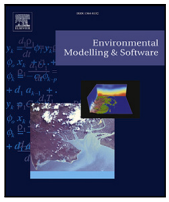
Environmental Modelling and Software 179 (2024) 106128

Contents lists available at ScienceDirect

Environmental Modelling and Software

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



Position Paper

Global optimization-based calibration algorithm for a 2D distributed hydrologic-hydrodynamic and water quality model

Marcus Nóbrega Gomes Jr. [a,b,*], Marcio Hofheinz Giacomoni [b], Fabricio Alonso Richmond Navarro [a], Eduardo Mario Mendiondo [a]

[a] University of São Paulo, Department of Hydraulic Engineering and Sanitation, São Carlos School of Engineering, Av. Trab. São Carlense, 400 - Centro, São Carlos, 13566-590, São Paulo, Brazil

[b] The University of Texas at San Antonio, College of Engineering and Integrated Design, School of Civil & Environmental Engineering and Construction Management, One UTSA Circle, BSE 1.310, San Antonio, 78249, TX, United States of America

ARTICLE INFO

Keywords:
Automatic calibration
HydroPol2D
Parameter estimation
Genetic algorithm

ABSTRACT

Hydrodynamic models with rain-on-the-grid capabilities are usually computationally expensive for automatic parameter estimation. In this paper, we present a global optimization-based algorithm to calibrate a fully distributed hydrologic-hydrodynamic and water quality model (HydroPol2D) using observed data (i.e., discharge, or pollutant concentration) as input. The algorithm finds near-optimal set of parameters to explain observed gauged data. This framework, although applied in a poorly-gauged urban catchment, is adapted for catchments with more detailed observations. The results of the automatic calibration indicate NSE = 0.99 for the V-Tilted catchment, RMSE = 830 mg L^{-1} for salt concentration pollutograph in a wooden-plane (i.e., 8.3% of the event mean concentration), and NSE = 0.89 in a urban real-world catchment. This paper also explores the issue of equifinality (i.e., multiple parameters giving the same calibration performance) in model calibration indicating the performance variation of calibrating only with an outlet gauge or with multiple gauges within the catchment.

Software availability

Name of software: HydroPol2D
Developer: Marcus Nóbrega Gomes Júnior, Ph.D.
Contact address: Department of Hydraulics and Sanitation, University of São Paulo, São Carlos School of Engineering, Av. Trab. São Carlense, 400 - Centro, São Carlos, 13566-590.
Email: marcusnobrega.engcivil@gmail.com
Software required: MATLAB version 2021a or higher.
Programming language: MATLAB
Program size: Approximately 25 MB
Availability: Open source (github license available in Gomes Jr. (2024))

1. Introduction

The advances in computational processing, high resolution GIS data availability, and relatively more complete physically-based models enables the application of fully distributed hydrodynamic and pollutant transport and fate models (Yang et al., 2010; Gomes Júnior et al., 2022). Although the application of fully distributed models (i.e., models that discretize the catchment domain into finite cells) remotes to the late 1970s (Zhang et al., 1990), the demands for high-resolution modeling, especially for flood and pollution assessment, make optimization-based calibration (i.e., herein referred to as automatic calibration) complex, time-consuming, and dependent on prior knowledge of the modeler (Blasone et al., 2008) due to the relatively high computational effort required to perform numerical hydrodynamic simulations with high-resolution data (Blasone et al., 2008; Brath et al., 2004).

The complexity comes because the Shallow-Water-Equations dynamic problem forms a set of hyperbolic partial differential equations with no analytical solution for complex real-world cases (Bermudez and Vazquez, 1994), requiring finite-volume or finite-difference numerical schemes to solve the problem (Brunner, 2016). The hydrodynamic problem can be simplified into diffusive-like problems when local acceleration and inertial terms can be neglected (Akan and Iyer, 2021) and

* Corresponding author at: University of São Paulo, Department of Hydraulic Engineering and Sanitation, São Carlos School of Engineering, Av. Trab. São Carlense, 400 - Centro, São Carlos, 13566-590, São Paulo, Brazil.
E-mail addresses: marcusnobrega.engcivil@gmail.com (M.N. Gomes Jr.), marcio.giacomoni@utsa.edu (M.H. Giacomoni), fabricionavarro@usp.br (F.A.R. Navarro), emm@sc.usp.br (E.M. Mendiondo).
URL: https://www.engenheiroplanilheiro.com.br (M.N. Gomes Jr.).

https://doi.org/10.1016/j.envsoft.2024.106128
Received 15 December 2023; Received in revised form 25 June 2024; Accepted 25 June 2024
Available online 2 July 2024

these can reduce computational effort (Gomes et al., 2023a); however, the number of required simulations for a full calibration still makes the process laborious. Since the governing equations of flow and pollutant routing are generally performed for each element in the grid domain either as a matrix-wise expression or as an element-wise approach, the numerical modeling process can be challenging for finer mesh grids, such as the ones required for flood mapping (do Lago et al., 2023). This might be one of the reasons why only a few articles attempted to develop automatic calibration algorithms for 2D hydrodynamic and pollutant transport and fate models (Afshar et al., 2011).

The parameter discretization of distributed models is usually performed by the discrete categorical values of Land Use and Land Cover (LULC) and Soil type classifications. This information is entered as georeferenced maps, and each cell of the computational domain is assigned with the parameters associated with each input map. For example, in HydroPol2D (Gomes et al., 2023a), hydrological parameters (i.e. Green-Ampt parameters) are assigned with the Soil raster, and hydrodynamic (i.e., Manning's roughness coefficient and initial abstraction) and water quality (i.e., build-up and wash-off parameters) are assigned according to the LULC raster.

The automatic calibration of distributed models can be a taunting task due to the degrees of freedom of the optimization problem and the number of calibration variables that could increase proportionally to the number of land uses and soil classifications (Debele et al., 2008). Moreover, due to the nonlinear behavior of hydrology and hydrodynamics, the use of convex optimization to find global optima is unfeasible, unless several simplifications are performed in the modeling equations (Wang et al., 2020). Additionally, defining the appropriate ranges for model parameters can also lead to unrealistic parameter estimations, especially when the physical boundaries of the parameters are incorrectly treated (Domeneghetti et al., 2012). Due to the complexity of automatic parameter estimations, several studies have successfully performed manual calibrations using distributed models (Ardıçlıoğlu and Kuriqi, 2019; Phillips et al., 2005; Li et al., 2021). Despite these challenges, calibrating a complex model with a relatively large number of parameters can be even more difficult.

Although successful calibrations are presented in the literature, one of the yet unsolved and considerably complex problem is the total reduction of equifinality for a relatively large number of model parameters and ranges (Fatichi et al., 2016). This paper does not attempt to provide a definitive solution to this issue; rather, we explore the factors associated with the chances of finding parameter equifinality in automatic calibration. We investigate the trade-offs between the number of observation points, intensity of rainfall events, and combination of different associations of gauges that can affect parameter equifinality.

Several physically-based hydrologic-hydrodynamic models are available in the literature, such as the Hydrologic Engineering Center - River System Analysis (Brunner, 2016), the Stormwater Management Model - SWMM (Rossman et al., 2010), and the Gridded Surface and Subsurface Analysis (Downer and Ogden, 2004). However, only a few studies used the aforementioned commercial models or developed new models that can take advantage of automatic calibration capabilities (Cho and Lee, 2015; Dung et al., 2011; Hong et al., 2019).

Research conducted in Cho and Lee (2015) used a genetic algorithm solver to calibrate observations with modeling results; however, as most studies in automatic calibration of hydrologic models (Gupta et al., 1999; Confesor and Whittaker, 2007) they used a semi-distributed model that cannot account for some important hydrodynamic features such as backwater effects or hydrologic characteristics such as spatial distribution of soil moisture and pollutants inside the subcatchments. Other recent research using the SWMM model attempts to develop automatic calibration algorithms for semi-distributed models, as shown in Behrouz et al. (2020) and Swathi et al. (2019).

The research conducted in Hong et al. (2019), however, considered a physically based fully distributed model that assumes various wash-off processes such as detachment and transport of particulate, resulting in six wash-off parameters that, in addition to the water quantity model parameters, must be calibrated altogether. This dramatically increases the decision variable space and might result in relatively longer simulations, as well as increase the chance of finding a different set of parameters that could explain the modeling results within a defined physically-based parameter range (i.e., Equifinality effect Beven and Freer, 2001). In addition to calibrating the model, an essential part is the model validation/evaluation that can be done to understand the model capability to represent the behavior of the system outside the calibration range (Shen et al., 2022).

The validation of conceptual lumped-parameter hydrological models as the ones presented in Shen et al. (2022) does not require detailed description of the surface topography. For fully distributed models, however, if all parameters are correct but the Digital Elevation Model (DEM) does not allow proper continuity of the flow, the model performance is affected. Poor DEMs are one of the limiting factors of applying distributed models. DEMs usually contain noises, bridges, and are affected by vegetation (Hawker et al., 2018). Raster-based algorithms such as HydroPol2D might be hydraulically affected by such issues in the DEM, requiring a pre-processing filtering to allow proper flow paths and continuity. To this end, filtering algorithms to smooth DEM flow paths (Schwanghart and Scherler, 2014), remove vegetation noises (De Paiva et al., 2013), reduce sharp elevations (Conrad et al., 2015), or smooth hillslopes (Milledge et al., 2009), can be applied to treat develop more accurate DEMs suitable for flood inundation modeling.

1.1. Paper objectives and contributions

As shown above in the literature, although several studies successfully calibrated distributed hydrodynamic models, most of the calibration studies were performed manually. With advances in computational processing and parallel computing, models that take advantage of these techniques can be applied and used for automatic optimization-based calibration. Only a few studies developed automatic calibration algorithms for fully distributed, high-resolution, hydrological-hydrodynamic models. This is likely due to complexity of due to the complex computational models with a high number of cells, states, and relatively high nonlinear underlying physical laws.

The objective of this paper is to derive a flexible framework to apply a formal HydroPol2D (Hydrodynamic and Pollution 2D model) calibration-optimization problem (i.e., a 2D distributed water quantity and quality model) using only source data at observed gauges as input. Although we use HydroPol2D in this paper, the methods developed here are valid for any other hydrodynamic model used to estimate information at gauge stations.

More specifically, in this paper, we develop a modeling framework that calibrates HydroPol2D for gauged information such as hydrographs and/or stage hydrographs and/or pollutographs using rainfall, initial soil moisture and initial water surface depths from the warm-up process as initial conditions for the model. All other hydrological-hydrodynamic and pollutant transport and fate parameters can be automatically obtained by the calibrator module developed in this paper. The method is of particular interest in catchments that already have point source gauged data in observed nodes (e.g., outlet), and these data can be used to generate spatiotemporal information within the catchment by running a calibrated HydroPol2D model (Brath et al., 2004).

The fundamental contributions of this paper are:

- We develop an automatic calibration routine to estimate HydroPol2D parameters requiring only point-source information such as depths, flows, or pollutant concentrations.
- We provide a framework capable of calibrating HydroPol2D for various events for water quantity and/or water quality modeling.

- We improved the model presented in Gomes et al. (2023a) by allowing not only a Von-Neumann (4-D) cell topology but also by adding a Moore grid (8-D) topology (i.e., four in the cardinal direction and four in the diagonal directions), allowing cells to have more flow directions, eventually decreasing the flow paths.
- We also expand Gomes et al. (2023a) by allowing raster-based DEM pre-processing algorithms to enhance flow continuity in coarse-resolution DEMs.
- We explore the equifinality problem for flood and water quality modeling.
- We explore the effects of calibrating a 2D hydrologic-hydrodynamic model with only outlet gauge data in comparison with a scenario where more stream gauges are available within the catchment.

2. Model background

HydroPol2D model is a 2D model hydrodynamic and water quality transport and fate model. The watershed is discretized into finite cells with known spatial resolution Δx and the equations of conservation of runoff mass, momentum and conservation and transport of pollutant mass are all solved matrix-wise. For a more complete description of the model, please refer to Gomes et al. (2023a).

The mass balance equation is solved for each finite element of the domain. For a cell (i, j) in the 2D meshgrid of catchment, the runoff mass balance is given by:

$$h^{t+\Delta t} = h^t + \Delta t\left(\sum Q_{\text{in}}^t - \sum Q_{\text{out}}^t + (r^t - f^t) + S^t\right), \tag{1}$$

where h is the water depth [L], Δt is the computational time-step, Q_{in}, Q_{out}, r, f, and S are the inflow, outflow, rainfall, infiltration rate, and sink/source discharge rates, respectively [$\text{L}\cdot\text{T}^{-1}$].

The infiltration capacity is estimated with an explicit Green-Ampt formulation such that (Green and Ampt, 1911):

$$C^t = k_{\text{sat}}\left[\frac{(h^t + |\psi|)\Delta\theta}{F^t}\right], \tag{2}$$

where C is the infiltration capacity [$\text{L}\cdot\text{T}^{-1}$], Δt is the time-step [T], k_{sat} is the saturated hydraulic conductivity [$\text{L}\cdot\text{T}^{-1}$], ψ is the wetting front suction head [L], $\Delta\theta$ is the moisture deficit [-], and F is the cumulative infiltrated depth into the soil media [L].

The evolution of the infiltrated depth is derived from a soil-matrix mass balance considering the infiltration rate at the surface (i.e., the minimum between the infiltration capacity and the inflow rate) and the exfiltration through the groundwater replenishing rate (f_g), resulting in:

$$F^{t+\Delta t} = F^t + \Delta t\left[\min\left(f^t,\ i^t + Q^t + h^t/\Delta t - f_g\right)\right], \tag{3}$$

where f_g is the groundwater replenishing rate [$\text{L}\cdot\text{T}^{-1}$] as function of the soil k_{sat} and moisture deficit (Rossman et al., 2010; Gomes et al., 2023b, 2024).

