772 agreement was evident across all three LID types. PP was the predominant LID type in
773 both approaches, with its coverage increasing almost monotonically with increasing
774 hydraulic performance. Similarly, both approaches allocated relatively limited BT areas
775 at low and intermediate performance levels, followed by a marked increase in BT
776 coverage toward the high-performance region. GR also exhibited a consistent pattern
777 between the two approaches, remaining largely absent from most solutions and becoming

778 relevant mainly near the maximum achievable reductions. Thus, both optimization
779 approaches revealed a similar LID allocation hierarchy, characterized by the preferential
780 use of PP, followed by increasing contributions from BT and, ultimately, GR as higher
781 levels of hydraulic performance were sought. A similar preference for PP and BT over
782 GR was reported by Lopes and Silva (2021) and Yang et al. (2023), suggesting that the
783 allocation patterns consistently identified by both optimization approaches are also in
784 agreement with previous studies. The similarity between the approaches was observed
785 under both return periods, although the 25-year event generally required greater LID
786 coverage to achieve the highest levels of hydraulic performance.
787 To further investigate whether these similarities were also reflected in LID
788 implementation and spatial allocation, the subsequent decision-variable analysis focused
789 on the high-performance region of the solution space. Only solutions simultaneously
790 achieving at least 70% reductions in both POF and TOV were considered. This criterion
791 resulted in 195 and 165 nondominated solutions for the NSGA-II + SWMM approach
792 under the 5- and 25-year rainfall scenarios, respectively, compared with 28 and 43
793 solutions for the NSGA-II + ANN approach followed by SWMM validation. The high-
794 performance solutions revealed broadly similar LID implementation patterns between the
795 two optimization approaches, although the degree of agreement varied among LID types
796 and rainfall scenarios. For the 5-year rainfall scenario, both BT and PP were present in all
797 high-performance solutions obtained with either approach, whereas GR occurred less
798 frequently. The average number of implemented locations was also similar for PP (12.80
799 and 12.84 for NSGA-II + SWMM and NSGA-II + ANN + SWMM validation,
800 respectively) and, to a lesser extent, for BT (8.01 and 8.77), while larger differences were
801 observed for GR (1.01 and 2.53).
802 Regarding spatial prioritization, a significant Spearman rank correlation was observed for
803 PP locations under the 5-year rainfall scenario ($\rho = 0.736$, $p = 0.004$), indicating similar

804 spatial preferences between the two approaches. Although the correlation for BT was not
805 statistically significant ($\rho = 0.475$, $p = 0.165$), some of the most frequently selected
806 locations, particularly BT1 and BT3, were consistently prioritized by both approaches. In
807 contrast, GR locations showed little rank agreement ($\rho = 0.122$, $p = 0.610$), indicating
808 greater variability in their spatial allocation.
809 For the 25-year rainfall scenario, the predominance of BT and PP became even more
810 pronounced. Both LID types occurred in all high-performance solutions, with an average
811 of 9.30 and 9.79 BT locations and 12.98 and 13.00 PP locations per solution for the
812 NSGA-II + SWMM and NSGA-II + ANN + SWMM validation approaches, respectively.
813 In particular, all PP locations were selected in every surrogate-assisted solution, while
814 their selection frequencies ranged from 98.46% to 100% in the SWMM-based solutions.
815 Consequently, Spearman's coefficient could not be computed for PP because of the
816 absence of variability in the surrogate-assisted selection frequencies. The results indicate
817 an almost identical and systematic prioritization of PP locations by both approaches.
818 GR exhibited greater variability between the approaches and rainfall scenarios. Its
819 occurrence increased under the 25-year event, but the spatial rankings remained weakly
820 correlated ($\rho = 0.278$, $p = 0.236$). Overall, the analysis of the high-performance solutions
821 indicates that the surrogate-assisted optimization preserved the main LID implementation
822 and spatial allocation patterns identified by the direct SWMM-based optimization,
823 particularly the extensive and systematic use of PP and BT. The greater variability
824 observed for GR suggests that multiple spatial configurations of this LID type may
825 provide comparable hydraulic performance rather than reflecting a systematic divergence
826 between the optimization approaches.
827 Fig. 10 compares the normalized LID layer thicknesses obtained for the high-performance
828 solutions using the direct SWMM-based and surrogate-assisted optimization approaches.
829 Overall, the two approaches produced broadly similar thickness distributions, indicating

that the ANN-guided optimization largely preserved the LID sizing characteristics identified by the direct NSGA-II + SWMM coupling. This agreement was particularly evident for the 5-year event, for which the median values and interquartile ranges were comparable between the two approaches for the BT soil, BT storage, and PP storage layers. For the 25-year event, similar distributions were also observed, although some differences became more apparent. The median PP storage thickness remained nearly identical between the approaches, whereas the surrogate-assisted solutions tended to exhibit a lower median BT soil thickness and a higher median BT storage thickness. The latter represented the most noticeable difference between the two approaches, although substantial overlap between their distributions was still observed. Overall, these results indicate that the surrogate-assisted optimization was able to reproduce the main LID sizing patterns of the high-fidelity optimization, while allowing some variability in individual design variables, particularly under the more severe rainfall event.

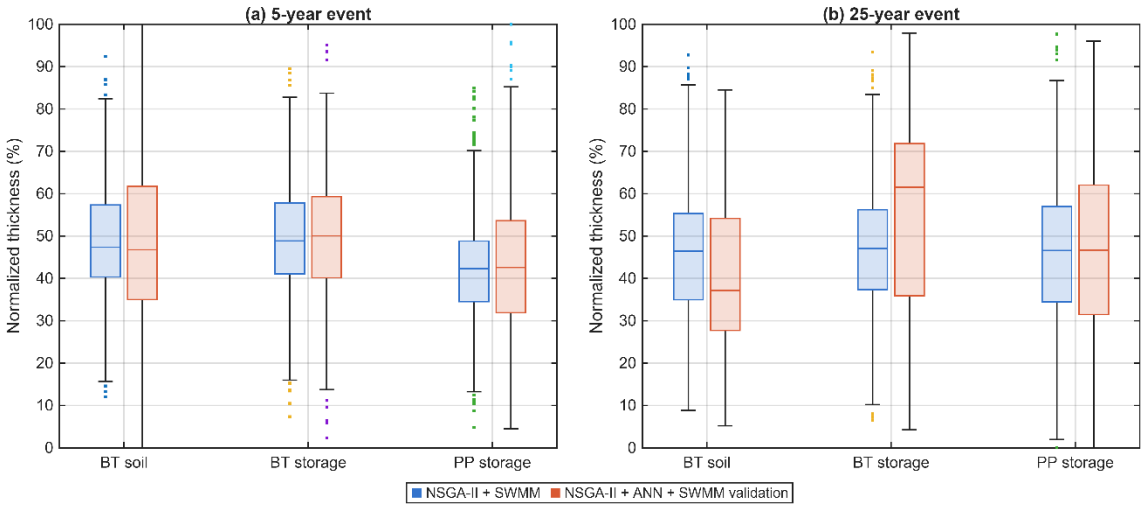


Fig. 10 Comparison of normalized LID layer thicknesses between the direct and surrogate-assisted optimization approaches for the high-performance solutions: (a) 5-year event and (b) 25-year event

3.5 Computational Efficiency Analysis

The direct NSGA-II + SWMM optimization required a total computational time of 83.64 h, comprising 38.71 h for the 5-year rainfall scenario and 44.93 h for the 25-year scenario. All optimization runs satisfied the stopping criterion based on a tolerance of 10^{-4}, requiring, on average, 135 and 155 generations for the 5- and 25-year scenarios, respectively. In contrast, the NSGA-II + ANN optimization required only 11.59 h, with 5.01 h and 6.58 h spent on the 5- and 25-year scenarios, respectively. All surrogate-assisted runs also satisfied the same stopping criterion, requiring an average of 113 and 123 generations, respectively. Thus, although the surrogate-assisted optimization converged with moderately fewer generations, the substantial reduction in computational time primarily reflects the replacement of repeated high-fidelity SWMM simulations by computationally inexpensive ANN evaluations during the evolutionary Search. When the computational overhead associated with surrogate model development was included, namely, training dataset generation (1.93 h), Bayesian optimization of the ANN architectures (26.55 h), and ANN training (5.31 h), together with the final SWMM re-evaluation of the surrogate-assisted solutions (approximately 100 s), the total computational cost of the surrogate-assisted framework was approximately 45.39 h. Even after accounting for all these additional steps, the proposed framework reduced the overall computational time by approximately 46% relative to the direct high-fidelity optimization. Notably, the final high-fidelity re-evaluation represented a negligible fraction of the total computational cost. When only the optimization stage was considered, excluding the one-time computational cost associated with surrogate model development, the computational time was reduced by approximately 86%.