The remaining terms of Eq. (1) ($\sum Q_{\text{in}}^t$ and $\sum Q_{\text{out}}^t$) are fundamentally derived from momentum equations to the flow direction cartesian axes, as in other models such as LISFLOOD (Van Der Knijff et al., 2010) or HEC-RAS (Brunner, 2016). Current version of HydroPol2D accounts for local-inertial, (Bates et al., 2010), kinematic-wave, and diffusive-like SWE approximations. In this paper, the latter is used. To ensure relatively faster computational times, a cellular automata approach is applied in the diffusive-wave solution by only calculating the friction slope for the steepest water surface elevation gradient only once per cell (Gomes et al., 2023a; Guidolin et al., 2016). This approximation ensures faster simulations and allows the application of heuristic calibration that relies on extensive exploration of the decision space.

The cellular automata algorithm $\mathcal{F}$ collects the states at time t and, according to heuristic mass-conservative rules, estimate flow discharges based on available void volumes from neighborhood cells using a weighted cellular automata approach (Guidolin et al., 2016). The total outflow rate of a cell (i, j), assuming a Von-Neumann grid, can be written as:

$$\sum_{m=1}^{n_{\text{b}}} Q_{\text{out,m}}^t = \mathcal{F}\left(s_{\text{f,max}}^t, n,\ h_0,\ V^t,\ V_{(i-1,j)}^t,\ V_{(i+1,j)}^t,\ V_{(i,j-1)}^t,\ V_{(i,j+1)}^t\right), \tag{4}$$

where n_{b} is the number of neighboring cells, n is the Manning's roughness coefficient ($\text{T}\cdot\text{L}^{-1/3}$), h_0 is the initial abstraction [L], and V is the void volume from the central cell to the neighborhood cells [L^3].

By using a flow direction matrix derived from the water surface elevation map at time t, one can define which cells receive runoff from the neighbors, such that:

$$\sum_{m=1}^{n_{\text{b}}} Q_{\text{in,m}}^t = \mathcal{G}\left(Q_{\text{out,(i-1,j)}},\ Q_{\text{out,(i+1,j)}},\ Q_{\text{out,(i,j-1)}},\ Q_{\text{out,(i-1,j+1)}}\right), \tag{5}$$

where $\mathcal{G}$ estimates the total inflow rate in cell (i, j) based on the outflow rates of the neighborhood cells and based on the flow directions.

Both equations Eq. (4) and (5) can be easily expanded for a Moore grid (Torres et al., 2022) by adding the flow components in the non-cartesian axes. For a detailed mathematical formulation, including a pseudocode of the cellular automata algorithm, please refer to the supplemental material (Gomes et al., 2023a).

During dry weather periods, the initial mass of pollutant available in the catchment domain varies according to each Land Use and Land Cover (LULC) classification (Rossman et al., 2010), and can be calculated using the build-up equation, for a cell (i, j) in the domain, as follows:

$$B^{t+\Delta t_{\text{d}}} = C_1\left(1 - \text{e}^{C_2\times\Delta t_{\text{d}}}\right) + B^t, \tag{6}$$

where B is the pollutant mass [M], C_1 is the buildup maximum accumulation ($\text{M}\cdot\text{L}^{-2}$), C_2 is a fitted decreasing factor that varies with the pollutant simulated [T^{-1}] and Δt_d is the antecedent dry days duration [T] measured within two consecutive rainfall events where pollutant buildup is assumed to increase (Deletic and Orr, 2005; Gomes et al., 2021).

During wet-weather periods, the total pollutant wash-off rate leaving the cell (i, j) to downstream cells is given by:

$$\sum_{m=1}^{n_{\text{b}}} W_{\text{out,m}}^t = \Delta t \sum_{m=1}^{n_{\text{b}}} \left[C_3 Q_{\text{out,m}}^t (h^t)^{C_4} B^t\right], \tag{7}$$

where $\sum_{m=1}^{n_{\text{b}}} \boldsymbol{W}_{\text{out}}^t(t)$ is the total wash-off rate [$\text{M}\cdot\text{T}^{-1}$] that leaves cell (i, j) for directions from 1 to n_{b}, C_3 is the wash-off coefficient [$(\text{L}\cdot\text{T}^{-1})^{C_4}\cdot\text{T}^{-1}$], and C_4 is the wash-off exponent [-].

Similarly to Eq. (5), a flow direction algorithm is used to determine which neighborhood cells contribute to incoming pollutant that enters cell (i, j) for a Von-Neumann grid, such that:

$$\sum_{m=1}^{n_{\text{b}}} \boldsymbol{W}_{\text{in}}^i(t) = \mathcal{G}\left(W_{\text{out,(i-1,j)}},\ W_{\text{out,(i+1,j)}},\ W_{\text{out,(i,j-1)}},\ W_{\text{out,(i-1,j+1)}}\right), \tag{8}$$

By using an Eulerian forward finite difference scheme, we can derive a pollutant mass balance dynamical equation, such that:

$$B^{t+\Delta t} = B^t + \Delta t\left(\sum_{m=1}^{n_{\text{b}}} \boldsymbol{W}_{\text{in}}^i(t) - \sum_{m=1}^{n_{\text{b}}} W_{\text{out,m}}^t + S_b\right), \tag{9}$$

where S_b is the source/sink pollutant rate [$\text{M}\cdot\text{T}^{-1}$].

The dynamical Eqs. (1), (3), and (9) are solved for the total simulation timespan.

2.1. Decision variables in HydroPol2D automatic calibration problem

In HydroPol2D, parameters are spatially derived as a function of the LULC and Soil rasters. The minimum and maximum values of each decision variable are entered for each classification of these rasters. Let n_l be the number of land use classifications and n_s be the number of soil classifications in a catchment. Also, let n_v^l be the number of decision

variables related to the LULC map and n_v^s be those related to the soil map; therefore, we have a decision vector $\boldsymbol{x} \in \mathbb{R}^{n_x}$ with $n_x = n_l n_v^l + n_s n_v^s$.

We classify the decision variables into water quantity variables (superscript r), water quality variables (superscript w), and soil related parameters (subscript s). Furthermore, we classify the variables as LULC-based and soil-based into subscripts 1 to n_l and 1 to n_s, respectively. The decision variable of the calibration problem can be written as $\boldsymbol{x} = [\boldsymbol{x}_l^r,\ \boldsymbol{x}_l^w,\ \boldsymbol{x}_s^r]^{\mathrm{T}}$, such that:

$$\boldsymbol{x}_l^r = [n_1,\ \ldots, n_{n_l},\ h_{0,1},\ \ldots,\ h_{0,n_l}]^{\mathrm{T}} \tag{10a}$$

$$\boldsymbol{x}_l^w = [C_{1,1},\ \ldots, C_{1,n_l},\ C_{2,1},\ \ldots,\ C_{2,n_l}, C_{3,1}, \ldots C_{3,n_l},\ C_{4,1},\ \ldots C_{4,n_l}]^{\mathrm{T}} \tag{10b}$$

$$\boldsymbol{x}_s^r = [k_{\mathrm{sat},1},\ \ldots, k_{\mathrm{sat},n_s},\ \Delta\theta_1,\ \ldots,\ \Delta\theta_{n_s}, \psi_1,\ \ldots \psi_{n_s}]^{\mathrm{T}}, \tag{10c}$$

2.2. Initial values for the simulation

In addition to the parameters, in the automatic calibrator of the HydroPol2D model, we can enter initial maps of water surface depth and soil moisture to represent the initial conditions for simulation for each event accurately. The code is designed to calibrate a maximum of 10 events with 10 observation points of discharge, water depth, and pollutant concentration within the catchment. Another input map that could be entered is the distributed pollutant mass in the catchment domain before the simulation; however, in this paper, we opted to calculate this mass in terms of the antecedent dry days and assume that it is uniformly distributed according to the land use classification due to the lack of observed data.

2.3. Fitness functions

In this section we show the fitness functions allowed in the automatic calibrator. These functions can be used to calibrate hydrographs, stage-hydrographs, and pollutographs at the catchment outlet. Detailed mathematical descriptions of these functions are available in the Supplemental Material. In this paper, we use the Nash-Sutcliffe-Efficiency (NSE) (Nash and Sutcliffe, 1970), the coefficient of determination (r^2), the root-mean-square-error (RMSE) (Fisher et al., 1920), the Peak Flow Bias, and the relative runoff volume error.

In addition to the previously defined functions, users can have the flexibility to write their fitness functions since all codes are open source. Some examples that are predefined in the model and not fully detailed here for the sake of parsimony are the (i) mean average error, (ii) event mean concentration, (iii) PBIAS, and (iv) runoff volume mean error.

2.4. Optimization constraints

The optimization problem is subject to four constraints. First, the model is constrained to HydroPol2D dynamical model that has conservation of mass and momentum constraints, and pollutant transport and fate dynamics. Moreover, the optimization problem can also have equality constraints (e.g., the case where a parameter is a linear combination of other parameters). Finally, we can set the minimum and maximum parameter ranges for each decision variable.

2.5. Objective function

We can calibrate a single or multiple events together. Therefore, let the index j represent the jth event used for calibration. Moreover, let f collect fitness functions such that a possible scenario of objective functions could be $f_1 = -\mathrm{NSE}$, $f_2 = -r^2$, $f_3 = -\mathrm{RMSE}$, and $f_4 = \eta_q$, for example. The problem is set as a single-objective minimization problem, and the objective function used in this paper can be written as a function of linear combinations between each fitness function for each event, such that:

$$O_f = \sum_{j=1}^{n_e} \beta_j \left[\left(\sum_{i=1}^{n_f} \alpha_i f_i \right) \right], \tag{11}$$

where n_f is the number of fitness functions used in the optimization α defines the relative weight of each function f_i, n_e is the number of events used in the calibration, and B_j is the weight given by the objective function values for each event.

It is important to mention that the factor α_i must be such that it normalizes the varied objective functions used to avoid over-weighting in fitness functions with different scales of magnitude and units.

2.6. Automatic calibration optimization problem

In this section, we define the automatic calibration optimization problem. Although the nature of hydrological model calibration can be inherently multiobjective (Shafii and De Smedt, 2009), for the sake of parsimony and to allow practical application, we focus on developing a single objective automatic calibration problem. It can be written by minimizing the objective function, subject to HydroPol2D dynamics. An example of calibration problem such that we can write the problem as:

$$\min_{\boldsymbol{x}}\ O_f = \sum_{j=1}^{n_e} \beta_j \left[\left(\sum_{i=1}^{n_f} \alpha_i f_i \right) \right] \tag{12a}$$

$$\text{s.t. HydroPol2D Dynamics in Eqs. (1)-(9)} \tag{12b}$$

$$\boldsymbol{A}_{\mathrm{eq}} \boldsymbol{x} = \boldsymbol{B}_{\mathrm{eq}} \tag{12c}$$

$$\boldsymbol{x}^l \leq \boldsymbol{x} \leq \boldsymbol{x}^m, \tag{12d}$$

where $\boldsymbol{x}^l$ and $\boldsymbol{x}^m$ are the lower and upper bounds of the decision vector $\boldsymbol{x}$.

The problem posed in Eq. (12) is non-linear and non-convex. The use of evolutionary strategies such as the Shuffled Complex Evolution (SCE) (Naeini et al., 2019) has been used for this type of calibration problems (Tigkas et al., 2016). In Matlab, several solvers are available to solve problems as Eq. (12), such as *Global-Search*, *Patter-Search* or *Genetic-Algorithms* (GA) (Higham and Higham, 2016). Herein, we choose GA due to its flexibility to deal with non-linear problems (Giacomoni and Joseph, 2017) and its ability to use parallelization in Matlab. Other software, such as SWMM and HEC-HMS, have applications that follow a similar approach to the one presented in this paper (Behrouz et al., 2020; Dariane et al., 2016). The GA is a population-based probabilistic optimization method that emulates the principles of genetics and natural selection (Tigkas et al., 2016). Since the goal of hydrologic-hydrodynamic model calibration is not essentially finding the global optima but a physically possible set of parameters trying to avoid equifinality, we set all problems to run with a relatively low number of generations, but still with a relatively large population size to allow a proper exploration of the decision variable space.

The genetic algorithm is defined for a certain number of generations and population size. By population we mean a set of feasible individuals. An individual is a feasible (e.g., within the minimum and maximum expected values) set of the model parameters, and each model parameter is the genes of each individual. During the generations, only the 50% best individuals (i.e., with the best fitness function) of the population survive, recombine, and mutate, leading to the novel population set for the next generation. The process is repeated until a convergence criterion is met or if the maximum number of generations is reached (Giacomoni and Joseph, 2017).

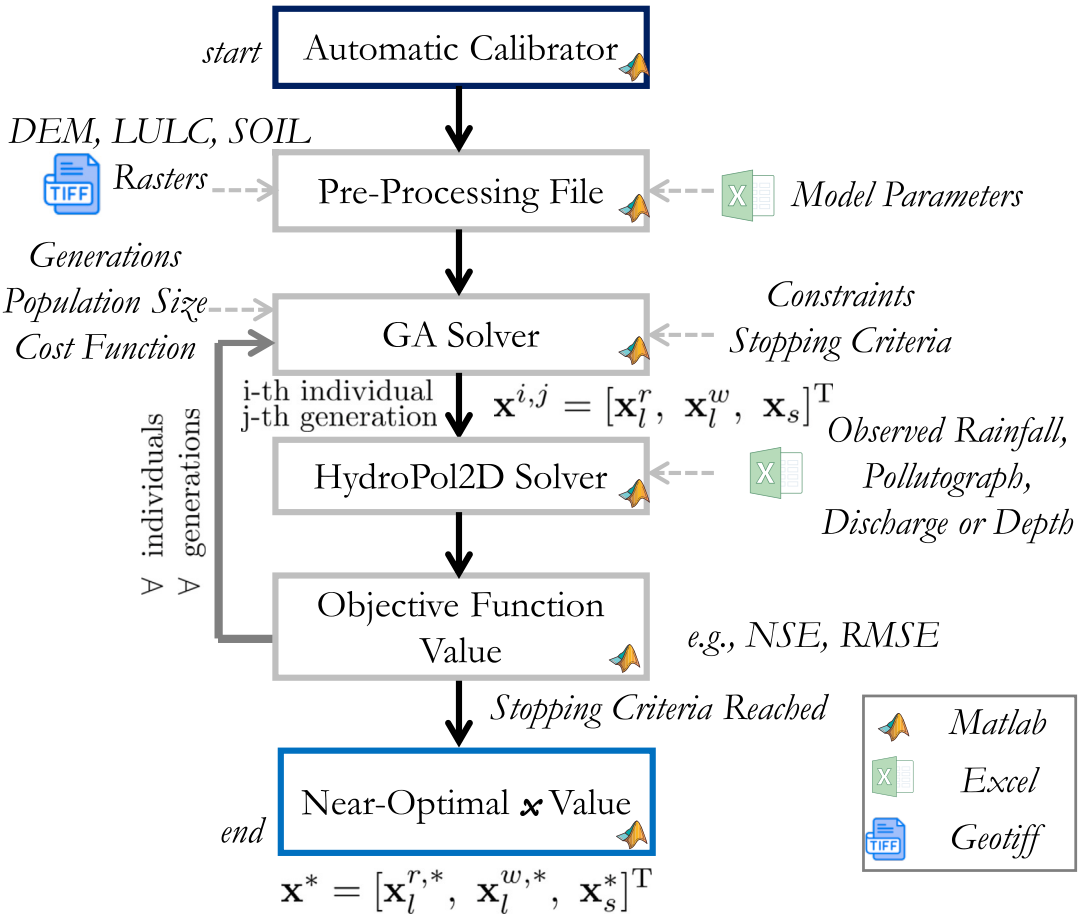


Fig. 1. Automatic calibrator flowchart. First, the model reads the automatic calibration inputs, and then it runs a pre-processing file, defining the required numerical input for HydroPol2D, such as the DEM, LULC, and Soil Maps. Following this phase, the model runs the GA solver, which uses HydroPol2D to estimate the objective function values, and this process is looped until the stopping criteria are reached.

2.6.1. Genetic algorithm properties

We set the problem to run for (10-40) generations with a 100 population. The stopping criteria were twofold: (i) simulation would stop if the number of maximum generations is reached, (ii) if no improvement in the objective function was found in 30 min. All genetic algorithm parameters are set as standard values from Matlab (Higham and Higham, 2016). A flowchart of the optimization process is presented in Fig. 1.

2.7. Input data

To run the automatic calibration procedure defined in this paper, it is necessary to build a HydroPol2D model and to enter the observed input data and climatologic forcing. All HydroPol2D model input data are entered in excel user interfaces allowing users to control time-stepping parameters, DEM, LULC, and SOIL raster directories, and etc. For detailed explanation on how to set up the model, please refer to the model manual in Gomes Jr. (2024). In addition to the built model, users then have to fill an excel spreadsheet with gauge coordinate and at least one time series of discharge and pollutant concentration. Screenshots of the interfaces to enter the data are presented in the SI.

3. Case studies

3.1. Numerical case study 1 - V-tilted catchment

The objective of this numerical case study is to test the ability of the optimization model to predict Manning's roughness coefficient of the catchment and to check if the model can predict that there are no initial abstractions and infiltration in this case study. Essentially, we want to answer the following question:

- Q1: *Does the automatic calibration algorithm can identify the Manning's roughness coefficients of hillslopes and main channel, as well as the initial abstraction values of these land uses? In addition, can it identify whether infiltration is being considered in this case study?*

We choose the V-Tilted catchment as a virtual experiment inverse problem (Fatichi et al., 2016). The catchment has 4050 cells of 20 x 20 m in size and has a reasonable fast computation, allowing the use of metaheuristics that rely on multiple computations of the objective function. To answer Q1, we define the decision vector of this problem as:

$$\boldsymbol{x} = [n_1, n_2, h_{0,1}, h_{0,2}, k_{\mathrm{sat},1}, \Delta\theta_1, \psi_1]^{\mathrm{T}},$$

where subscripts 1 and 2 are the LULC classifications in the catchment (i.e., 1 are the hillslopes, and 2 is the channel, see Fig. 2(a)). The optimization problem mathematical description of this case is presented in Eq. S1. It is assumed that there is only one type of soil in the catchment, such that $n_s = 1$. We add an equality constraint in Eq. S1 by entering the known parameters to set only the water quantity variables as decision variables. In this problem, we choose the NSE as the objective function since we are focused on calibrating the modeled flow discharge with the observed discharge at the outlet. This case study has no infiltration or initial abstraction and is a reverse problem since we know the right parameters (Kollet and Maxwell, 2006); however, we decided to include infiltration variables in the optimization problem formulation to see if the algorithm can identify this condition.

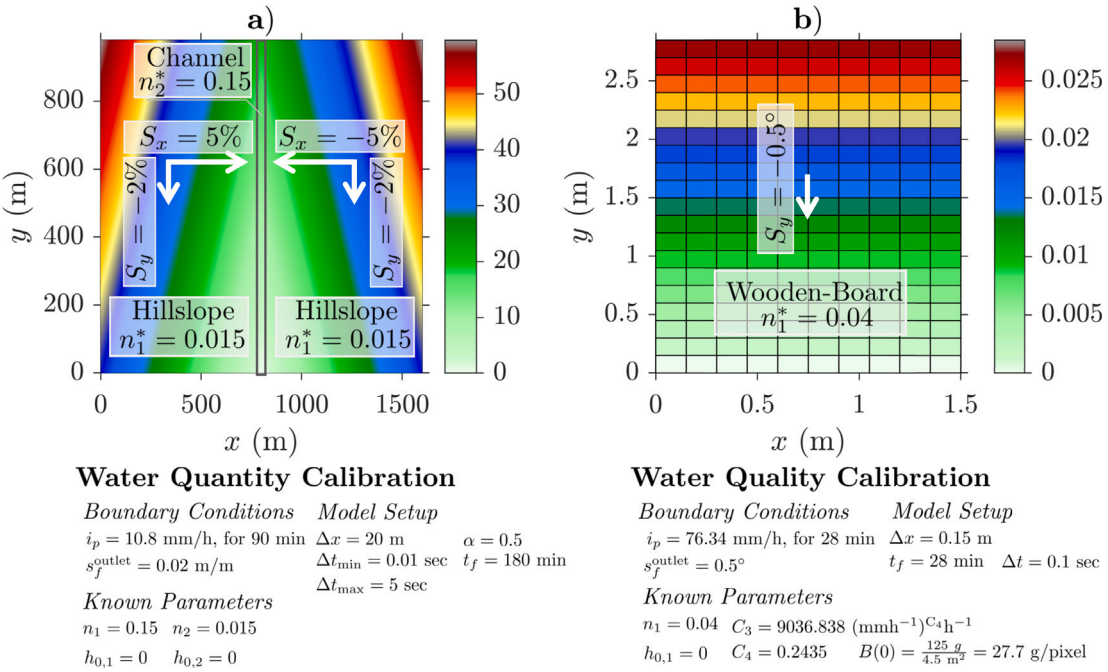


Fig. 2. Numerical Case Study 1 and 2 digital elevation models in meters. Part (a) is the V-Tilted Catchment, whereas part (b) is the Wooden-Board catchment. The parameters of each case are shown in the figure, where i_p is the constant rainfall rate, s_f^{outlet} is the friction slope at the outlet, Δx is the pixel size, Δt is the constant time-step assumed, t_f is the end of the simulation, and α is the Courant number.

This is a problem with a relatively short decision space, in which equifinality effects are hypothesized to be minimized.

3.2. Numerical case study 2 - Wooden-Board catchment - pollutant concentration

The objective of this numerical case study is to test the ability of the optimization model to predict the salt concentrations at the outlet of the catchment, the initial salt mass, and the wash-off parameters of the model. This is a fairly more complex optimization problem if not only the water quantity parameters are required to calibrate, but also the water quality ones. In addition, the water quality parameters have a wider sensitivity, as shown in Gomes et al. (2023a). The optimization problem mathematical description of this case is presented in Eq. S2. In this problem, we assume that the water quantity parameters (i.e., n and h_0) are already calibrated (Zhang et al., 2020), so that the decision vector for this problem is $\boldsymbol{x}_l^w = [C_{1,1},\ C_{2,1},\ C_{3,1},\ C_{4,1}]^{\text{T}}$. To set only the water quality variables as the decision variables, we add an equality constraint in Eq. S2. In this problem, we only choose RMSE as the fitness function.

This case study is a controlled experiment in a wooden-catchment as shown in Fig. 2(b). The wooden-board has 4.5 m^2 and 300 cells. The initial mass of the solute is 125 g, and it is assumed that it is uniformly distributed in the catchment area (Zhang et al., 2020). However, in this paper, we do not assume that the initial solute mass is known, and we let the model search for the near-optimal solute mass considering C_1 and C_2 as decision variables. Naturally, this controlled experiment is not a direct case of applying the build-up equation that calculates the available mass of the pollutant in terms of the Δt_d, as shown in Eq. (6). However, fixing $\Delta t_{\text{d}} = 10$ days, for example, we can estimate C_1 and C_2, calculate the initial build-up by solving Eq. (6) and compare with the initial mass of 125 g known from the experiment. Ultimately, what matters for HydroPol2D is the initial pollutant mass available in each domain cell. The reason we consider build-up as a function of Δt_d and LULC is that in most cases, the initial pollutant mass varying cell-by-cell is unknown, and these direct measurements are either intractable, hardly ever available, and would result in an intractable decision-space if all cells are treated individually in the optimization problem. The parameter ranges were estimated using a 60% variation from the previous calibrated ones, assuming that the initial pollutant mass was 125 g (Gomes et al., 2023a). In this Numerical Case Study, we want to answer the following question:

- Q2: *Assuming the water quantity parameters known, can the algorithm find the initial mass of salt (build-up model parameters) and the wash-off parameters to match with the observed pollutograph at the outlet?*

3.3. Numerical case study 3 - Gregorio catchment in Sao Carlos / Brazil

The Gregório catchment is located in the municipality of São Carlos in the state of São Paulo, Brazil. The climate in the state of São Paulo is influenced by Atlantic Tropical and Continental and Atlantic Polar air masses, complemented by Continental Equatorial air masses coming from the Western Amazon. The months with the largest rainfall events are in summer, from October to March, and the dry weather period varies from April to September in winter. The average annual precipitation of the city of São Carlos is approximately 1492 mm (de Meteorologia, 2022) and the city has been prone to critical rainfall events yearly (Abreu, 2019). The catchment area is 18.64 km^2, the length of the main channel is 8.6 km and its morphological characteristics indicate an elongated to strongly elongated catchment, which presents a compactness coefficient (C_c) of 2.030, a circularity ratio (R_c) of 0.120 and a form factor (R_f) of 0.289, as shown in Fig. 3. Although the morphometric characteristics would indicate a resilient catchment to floods, the large impervious rate, mean slope, and the channelization of the main creek increase the vulnerability of the area in terms of floods.

To perform hydrodynamic modeling, we built maps of the Digital Elevation Model (DEM), Land Use Land Cover (LULC) and Soil Texture as presented in Fig. 4. Due to the lack of high-resolution data in the catchment, we use freely available worldwide datasets for all input maps; therefore, the methods applied here are replicable in other poorly-gauged catchments (Gomes et al., 2023a). However, when available, higher resolution maps can be used.

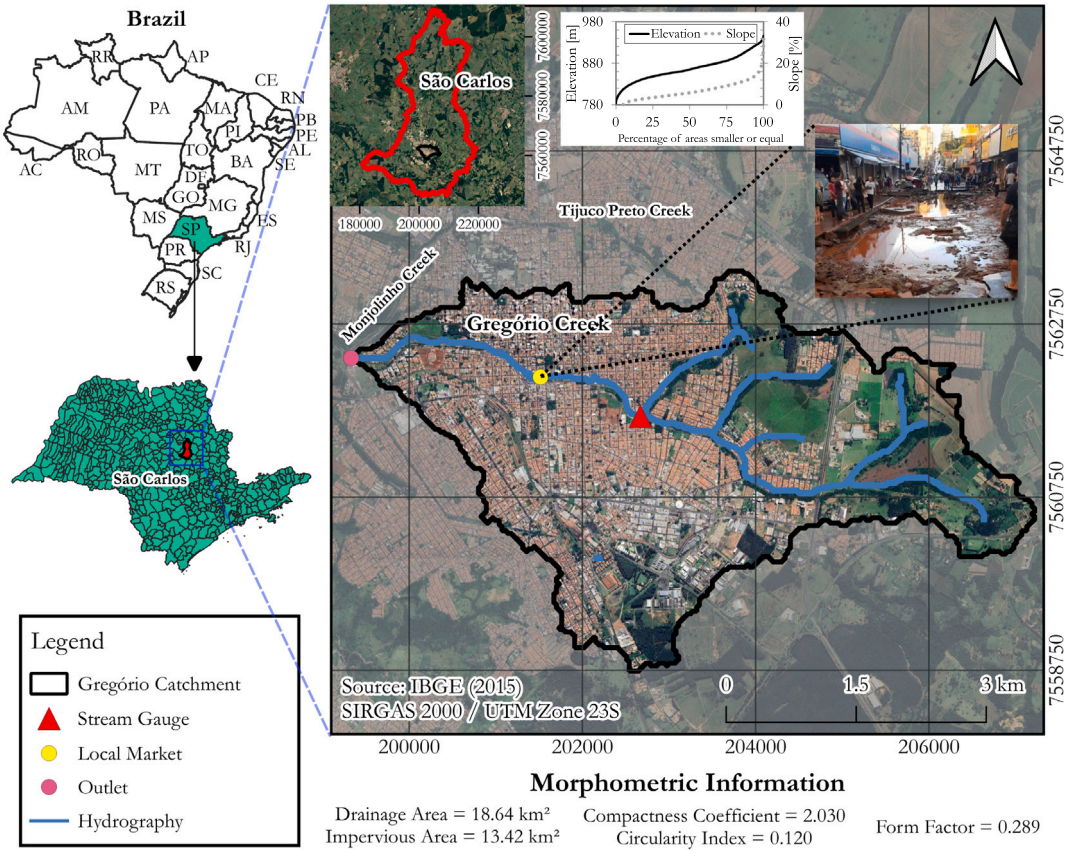


Fig. 3. Gregorio Catchment location map with hypsometric curves of elevation and slope, and a figure of the flood-related impacts in the Local Market point of Sao Carlos.