4. Conclusions

This study evaluated the reliability of a surrogate-assisted multiobjective optimization framework for the design and allocation of LID systems by combining ANNs, NSGA-II,

and SWMM simulations. The results demonstrate that ANN surrogate models can substantially reduce the computational burden associated with simulation-based optimization while preserving the main design characteristics of the solutions obtained through direct high-fidelity optimization.

The ANN models showed high predictive performance across the rainfall scenarios and LID configurations considered in the training domain. However, prediction accuracy decreased in the high-performance region preferentially explored during optimization, particularly for POF. Despite these local prediction errors, the surrogate models remained sufficiently accurate to guide NSGA-II toward promising regions of the search space. This finding indicates that accurate reproduction of the high-fidelity model throughout the entire decision space is not necessarily required for effective surrogate-assisted optimization. Instead, the surrogate must provide sufficiently reliable information to guide the evolutionary search, while subsequent high-fidelity re-evaluation can be used to obtain reliable objective-function values and reassess the final nondominated solutions. This strategy proved effective for both rainfall scenarios. Re-evaluating the ANN-guided candidate solutions with SWMM and reconstructing the nondominated sets reduced the IGD from 0.0288 to 0.0176 for the 5-year event and from 0.0240 to 0.0173 for the 25-year event, corresponding to improvements of 38.9% and 28.0%, respectively. Thus, incorporating high-fidelity re-evaluation substantially improved the agreement between the surrogate-assisted and direct optimization approaches while ensuring that the final objective values and dominance relationships were determined from SWMM simulations. Importantly, the agreement between the two approaches extended beyond the objective space. The surrogate-assisted framework reproduced the main relationships between LID area coverage and hydraulic performance and revealed a similar allocation hierarchy, with PP being preferentially implemented, followed by increasing contributions from BT and GR as higher levels of hydraulic performance were sought. Analysis of the high-

performance solutions further showed that the main implementation and spatial allocation patterns were generally preserved, particularly for PP and BT. Greater variability was observed in the spatial allocation of GR, suggesting that multiple configurations of this LID type can provide comparable hydraulic performance. The two approaches also produced broadly similar LID layer-thickness distributions, although some differences emerged under the more severe rainfall event. These findings indicate that the surrogate-assisted approach can preserve not only the overall hydraulic trade-offs but also important physical characteristics of the resulting LID designs.

The computational benefits were substantial. The direct NSGA-II + SWMM optimization required 83.64 h, whereas the complete surrogate-assisted workflow required approximately 45.39 h, including dataset generation, ANN architecture optimization and training, surrogate-assisted optimization, and final SWMM re-evaluation. This corresponds to an overall computational-time reduction of approximately 46%. When only the optimization stage was considered, the reduction reached approximately 86%. Moreover, the final SWMM re-evaluation required only approximately 100 s, demonstrating that high-fidelity verification can be incorporated into the surrogate-assisted framework with negligible additional computational cost.

Overall, the results demonstrate that the main advantage of surrogate modeling in LID optimization does not necessarily depend on perfectly reproducing the high-fidelity model across the entire search space. Rather, ANNs can serve as computationally efficient search guides, while targeted high-fidelity re-evaluation ensures that the final objective values and dominance relationships are determined using the process-based model at a small additional computational cost. The proposed framework therefore provides an efficient and reliable decision-support approach for computationally intensive multiobjective LID optimization, enabling the exploration of hydrological and economic

trade-offs among alternative LID configurations while combining the speed of data-driven surrogate models with the reliability of process-based hydrological simulations.

The main limitation of this study is that the SWMM model was not calibrated against observed hydrological or hydraulic data, which limits the direct interpretation of the optimized solutions as site-specific design recommendations. However, this limitation does not directly affect the internal comparison between the direct and surrogate-assisted optimization frameworks, as both approaches were evaluated using the same SWMM representation. Future research should therefore evaluate the proposed framework using calibrated models in other urban catchments and under different hydrological conditions, explore adaptive sampling strategies to enrich the training dataset in Pareto-optimal and near-Pareto-optimal regions, and extend the surrogate models to predict hydraulic responses at individual drainage-system nodes. Such spatially distributed predictions could broaden the applicability of surrogate-assisted optimization to drainage-system rehabilitation by enabling the identification of critical locations and the evaluation of targeted interventions.