The pedology of the catchment is composed of yellow-red latosoil (YRL) and small areas of purple latosoil (PL) (de Geografia e Estatística (IBGE), 2022). The soil texture within the catchment can be classified into medium and clayey texture (de Geografia e Estatística (IBGE), 2022) (see Fig. 4). The headwaters of Gregorio catchment remains relatively undeveloped, with a predominance of pervious areas with crops, grass, and shrub areas. Downstream the creek, impervious rates dramatically increase with urbanization, which almost makes the catchment impervious towards the outlet, as presented in Figs. 3 and 4. This catchment covers large proportion of the urbanized area in the municipality of São Carlos with the most commercial activities of the city being carried out in this area. For this reason, due to the climatic and hydraulic characteristics of the catchment, floods are constantly recorded, especially in the summer. A recent flood picture is shown in Fig. 3 in the Local Market (Abreu, 2019; Sarmento Buarque et al., 2020).

The specific question of this Numerical Case Study is:

- Q3: *Given the reality of scarce data in poorly-gauged catchments, can the algorithm find the near-optimal hydrodynamic parameters, within physical limits, to match with the observed hydrograph? Can this set of parameters be used to estimate catchment-scale information?*

3.3.1. DEM - preprocessing

A 30-m DEM might be considered high-resolution for rural catchments. For urban areas, however, the complexity of the built environment with detailed infrastructure would require a more detailed resolution. Information in such detail is typically unavailable in developing countries. Nonetheless, poorly-gauged areas are usually those that often suffer from floods (Fava et al., 2020). Raster-based flood routing models are, therefore, affected by DEM quality. Typically, DEMs are required to be hydrologically corrected, ensuring that the flow directions are continuous and connected towards the outlet. However, especially in urban areas with bridges, culverts, and stormwater reservoirs, DEMs usually have to be burned to allow proper flow directions and connection. Herein, we provide 4 algorithms to treat low-quality DEMs (i.e., the gaussian filter Young and Van Vliet, 1995), the constrained regularized smoothing of the channel length profile (CRS) (Schwanghart and Scherler, 2014), and the method of reducing the DEM elevation to consider water surface depths based on De Paiva et al. (2013). These methods are detailed in the Supplemental Material.

3.3.2. Data collection

Rainfall intensity and depth of the water surface are recorded in a limited way each minute and upscaled to 5 min intervals, and the rainfall and stream gauge station is shown in Fig. 3. From the monitoring campaign provided in Souza (2008), only one event had a sufficiently large rainfall volume and quality discharge observations. To collect data, a Campbell Scientific® CR10 station was installed and calibrated to record the data to be collected after each rainfall event (Souza, 2008). A calibrated rating curve (Gomes et al., 2023c) converts water depth into flow discharge by the following relationship (Lima et al., 2007):

$$Q(h) = 8.278h^{2.2517},\ r^2 = 0.99, \tag{13}$$

where Q is the observed flow discharge at the gauge station, and h is the measured water depth taken from the channel invert.

The recorded level was converted into flow discharges using Eq. (13) and used for the calibration of HydroPol2D.

3.3.3. Initial conditions

We assumed the initial soil moisture in the soil calculated with the cumulative rainfall prior to the event coupled with the SCS-CN (SCS, 1986) infiltration model. A Curve-Number map was developed by Souza (2008) to estimate spatial infiltration in pervious areas.

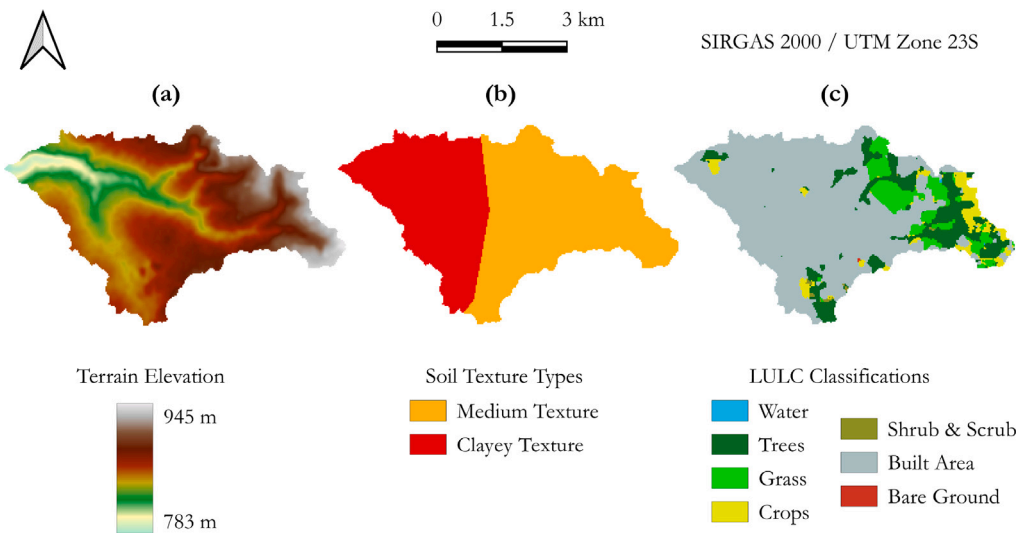


Fig. 4. Input maps to the hydrodynamic simulation in HydroPol2D for the Gregorio catchment, where (a) is the digital elevation model, (b) is the soil texture map, and c) is the land use and land cover map.

Table 1
Parameter ranges and calibrated values for the LULC-Based parameters of HydroPol2D for Numerical Case Study 3.

Classification	n_{min} [$s \cdot m^{-1/3}$]	n_{max} [$s \cdot m^{-1/3}$]	n [$s \cdot m^{-1/3}$]	$h_{0,min}$ [mm]	$h_{0,max}$ [mm]	h_0 [mm]
Water	0.0250	0.0400	0.0400	0.00	0.00	0.00
Trees	0.0250	0.0400	0.0268	0.00	10.00	0.8259
Grass	0.0200	0.0350	0.0244	0.00	5.00	0.2740
Flooded vegetation	0.0250	0.0400	0.0381	0.00	10.00	1.2281
Crops	0.0200	0.0350	0.0236	0.00	10.00	0.1564
Scrub/Shrub	0.0300	0.0400	0.0358	0.00	10.00	7.6289
Built areas	0.0150	0.0300	0.0216	0.00	2.00	0.0625
Bare ground	0.0200	0.0300	0.0260	0.00	2.00	1.6101

For the initial water surface depth, previous modeling results indicate that no warm-up is necessary, and current visits to the study area indicate a minimum effect of non-hortonian flows. It is also seen from the observed hydrographs that the initial flow is null, indicating an intermittent creek.

3.3.4. Boundary conditions and running control parameters

The model is simulated with a space-invariant and time-variant rainfall hyetograph distributed to all cells of the grid. In addition, we assume a gradient outlet boundary condition at the outlet with the friction slope $s_f = 0.02$ m · m. Although we assume normal flow at the catchment outlet, the flow is considered transient in the gauging station as it is an internal domain node (see Fig. 3). To guarantee numerical stability, we define minimum and maximum time steps of 0.1 and 5 sec, respectively, and we set HydroPol2D model to change time-steps each 60 s of the simulation time. We use a Courant number bound of 0.4, so time-steps are adapted to match this input (Gomes et al., 2023a). The model is run for 120 min of simulation, and point and raster results are retrieved each 5 min.

3.3.5. Parameter ranges

The parameter ranges used for calibration for the construction of the calibration optimization problem of Eq. (12) are given in Tables 1 and 2. The physically bounds used in this paper were derived from the literature and recently published papers, and manuals (Soliman et al., 2022; Rossman et al., 2010; Brunner, 2016).

3.3.6. Sensitivity analysis

A one-at-the-time sensitivity analysis is performed in the model to identify the most sensitive parameters before the automatic calibration procedure (Gomes et al., 2023b). We define three output functions and calculate the variance of each perturbation in the decision variables in terms of the variance in the output functions. We evaluated the output variance of Peak Flow, Runoff Volume, and Time to Peak, as they are closely related to the hydrograph properties. We also assessed flood areas. More details of the output functions are found in the Supplementary Material. The parameters ranged from 10% to 190% of the baseline parameters, with 10% intervals. In addition, they are defined as the arithmetic average of the parameter ranges presented in Tables 1 and 2.

3.4. Numerical case study 4 - exploring equifinality

In this case study, our objective is to explore the parameter equifinality problem in the calibration of the hydrological and hydrodynamic model. Equifinality tends to decrease with the number of observations and with the decrease in the model parameters (Her et al., 2019). To this end, we create a synthetic case study without uncertainty in rainfall, initial soil moisture, and observed discharge, mimicking a perfect gauging system. Therefore, the error in the parameter calibration is most likely due to equifinality, although calibration solver properties such as the number of generations, population size, or genetic algorithm properties that might change the behavior of the exploration of the decision space can play a role. The spatial-variability of rainfall is a challenge that could also be explored, but is out of the scope of this paper. To explore the equifinality problem in a scenario of certain rainfall and perfect measurements in the gauges, we formulate the following question.

- *Q4-1: How does parameter equifinality affect the calibration of HydroPol2D for different parameter ranges, number of events, the magnitude of the rainfall intensity, and location of the gauging stations?*

Therefore, we assess the near-optimal calibrated parameters in an inverse problem using the V-Tilted catchment as a surrogate case study, varying the number of gauges, the number of rainfall events, and their intensities. This catchment is used as a virtual laboratory to test the hypothesis raised in this case study (Fatichi et al., 2016). We altered the original watershed to have 3 classifications of soils and LULC, following the left hillslope (1), the middle channel (2), and the right hillslope

Table 2
Parameter ranges and calibrated values for the SOIL-Based parameters of HydroPol2D for Numerical Case Study 3.

Type	$k_{sat,min}$ [mm.h^{-1}]	$k_{sat,nax}$ [mm.h^{-1}]	k_{sat} [mm.h^{-1}]	$\Delta\theta_{min}$ [-]	$\Delta\theta_{max}$ [-]	$\Delta\theta$ [-]	ψ_{min} [mm]	ψ_{min} [mm]	ψ [mm]
Medium	1.00	10.00	1.14	0.25	0.60	0.29	0.00	230.00	33.56
Clayey	0.20	10.00	6.23	0.25	0.60	0.40	0.00	312.50	281.10

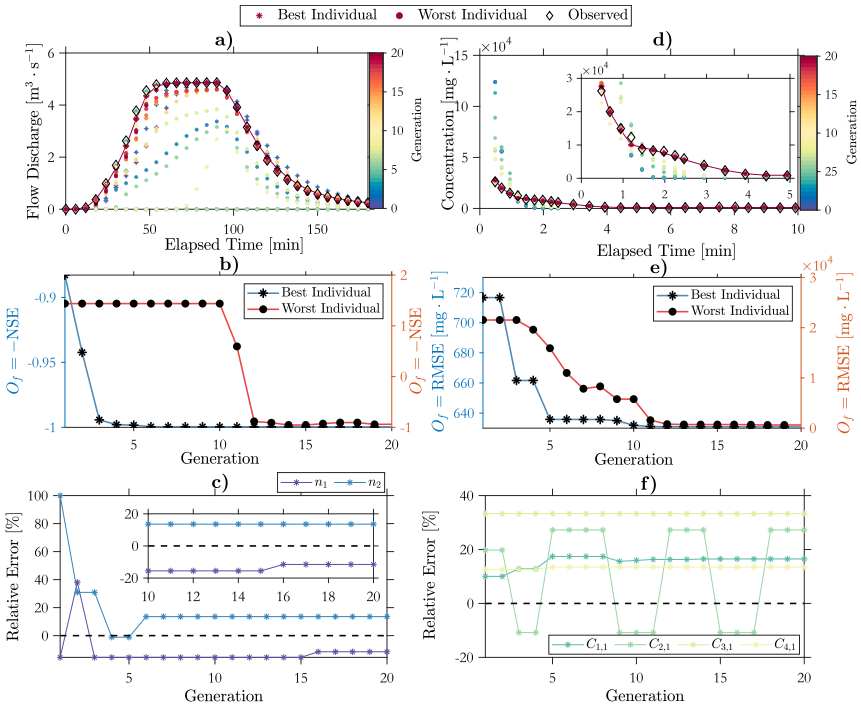


Fig. 5. Modeling Results of Numerical Case Studies 1 and 2. Parts (a) - (c) are the hydrograph, objective function chart, and relative error of parameters for the V-Tilted catchment. Parts (d) - (f) represent the pollutograph, objective function chart, and relative error chart for the Wooden-Plane catchment. Only the best and the worst individuals of each generation are plotted in (a) and (d). Only the parameters of the best individuals are plotted in (c) and (f).