References

Bibi TS, Kara KG, Bedada HJ, Bededa RD (2023) Application of PCSWMM for assessing the impacts of urbanization and climate changes on the efficiency of stormwater drainage systems in managing urban flooding in Robe town, Ethiopia. J Hydrol Reg Stud 45:101291. https://doi.org/10.1016/j.ejrh.2022.101291

Brazão AJC, Silva CM, Costa MEL, Koide S, Silva GBL (2026) Evaluation of artificial neural network surrogate models for detention-based flood control in urban areas. J Water Resour Plan Manag 152(11):04026052. https://doi.org/10.1061/JWRMD5.WRENG-7598

949 Brown RA, Line DE, Hunt WF (2012) LID treatment train: Pervious concrete with
950 subsurface storage in series with bioretention and care with seasonal high water tables. J
951 Environ Eng 138(6):689–697. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000506
952
953 Chiu YJ, Chang LC, Lin YT (2025) Hydrologic performance assessment of low impact
954 development facilities based on monitoring data and SWMM modeling in an urban
955 catchment in Taiwan. Sci Rep 15:27251. https://doi.org/10.1038/s41598-025-11586-7
956
957 Costa MEL (2013) Monitoramento e modelagem hidrológica de águas de drenagem
958 urbana na bacia do Lago Paranoá. Master's thesis, Universidade de Brasília, Brasília,
959 Brazil
960
961 Eckart K, McPhee Z, Bolisetti T (2018) Multiobjective optimization of low impact
962 development stormwater controls. J Hydrol 562:564–576.
963 https://doi.org/10.1016/j.jhydrol.2018.04.068
964
965 Ekmekcioğlu Ö, Yılmaz M, Özger M, Tosunoğlu F (2021) Investigation of the low impact
966 development strategies for highly urbanized area via auto-calibrated Storm Water
967 Management Model (SWMM). Water Sci Technol 84(9):2194–2213.
968 https://doi.org/10.2166/wst.2021.432
969
970 Farina A, Di Nardo A, Gargano R, van der Werf JA, Greco R (2023) A simplified approach
971 for the hydrological simulation of urban drainage systems with SWMM. J Hydrol
972 623:129757. https://doi.org/10.1016/j.jhydrol.2023.129757

973 Feng J, Yang Y, Li J (2023) Optimization of the low-impact development facility area
974 based on a surrogate model. Water Resour Manag 37:5789–5806.
975 https://doi.org/10.1007/s11269-023-03630-w
976
977 Garzón A, Kapelan Z, Langeveld J, Taormina R (2022) Machine learning-based surrogate
978 modeling for urban water networks: Review and future research directions. Water Resour
979 Res 58(5). https://doi.org/10.1029/2021WR031808
980
981 Ionno A, Arsenault R, Troin M, Martel JL, Brissette F (2024) Impacts of climate change
982 on flood volumes over North American catchments. J Hydrol 630:130688.
983 https://doi.org/10.1016/j.jhydrol.2024.130688
984
985 Krebs G, Kokkonen T, Valtanen M, Setälä H, Koivusalo H (2014) Spatial resolution
986 considerations for urban hydrological modelling. J Hydrol 512:482–497.
987 https://doi.org/10.1016/j.jhydrol.2014.03.013
988
989 Lee D, Park I (2026) Surrogate-based optimization of land-use-informed LID strategies
990 for enhanced urban flood resilience. J Environ Manage 409:130017.
991 https://doi.org/10.1016/j.jenvman.2026.130017
992
993 Li H, Zhang C, Chen M, Shen D, Niu Y (2023) Data-driven surrogate modeling:
994 Introducing spatial lag to consider spatial autocorrelation of flooding within urban
995 drainage systems. Environ Model Softw 161:105623.
996 https://doi.org/10.1016/j.envsoft.2023.105623