(3), each of them with different n, h_0, k_{sat}, $\Delta\theta$, and ψ. The left and right hillslopes can be classified into hydrologic units with the same roughness and infiltration properties. Consequently, having a gauge station in each of the hillslopes would possibly reduce the uncertainty in the parameters.

The problem has ten unknown parameters (i.e., five parameters for each hillslope) and five known parameters (i.e., the main channel parameters are kept constant). We apply the model to calibrate three different storms of 10.8 mm · h^{-1}, 21.6 mm · h^{-1}, and 32.4 mm · h^{-1} with 90-min duration and later we calibrate only using the first event of 10.8 mm · h^{-1}. Detailed results of the modeling of each of the three rainfall events are presented in the Supplemental Material. The number of gauges (1 - outlet, 2 - left hillsope, and 3 - right hillslope) is combined, resulting in 7 possible cases (1-2-3, 1-2, 1-3, 2-3, 1, 2, and 3). The left and right gauges are defined by the channel neighborhood cell located at the half middle of the V-tilted length (i.e., y = 500 m) spanned 20 m from the channel i.e (x = 780 m, and x = 820 m, see Fig. 2(a)).

We formulate the calibration problem with a wide parameter range (see Supplemental Material) mimicking no prior knowledge of the system except by the input data that discretize the domain into 3 areas (i.e., left hillslope, channel, right hillslope). We compare the results of this case with a condition with more knowledge of the system, that is, the parameter range is half of the previous one, hence reducing the decision space.

The calibration of hydrological models is inherently multi-objective (Shafii and De Smedt, 2009). For example, minimizing RMSE might give good objective function values, correctly matching the peaks, but might fail during the recession time, thus altering the overall mass balance that is accounted for in soil moisture, for example (Lindström, 1997). To this end, we use two metrics as our composite objective function, that is, the NSE and the relative volume error. The NSE varies from $-\infty$ to 1 and the relative volume error should be minimized, such that we would want to maximize the NSE while minimizing the volume error. By introducing a penalizing factor as a function of the relative volume error in the NSE, we seek solutions with a good NSE and reduced volume errors. Therefore, the objective function also varies from $-\infty$ to 1, where 1 indicates a perfect NSE and no volume error.

To transform this hypothesis into a minimizing optimization problem, we assume that each gauge has the same importance (i.e., $\gamma = 1/n_g$, with n_g as the number of gauges), assuming the NSE with weight $\alpha_1 = 1$ and the volume error with weight $\alpha_2 = 0.5$ as well, and we assume that each event also has the same importance (i.e., $\beta_j = 1/n_e \; \forall \, j$, with n_e being the number of events). Therefore, we can write the objective function (11) as (Lindström, 1997):

$$O_f = \frac{-1}{n_e}\sum_{j=1}^{n_e}\Big[\frac{1}{n_g}\sum_{i=1}^{n_g}\Big(\alpha_1 \overbrace{\mathrm{NSE}^{i,j}}^{f_1} - \alpha_2 \overbrace{\Big|\frac{\sum_{k=1}^{T_j}(Q_m^{i,j,k} - Q_{obs}^{i,j,k})}{\sum_{k=1}^{T_j}(Q_{obs}^{i,j,k})}\Big|}^{f_2}\Big)\Big], \tag{14}$$

where i is the gauge index, j is the event index, and k is a time-step index.

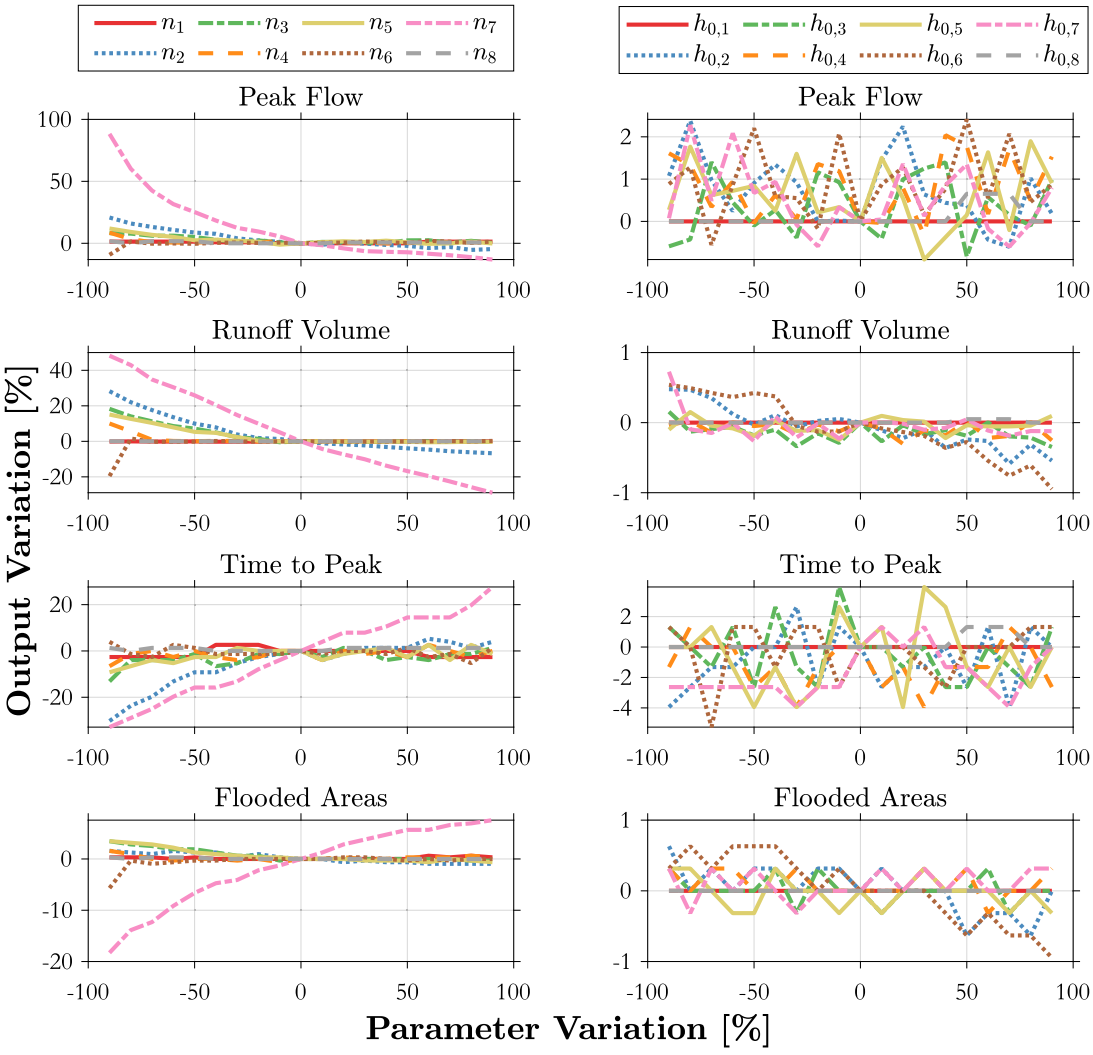


Fig. 6. One-at-the-time sensitivity analysis of the LULC-Based parameters using the average parameters from the parameter range presented in Tables 1 and 2 for Numerical Case Study 3. LULC-Based subscripts 1 = water, 2 = trees, 3 = Grass, 4 = Flooded Vegetation, 5 = Crops, 6 = Schrub/Scrub, 7 = Built Areas, and 8 = Bare Ground.

The previous objective function attempts to maximize NSE while trying to maintain important hydrological features such as volume conservation (Gomes et al., 2023b; Lindström, 1997). The negative sign in the first fraction is to transform the objective function suitable for the minimization of the optimization problem.

The calibration of hydrologic-hydrodynamic models is a necessary but not sufficient condition to apply the model under different input ranges. The validation process is usually performed with different hydrologic conditions, typically represented by storms different from those used for the calibration. We provide a validation test under different storm volumes, intensities, temporal distributions, and volumes to gain confidence in the estimated parameters. To address these issues, we formulate the following question:

- *Q4-2: Using only the observed data at the outlet, is it possible to obtain a sufficiently accurate model that can be used not only for the calibration events but also for the validation under different storms intensities, durations and temporal distributions?*

To answer Q4-2, we calibrate the model with only the outlet gauge as the source information for the optimization calibration algorithm. To ensure different rainfall characteristics, we change the durations and volumes andthe rainfall temporal distribution. The rationale is to have rainfall events with 50 or 150% values from the calibration events whenever possible to represent relatively different conditions from the calibration phase. Therefore, we alter the durations from the 90-min rainfall duration used from the calibration events, resulting in rainfall durations of either 45 or 135 min. However, reducing the rainfall volumes to 50% of the smallest rainfall event used for calibration would generate events without runoff. Therefore, for this case, we fix the intensity as 10.8 $\text{mm} \cdot \text{h}^{-1}$ but change the duration of the rainfall. To consider the effect of unsteady-state rainfall, we use the Huff 1st quartile hyetograph (Huff, 1967) as a proxy rainfall distribution to represent the temporal dynamics of the rainfall.

4. Results and discussion

4.1. Numerical case study 1

The modeling results of the V-Tilted catchment calibration problem are presented in Fig. 5(a)-(c). The near-optimal value of the decision vector is found after 20 generations with 100 population are n_1 = 0.0132 $\text{s} \cdot \text{m}^{-1/3}$, n_2 = 0.1703 $\text{s} \cdot \text{m}^{-1/3}$, $h_{0,1}$ = 0.02 mm, $h_{0,2}$ = 0.46 mm, k_{sat} = 0, $\Delta\theta$ = 0.07 $\text{cm}^3 \cdot \text{cm}^{-3}$, and ψ = 15.21 mm. The hydrographs

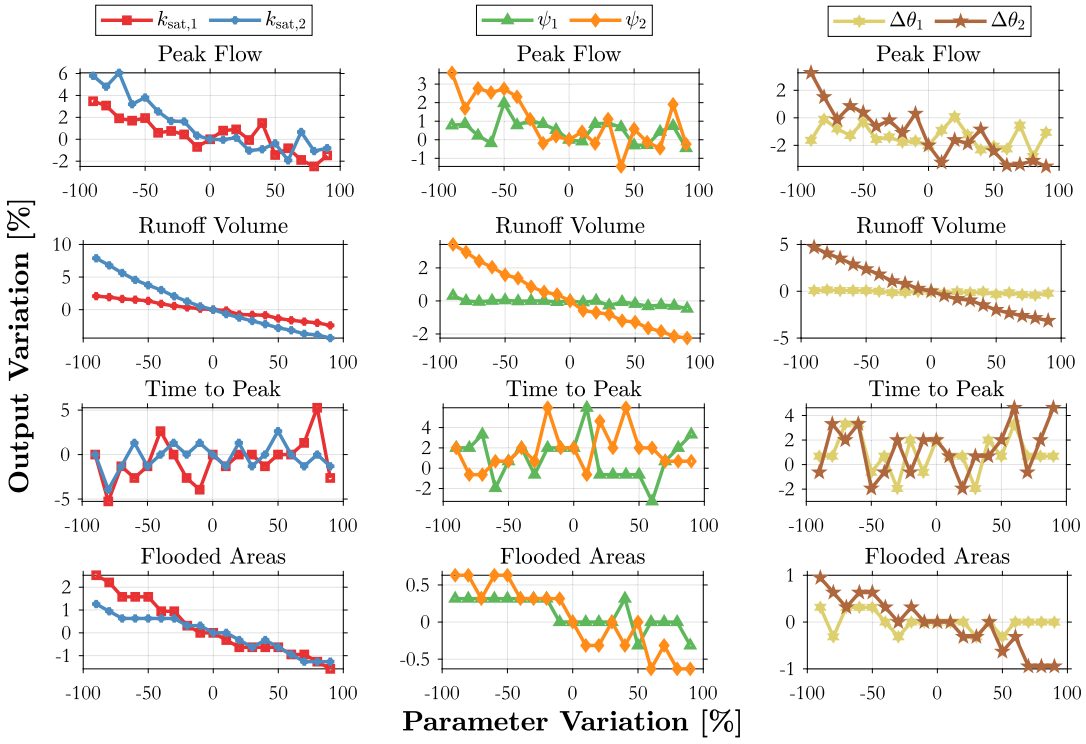


Fig. 7. One-at-the-time sensitivity analysis of the SOIL-Based parameters using the average parameters from the parameter range presented in Tables 1 and 2 for Numerical Case Study 3. Soil-Based subscripts 1 = Medium and 2 = Clayey.

at the catchment outlet are presented in Fig. 5(a), where the best and worst individual's hydrographs are plotted for each generation. By individual, we mean a feasible set of parameters following Eq. (S1) that produces an outlet hydrograph response in the catchment. Some of the worst individuals had no outflow due to large values of initial abstraction and/or k_{sat}, likely due to the decision space exploration of GA. However, as the generation moves, the worst individuals predict better hydrographs with NSE closing to unity as shown in Fig. 5(b), that shows the performance of the best and worst individuals throughout the generations. In nearly 1 generation, it is possible to note that the best individual already gets good results for hydrological models (i.e., NSE > 0.85) (Nash and Sutcliffe, 1970).

The best individuals rapidly move to NSE closer to the unit, even though the parameters are not 100% correct. Some of the parameter ranges used in this case study had nearly a 200% variation from the minimum and maximum values (e.g., n_1 and n_2, as shown in Eq. S1 and the model still had minor errors compared to the expected values, as shown in Fig. 5c). The model also predicted that no infiltration would occur in this catchment since the near-optimal k_{sat} is 0 mm · h^{-1}. However, it predicted some initial abstraction of 0.02 mm, but this value is nearly negligible. Overall, the optimization problem resulted in a near-optimal solution that, at least for hydrological purposes, is sufficient and physically based. It preserves the peak flow and overall shape of the hydrograph and has an optimal NSE compared to the outlet hydrograph. However, relying solely on NSE might produce acceptable solutions within the parameter ranges, equifinality is presented as shown in Fig. 5(c). One can note in this figure that Manning's roughness coefficient had a range of approximately 20% around the expected values, although NSE values were optimized.

Table 3
Known parameters of the inverse problem of Numerical Case Study 4.

Classification	n [s · m$^{-1/3}$]	h_0 [mm]	k_{sat} [mm · h^{-1}]	$\Delta\theta$ [–]	ψ [mm]
Left Hillslope	0.06	1	8	0.6	20
Channel	0.15	0	0	0.1	0
Right Hillslope	0.015	4	2	0.15	100

4.2. Numerical case study 2

The modeling results of the Wooden-Plane catchment are presented in Fig. 5(d)-(f). In this problem, the RMSE was chosen as the objective function, and the nearly-optimal objective function value was approximately 630 mg · L^{-1}. For other modeling pollutants, such as copper, zinc, or phosphate, a RMSE of this magnitude would result in an inaccurate model (Batalini de Macedo et al., 2021); however, we are modeling salt concentrations that had mean concentrations of approximately 30.000 mg · L^{-1}, as shown in Fig. 5(d). The goodness of fitness can also be visualized in the inserted chart in Fig. 5(d), where the model nearly predicted the same concentrations as the observations.

The values of the objective function for each generation's best and worst individuals re also shown in Fig. 5(e). Results are already relatively good for the 1st generation and find a near-optimal plateau after the 5th generation. Although the optimization model found good results for fitting the observed concentrations, it came at the cost of estimating a larger mass of salt at the beginning of the simulation. The overprediction of C_1 and C_2 can be seen in Fig. 5(f), where $C_1 = 323.68$, $C_2 = 1.081$, $C_3 = 12.045,03$, $C_4 = 0.2763$. In particular, the combined values of C_1, C_2, and Δt_d would result in an initial salt mass of 145 g, which is approximately 16% more than the value reported by Hong et al. (2019). The black dashed line in this figure is the result of the same problem, using the same model (HydroPol2D), calibrating

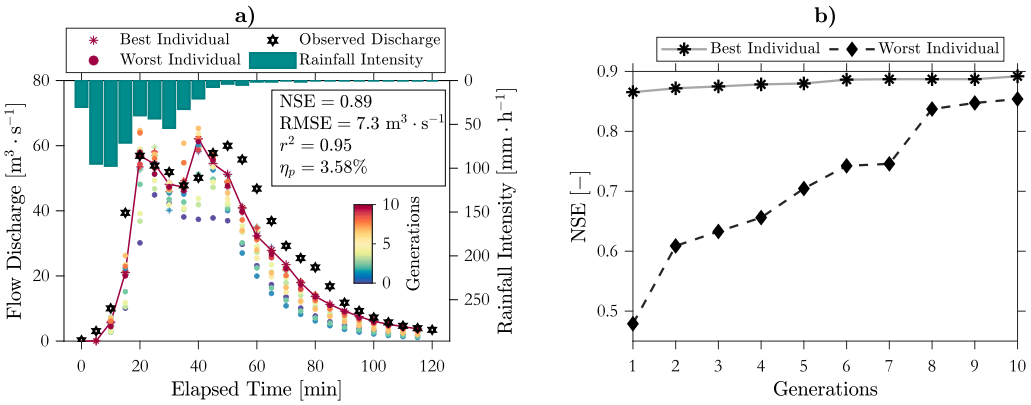


Fig. 8. Calibration problem results with the catchment simulated with 30 m spatial resolution under an observed rainfall event. Part (a) shows the hydrographs of the best and worst individuals for each generation, as well as the rainfall intensity. Part (b) shows the objective function (i.e., −NSE) values for the best and worst individuals.

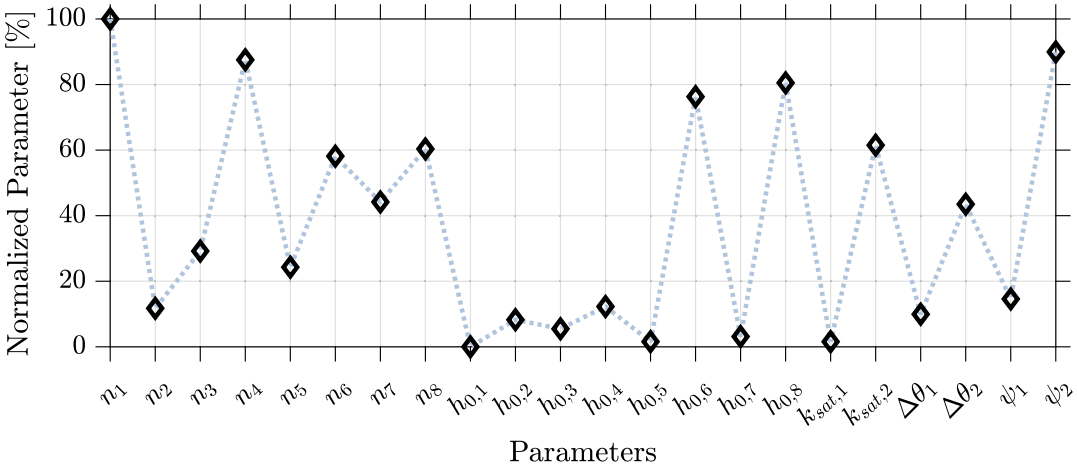


Fig. 9. Near-Optimal normalized parameter, where 0 and 100% are the boundaries of the decision vector $\boldsymbol{x}$. LULC-Based subscripts 1 = water, 2 = trees, 3 = Grass, 4 = Flooded Vegetation, 5 = Crops, 6 = Schrub/Scrub, 7 = Built Areas, and 8 = Bare Ground. Soil-Based subscripts 1 = Medium and 2 = Clayey.

for C_3 and C_4, but assuming the initial mass of 125 g (Gomes et al., 2023a). It is inferred that a larger mass was expected with a larger washing capacity, since all water quality parameters were larger for the simulations presented in this numerical case study.

4.3. Numerical case study 3

The model has 8 LULC and 2 Soil classifications, resulting in 22 hydrologic-hydrodynamic parameters (16 from LULC and 6 from the soil parameters). Using the average of the parameter range presented in Tables 1 and 2, the one-at-the-time sensitivities of these parameters are depicted in Figs. 6 and 7. The average of the parameters might be the baseline used when only the parameter ranges are known. The output functions used for this evaluation are mainly related to the hydrograph shape. We used peak flow, runoff volume at the end of the event, and time-to-peak variances as hydrograph shape evaluation functions. Furthermore, we evaluate the sensitivity of the parameters to flood areas (i.e., areas with maximum flood depth greater than 0.5 m). From Fig. 6 it is observed that the most sensitive parameters in terms of hydrograph shape are the Manning's roughness coefficient of the Built Areas (i.e., the watershed has nearly 70% of built areas), followed by those in the areas of trees. A reduction in roughness is fairly more expressive than an increase in peak flow. However, for runoff volume, time to peak, and flooded areas, Manning's variation seems to follow a linear relationship with these outputs.

It is interesting to note that reducing the n values reduces the total volume that leaves the catchment at the end of the event and decreases flooded areas, which does not necessarily mean that areas with risks of human instability would also decrease (Rotava et al., 2013). The larger velocities that resulted from the reduced Manning's coefficient might increase areas of instability risks. The initial abstraction (h_0) also had some sensitivity, but presented a very non-linear behavior for all output functions used, indicating that it could be assumed in some cases rather than calibrated. This non-linear behavior might be due to allowing storage and infiltration in cells even when depths are smaller than or equal to h_0. We hypothesize that h_0 would have more influence for values larger than 10 mm. This parameter can represent the storage of low-impact development (LID) facilities that do not change surface roughness, such as rain barrels. Parameters n and h_0 can be proxy representations of LID facilities such as rain barrels (i.e., increasing h_0 in pixels), green roofs, permeable pavements, or bioretention systems, as they represent the storage and delay of the flood wave passing through cells (Damodaram et al., 2010).

These parameters associated with the infiltration parameters shown in Fig. 7 can be used to assess the effects of retrofitting the catchment into a more sustainable scenario with green infrastructure (McClymont et al., 2020) or can also represent a scenario of increase in urbanization and hence represent the effects of post-development conditions (Gomes et al., 2023b). Due to the limited area for infiltration in the catchment, the results presented in Fig. 7 indicate that the soil properties have less influence than the roughness coefficients but greater influence than the initial abstraction, especially $k_{sat,2}$, which can be seen from Fig. 4 that most of the pervious areas are derived from this type of soil.

The model calibration results are presented in Fig. 8. Part (a) shows the hydrograph for the best and worst individuals of each generation,

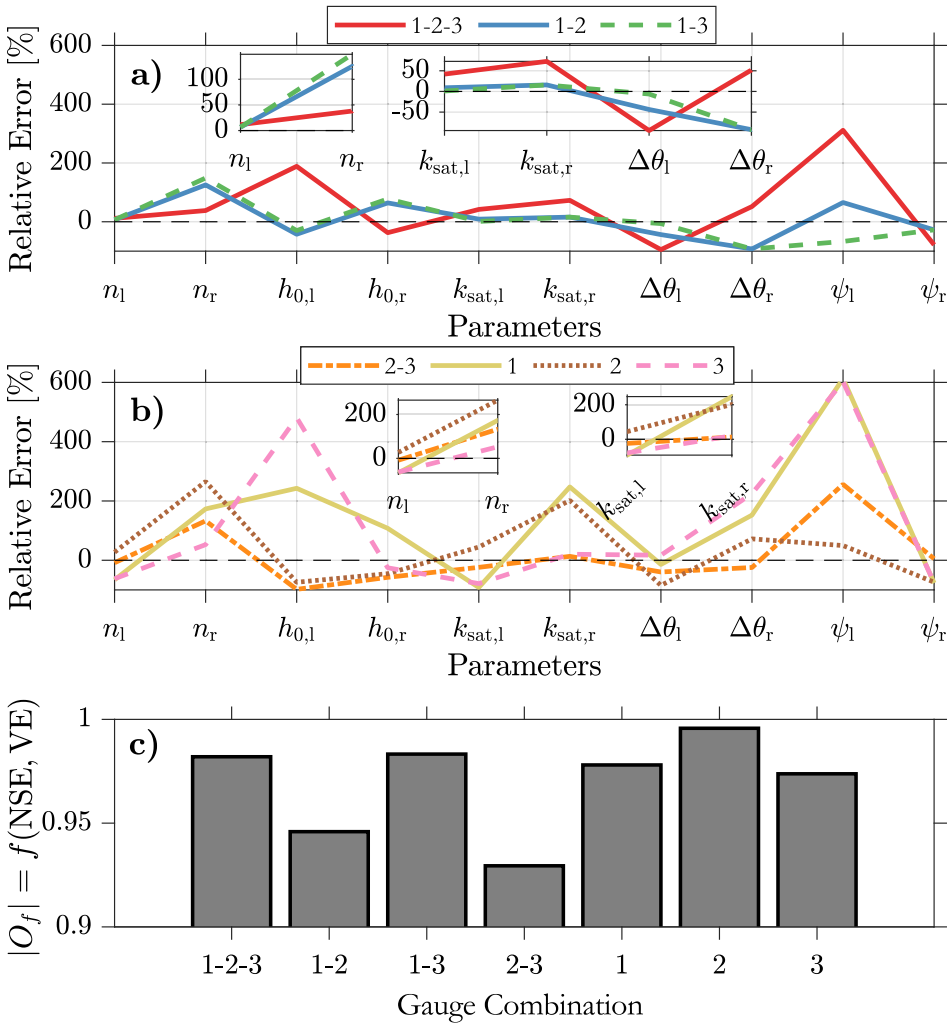


Fig. 10. Relative Parameter Error for Numerical Case Study 4, assuming no prior knowledge of the parameter boundaries where 1 = Outlet, 2 = left gauge, 3 = right gauge. Black dashed lines are the expected values. Part (a) is the relative error for cases where the outlet and at least one more gauge are observed and case, Part (b) is single gauges or a combination of gauges that are not at the outlet, and Part (c) is the objective function values given by Eq. S1. All cases were simulated with 10 generations and 100 populations.

while part (b) shows the evolution of the objective function (i.e., - NSE) through the generations. The calibration results present a NSE = 0.89, RMSE = 7.3 $m^3 \cdot s$, and $r^2 = 0.95$, and $\eta_p = 3.58\%$. Although we ran the optimization algorithm for 10 generations, 100 population, the model still could not properly capture the two-peak observed hydrograph and the observed runoff volume in the falling limb of the hydrograph. Other objective functions could also be tested, especially those that account for peak flows and runoff volumes, but for the sake of simplicity, we only used the NSE. Several factors might have influenced this behavior, and we hypothesize that the most important ones are the quality of the digital elevation model, the spatial variability of rainfall in the catchment, and the uncertainty in the rainfall measurements and transformation of stage into discharge, as well as the inherent uncertainty of the conceptual model of HydroPol2D. The calibrated parameters are shown in Tables 1 and 2.