997 Li Z, Zhang X, Ma Y, Feng C, Hajiyev A (2019) A multi-criteria decision making method
998 for urban flood resilience evaluation with hybrid uncertainties. Int J Disaster Risk Reduct
999 36:101140. https://doi.org/10.1016/j.ijdrr.2019.101140
1000
1001 Liu K, Li J, Xia J, Gao X, Gao J, Jiang C (2022) Study on LID facilities comprehensive
1002 effect evaluation: A case in campus. Ecohydrol Hydrobiol 22(3):530–540.
1003 https://doi.org/10.1016/j.ecohyd.2022.04.001
1004
1005 Lopes MD, Silva GBLD (2021) An efficient simulation-optimization approach based on
1006 genetic algorithms and hydrologic modeling to assist in identifying optimal low impact
1007 development designs. Landsc Urban Plan 216:104251.
1008 https://doi.org/10.1016/j.landurbplan.2021.104251
1009
1010 Macasieb RQ, White JT, Pasetto D, Siade AJ (2025) A probabilistic approach to
1011 surrogate-assisted multi-objective optimization of complex groundwater problems. Water
1012 Resour Res 61(5). https://doi.org/10.1029/2024WR038554
1013
1014 NRCS (2009) National Engineering Handbook, Part 630 Hydrology, Chapter 7:
1015 Hydrologic Soil Groups. U.S. Department of Agriculture, Natural Resources
1016 Conservation Service, Washington, DC
1017
1018 Rossman LA, Simon MA (2022) Storm Water Management Model User's Manual Version
1019 5.2. EPA/600/R-22/030, U.S. Environmental Protection Agency, Cincinnati, OH
1020

Sabbagh M, Browne D, Pickering L, Lintern A, Winfrey B (2025) Urban stormwater bioretention reduces runoff and improves water quality: A global meta-analysis of field studies. J Hydrol, 134163. https://doi.org/10.1016/j.jhydrol.2025.134163

Seemuangngam A, Lin HL (2024) The impact of urbanization on urban flood risk of Nakhon Ratchasima, Thailand. Appl Geogr 162:103152. https://doi.org/10.1016/j.apgeog.2023.103152

Seyedashraf O, Bottacin-Busolin A, Harou JJ (2021) A disaggregation-emulation approach for optimization of large urban drainage systems. Water Resour Res 57. https://doi.org/10.1029/2020WR029098

Souza FP (2014) Monitoramento e modelagem hidrológica da sub-bacia do Lago Paranoá - Brasília/DF - e avaliação de bacia de detenção. Master's thesis, Universidade de Brasília, Brasília, Brazil

Sui X, van de Ven FHM (2023) The influence of Low Impact Development (LID) on basin runoff in a half-urbanized catchment: A case study in San Antonio, Texas. J Hydrol 616:128793. https://doi.org/10.1016/j.jhydrol.2022.128793

Wang Q, Zhou Q, Lei X, Savić DA (2018) Comparison of multiobjective optimization methods applied to urban drainage adaptation problems. J Water Resour Plan Manag 144(11):04018070. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000996

Yang W, Brüggemann K, Seguya D, Ahmed E, Kaeseberg T, Dai H, Hua P, Zhang J, Krebs P (2020) Measuring performance of low impact development practices for the surface runoff management. Environ Sci Ecotechnol 1:100010

Yang Y, Li Y, Huang Q, Xia J, Li J (2023) Surrogate-based multiobjective optimization to rapidly size low impact development practices for outflow capture. J Hydrol 616:128848

Yi K, Yang P, Yang S, Bao S, Xu Z, Tan Q (2025) On the accuracy requirement of surrogate models for adequate global sensitivity analysis of urban low-impact development model. J Hydrol 657:133102. https://doi.org/10.1016/j.jhydrol.2025.133102

Zhang Z, Tian W, Lu C, Liao Z, Yuan Z (2024) Graph neural network-based surrogate modelling for real time hydraulic prediction of urban drainage networks. Water Res 263:122142. https://doi.org/10.1016/j.watres.2024.122142

Statements and Declarations

Funding

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that they have no competing financial or non-financial interests that are relevant to the content of this article.

Author Contributions

Abraão J. C. Brazão: Conceptualization, Methodology, Software, Formal analysis, Investigation, Visualization, Writing – original draft.

Gustavo B. L. Silva: Conceptualization, Methodology, Supervision, Writing – review and editing.

All authors read and approved the final manuscript.

Data Availability

The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Ethics Approval

Not applicable.

Consent to Participate

All authors consented to their participation.

Consent for Publication

All authors consent to this work being published.