As in Numerical Case Studies 1 and 2, after the first generation, the model performs similarly to the last generation (i.e., NSE = 0.85), indicating that relatively fewer simulations can be required to reach accepted modeling results (Moriasi et al., 2015). As shown in Fig. 9, none of the parameters reached the boundaries of the range defined for the upper and lower bounds. However, a relatively large Manning's roughness coefficient is noted. Using a coarser DEM and filtering the DEM with Gaussian filters, CRS, and carving water surface depths in the channel, we hypothesized that the flow paths were shortened and smoothed so that increasing n was necessary. However, not including these filters might cause water ponding in areas with DEM noise originating from natural imprecisions of the DEM or the influence of vegetation (De Paiva et al., 2013). Furthermore, using a 30-m DEM, the terrain details that would be captured with a higher resolution DEM are not considered. Higher resolution DEMs can eventually create more contact area between the runoff and the surface, resulting in larger head losses. Therefore, to compensate for the relatively smoother terrain, a larger value n was required. Similar findings of the one presented here can be seen in Bellos and Tsakiris (2015).

Using calibrated parameters, HydroPol2D can be applied to determine flood maps, human instability maps, infiltration, and other components of the water balance as shown in Gomes et al. (2023a).

4.4. Numerical case study 4

The relative parameter error assuming a considerably wide parameter range representing no prior knowledge about the parameters is shown in Fig. 10. Overall, the algorithm can find suitable sets of n and k_{sat}, i.e., the most sensitive parameters using the outlet gauge in combination with some of the other gauges, but have larger errors in the other less sensitive Green-Ampt parameters and in the initial abstraction, as shown in Fig. 10(a). These other parameters are less

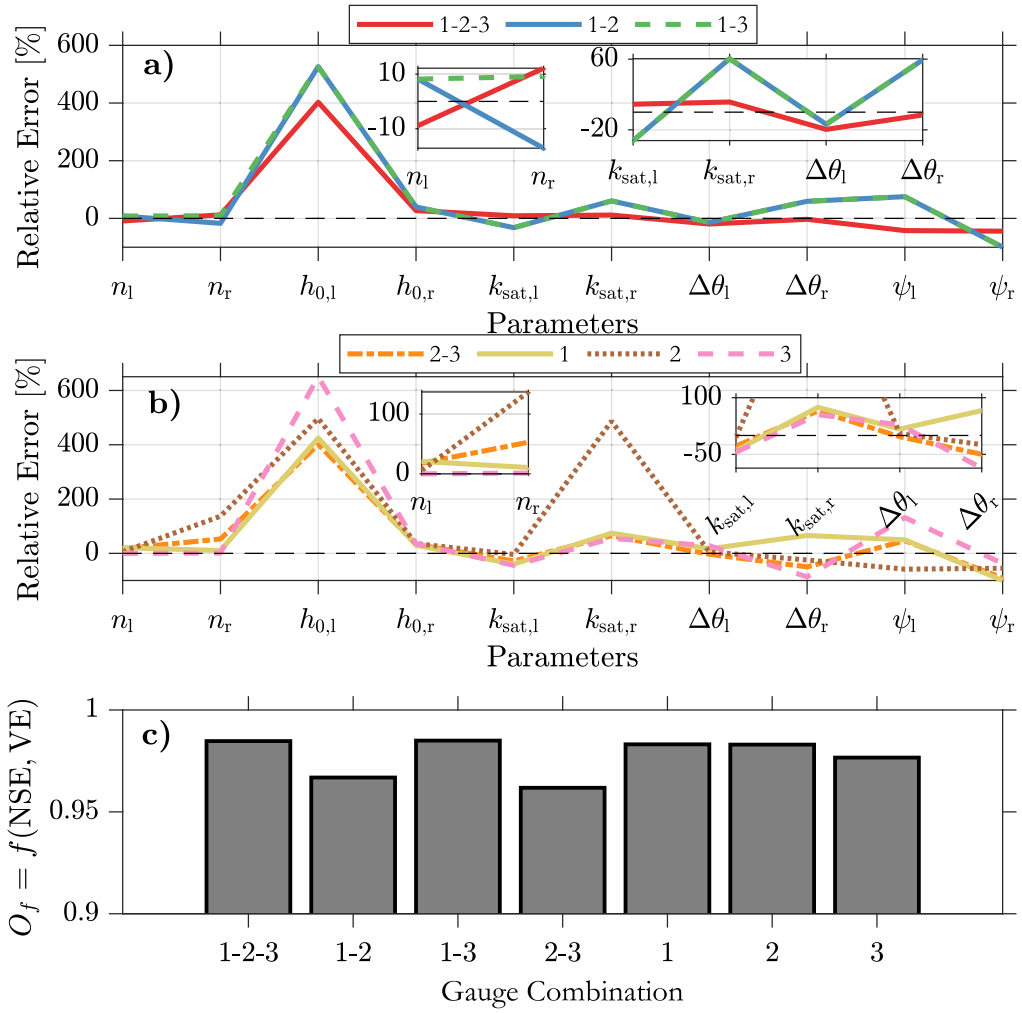


Fig. 11. Relative Parameter Error for Numerical Case Study 4, assuming a prior knowledge, that is, half of the parameter range from Fig. 10 of the parameter boundaries where 1 = Outlet, 2 = left gauge, 3 = right gauge. Black dashed lines are the expected values. Part (a) is the relative error for cases where the outlet and at least one more gauge are observed and case, Part (b) is single gauges or a combination of gauges that are not at the outlet, and Part (c) is the objective function values given by Eq. S1. All cases were simulated with 10 generations and 100 population.

sensitive and could have been assumed otherwise without prejudice to the overall representation of the hydrology and hydrodynamics.

When using gauges 2 or 3, that is, gauges from the left and right hillslopes, the error in the parameters of the opposite hillslope is relatively larger since the parameters were randomly chosen because no information about the discharges of that area is available, as shown in Fig. 10(b). Using gauges 2 and 3 together, the parameter error is reduced since information from both hillslopes (i.e., areas with similar hydrologic characteristics) is available. Using only gauge 1 as the source of information for calibration, a good objective function value is found (see Fig. 10(c)). Still, a very poor description of the parameters is found, as shown in Fig. 10(b), indicating a high chance of equifinality if no minimum knowledge of the parameters is known. This result shows the importance of experts defining proper parameter ranges for the system prior to the calibration phase. All objective function performances could be considered feasible for different hydrologists since solutions with high fitness values were found; although, parameters were off from the correct ones, illustrating the equifinality issue in calibration.

When comparing the performance of solutions without prior knowledge of the parameters with solutions with a smaller decision space, the equifinality tends to decrease, as shown in Fig. 10 compared to Fig. 11 and the values of the objective function are generally higher. This result indicates that reducing decision space to a more reduced space can substantially decrease equifinality. Some parameters less influential as $h_{0,l}$ had larger errors but little sensitivity and could have been assumed rather than calibrated.

The number of events in the calibration also plays an important role in reducing equifinality. By using a relatively small rainfall event that is not a runoff-producing event in the left hillslope, no quality information is available to calibrate the hydrodynamic and infiltration parameters of this hillslope. The left gauge in event 1, with 10.8 $\text{mm} \cdot \text{h}^{-1}$ did not record runoff. Therefore, any combination of parameters such that all water infiltrates in the soil is a solution with full performance of the objective function. An infinite number of combinations of parameters would satisfy this condition (e.g., $h_{0,l} > 16.2$ mm, $k_{sat,l} > 10.8$ $\text{mm} \cdot \text{h}^{-1}$), leading to high equifinality due to poor gauging location and selection of the event for calibration. However, for the right gauge, runoff is observed, and the parameters can be relatively well estimated, although not perfectly. The problem of distributed physically-based modeling in small catchments with events that produce little or no runoff is a complex problem, and models typically have lower performance for hortonian small flows (Senarath et al., 2000; Downer and Ogden, 2004).

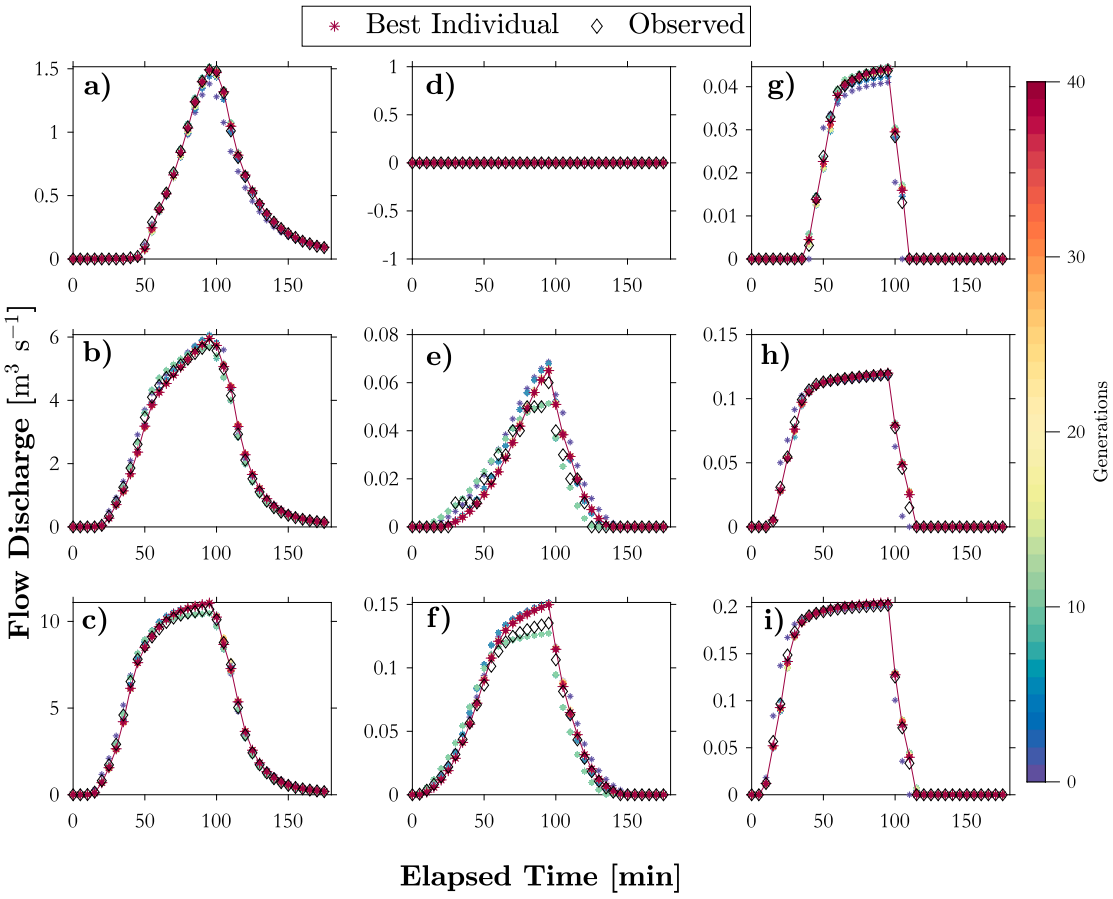


Fig. 12. Automatic calibration results of considering the 3 events but only the outlet as the observable gauge. Therefore, the charts (d)-(i) are shown, but the modeled results were not considered in the calibration of the model and were simulated with the parameters that were obtained by calibrating the model only with the outlet gauge. The first row are the events of 10.8 mm h^{-1}, followed by 21.6 mm h^{-1} and 32.4 mm h^{-1}. Parts (a)-(c) are results for the outlet, whereas (d)-(f) are from the left gauge and (g)-(i) are from the right gauge.

This idea is illustrated in Table 4. By choosing only one poor event, the calibration performance is nearly optimal for all combinations of gauges, but the parameter estimation is faulty. Therefore, calibrating the model for more extended hydrological periods or choosing a combination of events encompassing relatively high and low flows is desirable to increase the available information and reduce parameter equifinality. However, the uncertainty in the boundary conditions and the initial simulation values, especially the initial soil moisture (Senarath et al., 2000), is challenging. Even in a perfect virtual experiment without uncertainty in rainfall values, initial soil moisture, model boundary conditions, and perfect gauging data, the uncertainty in the parameters is substantially affected by a poor parameter range.

One of the advantages of this calibration approach is using the model to calibrate the parameters using only the outlet data as the sole gauge, which would be the case for many poorly gauged and flood-prone catchments such as the Gregorio Catchment. To this end, we use a relatively high optimization resource; that is, we run the optimization model for 40 generations and 100 population size and optimize Eq. S1 using only the outlet as the observed gauge for all events available. The rationale behind using a larger number of generations is to extract the maximum resource of the single-point observed information since it is only at one gauge. Using a larger number of generations would likely decrease the possibility of finding local optima in the optimization model. Therefore, the uncertainty would probably be due to equifinality since there is no uncertainty in rainfall and observations in this inverse problem.

In this analysis, we assume the initial abstractions of the left and right gauges are the correct parameters since they do not play an important role in the hydrological response of the catchment, as mentioned above in this section. The other parameter ranges are the same as those used in the simulated cases with prior knowledge of the system (see Tab. S1). The results in Fig. 12 show the hydrographs for the outlet (a)-(c) and for the other gauges not considered in the calibration (i.e., left gauge (d)-(f), and right gauge (g)-(i)). Even calibrating with only the catchment outlet, the model can still find a reasonable, physically-based, and bounded parameter set. However, the parameters are not equal to those of the inverse problem. This result points to the scenario that, given a sufficient number of runoff-producing events and reasonable computational resources, it is possible to calibrate HydroPol2D only with data at the outlet and later use the calibrated model to derive important catchment response information such as infiltrated depths, flood depths, and velocities. The calibrated parameters of this analysis are $n_l = 0.0536$ s · $m^{-1/3}$, $n_r = 0.0168$, $k_{sat,l} = 3.56$ mm · h^{-1}, $k_{sat,r} = 1.77$ mm · h^{-1}, $\Delta\theta_l = 0.625$, $\Delta\theta_r = 0.346$, $\psi_l = 91.36$ mm, $\psi_r = 49.92$ mm.

By comparing the calibrated parameters with the ones of the inverse problem shown in Table 3, it is noticed that a trade-off between $k_{sat,l}$ and ψ_l is found for the left hillslope. While $k_{sat,l}$ decreases, $\Delta\theta_l$ and ψ_l increase, counterbalancing the reduction in $k_{sat,l}$. However, even though the parameter equifinality is evident, the model performance and the errors are visually minimal, as shown in Figs. 13 and 14. In addition, the performance metrics are also accepted in most gauges, as shown in Table 5.

The model presented acceptable results for steady-state events, with all volume errors smaller than 6% and all NSE larger than 0.996. For unsteady-state hyetographs, as expected, the model presented a relatively reduced performance for the left gauge, especially for event

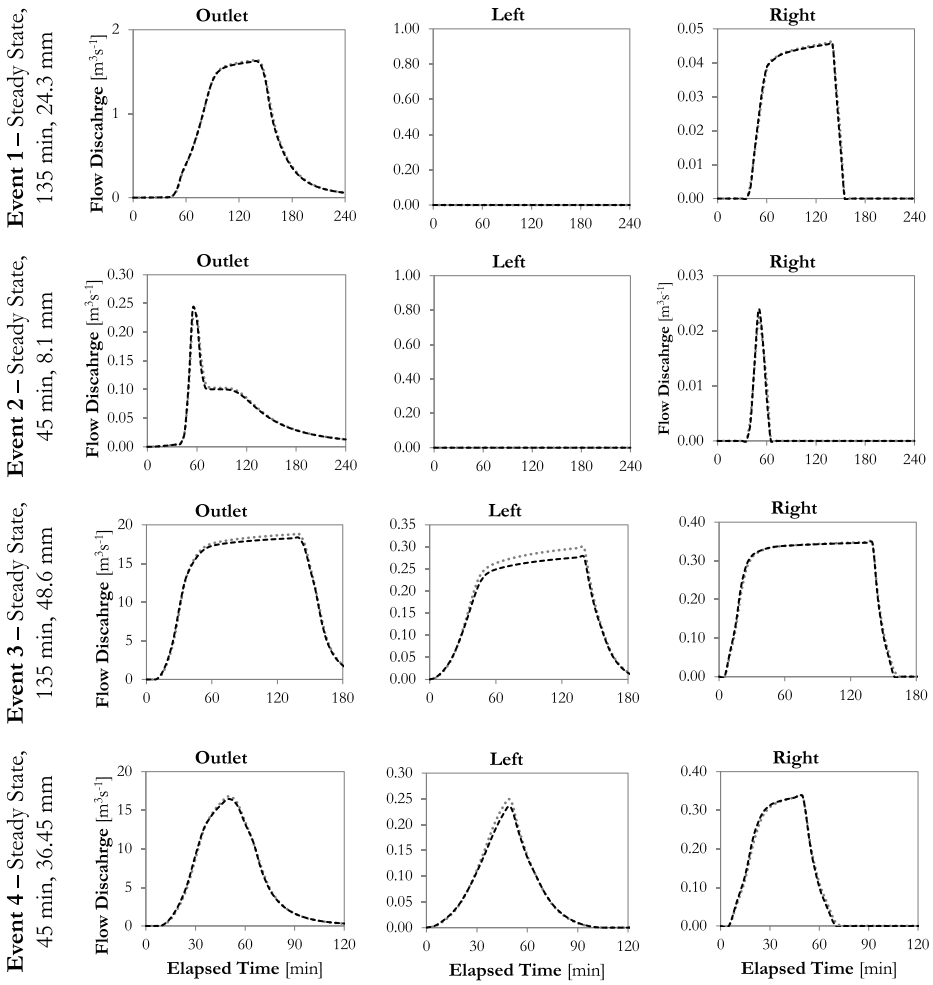


Fig. 13. Steady-state rainfall validation hydrographs of Numerical Case Study 4, for Events 1 to 4 described in Table 5. Gray dotted lines are modeled results with the calibrated model using only the outlet gauge data, and black dashed lines are the results with the parameters of the inverse problem.

Table 4
Near-optimal solutions for different combinations of gauges and for only 1 storm of 10.8 mm · h^{-1} during 90 min. The known parameters are $n_l = 0.06$ s · m$^{-1/3}$, $n_r = 0.015$, $h_{0,l} = 1$ mm, $h_{0,r} = 4$ mm, $k_{sat,l} = 8$ mm · h^{-1}, $k_{sat,r} = 2$ mm · h^{-1}, $\Delta\theta_l = 0.6$, $\Delta\theta_r = 0.15$, $\psi_l = 20$ mm, and $\psi_r = 100$ mm..

Gauges	n_l [s · m$^{-1/3}$]	n_r [s · m$^{-1/3}$]	$h_{0,1}$ [mm]	$h_{0,2}$ [mm]	$k_{sat,l}$ [mm · h^{-1}]	$k_{sat,r}$ [mm · h^{-1}]	$\Delta\theta_l$ [-]	$\Delta\theta_r$ [-]	ψ_l [mm]	ψ_r [mm]	OF [-]
1-2-3	0.036	0.019	9.17	5.25	8.986	3.44	0.65	0.04	92.34	8.33	−0.99
1-2	0.047	0.015	9.73	7.68	10.059	1.30	0.52	0.11	14.49	13.80	−0.99
1-3	0.068	0.015	5.36	7.68	5.395	1.78	0.64	0.11	15.26	13.80	−0.98
2-3	0.050	0.014	5.24	5.61	2.761	2.48	0.68	0.16	54.34	20.14	−0.98
1	0.040	0.013	9.95	5.51	3.191	3.12	0.77	0.27	0.61	0.17	−0.99
2	0.063	0.048	7.12	8.76	9.004	0.91	0.74	0.17	36.91	51.51	−1.00
3	0.032	0.020	6.44	5.14	6.630	3.36	0.73	0.04	38.00	23.21	−0.97

6, that is, the event with the smallest duration and volume. As shown in Fig. 14, event 6 generated a very low runoff rate observed in the inverse problem and not predicted by the calibrated model. In addition, for the right gauge, a relatively large volume error can be observed. For the outlet, however, the results are still quite accurate; although relatively faulty for the left and right gauges. Disregarding this event, the simulation results had volume errors smaller than 20% and NSE larger than 0.992. Overall, using the calibrated parameters obtained only with the outlet gauge is sufficient to explain the events used for calibration and can accurately represent the hydrological response of events outside the hydrological characteristics of the events used for calibration. Even though some errors are found in the internal gauges, the model's performance measured in the outlet can be considered very good for all validation events.

4.5. Limitations, challenges, and opportunities of this modeling approach

Calibrating a fully-distributed hydrodynamic and water quality model requires not only field observations but also depends extensively on the quality and resolution of the terrain, soils, land use, and land cover models. However, as presented in this paper, this modeling approach can be easily applied worldwide in catchments with scarce time-series rainfall observations and a representing variable of the flow dynamics, such as discharge depths, and a variable representing the pollutant, such as the pollutant concentration. Since pollutant concentrations are inherently associated with accurate discharge modeling, calibrating water quantity, and quality parameters altogether might result in high equifinality if only pollutographs are the optimization criteria.

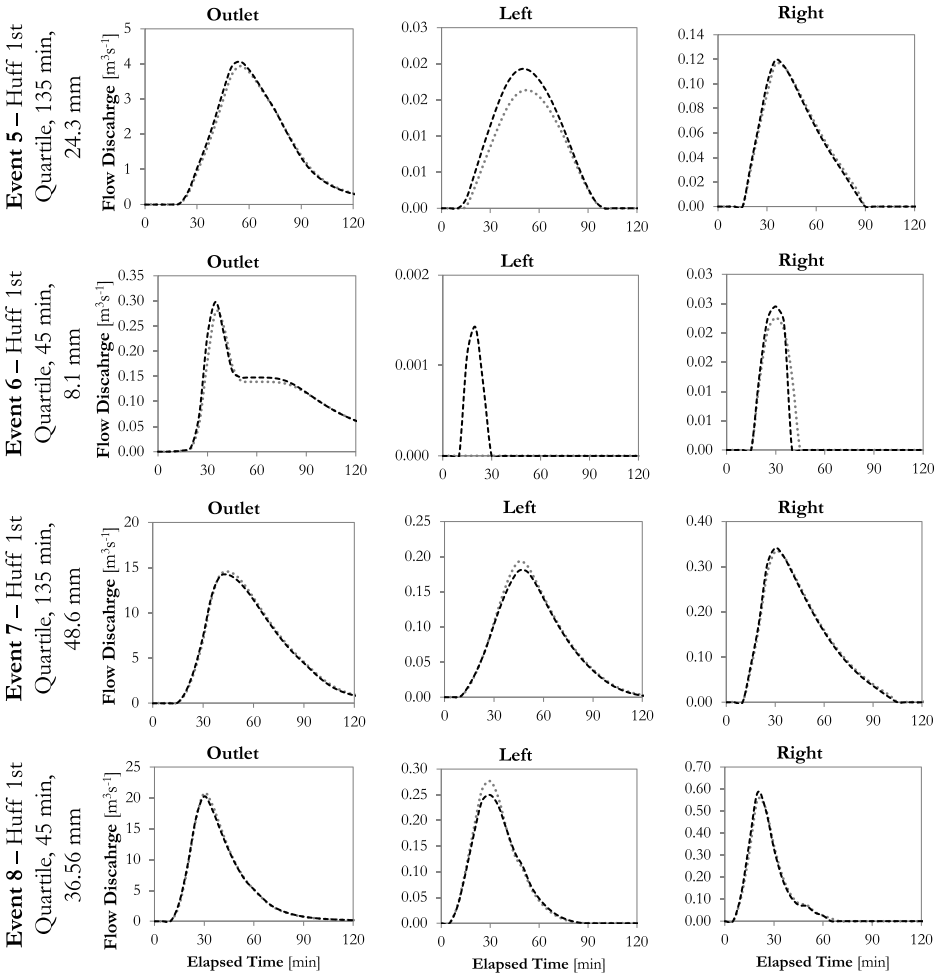


Fig. 14. Unsteady-state rainfall results of Numerical Case Study 4, for Events 5 to 8 described in Table 5. The rainfall is simulated with Huff 1st quartile hyetograph. Gray dotted lines are modeled results with the calibrated model using only the outlet gauge data, and black dashed lines are the results with the parameters of the inverse problem.

This approach can be enhanced and easily expanded by allowing calibration, not only with time series but also with maps of flood extent, magnitude, or by socio-hydrological information (Fava et al., 2022) such as maximum depths in certain flood points, especially in catchments with no gauge stations (Gomes et al., 2023a). The challenge, however, is to find a the suitable single objective cost function that can normalize different optimization criteria into a single and homogeneous cost function. In addition, the minimum requirement, however, is the rainfall intensity time-series in a proper resolution that depends on the catchment response. Regarding rainfall, this approach could also be improved by allowing space-variant rainfall that could be derived from radar, satellite imagery, or by interpolation of source-gauged rainfall stations.

It is recommended that a sensitivity analysis be performed before automatic calibration to avoid wasting computational resources on variables that do not play a substantial role in the catchment's hydrologic-hydrodynamic behavior. Although the results presented in this paper indicate that some parameters might be more sensitive than others, the results can vary dramatically for different catchments with different topography, LUCL, and soil properties.

The use of worldwide datasets to represent LULC and SOIL allows a proper definition of model parameters such as n or k_{sat}, as shown in Soliman et al. (2022) and Gupta et al. (2021). Studies such as the aforementioned ones might facilitate the parametrization of fully distributed models and can be opportunities for worldwide application.

5. Conclusions

An optimization-based algorithm was developed and applied to calibrate a fully distributed hydrological-hydrodynamic and water quality model (HydroPol2D). The algorithm can find near-optimal parameters to explain the observed gauged information, such as flow discharge, pollutant concentration, or flood depths. The answers (A) to the posed questions of Numerical Case Studies 1 (Q1), 2 (Q2), from the real-world case study in Numerical Case Study 3 (Q3), and the Equifinality analysis in Numerical Case Study 4 (Q4-1 and Q4-2) support the following:

- A1: The model can accurately predict not only the Green-Ampt infiltration parameters but also Manning's roughness coefficients and initial abstraction values, as shown in Numerical Case Study 1. The Predicted hydrographs match, with NSE > 0.99 the considered real observed hydrograph in Numerical Case Study 1.
- A2: The algorithm can find the wash-off parameters and the initial mass (error < 15%) of the pollutant in the wooden-plane catchment to match the observed pollutograph, as shown in Numerical Case Study 2. Even though the pollutograph was predicted correctly, an approximately 20% error was observed in the water quality parameters, indicating the equifinality.
- A3: The model can still find a physically bounded near-optimal set of parameters to calibrate the observed hydrograph with NSE = 0.89, indicating good accuracy. In addition, using this set of parameters, it was possible to determine distributed model outputs such as (i) infiltration maps, maximum flow velocities, (ii)

Table 5
Validation metrics for events outside of the calibration conditions. Each row represents the results of one event for different gauges. Events 1-2 are with the smallest rainfall intensity used in calibration (i.e., 10.8 $mm \cdot h^{-1}$, while events 3-4 are with the largest one (i.e., 32.4 $mm \cdot h^{-1}$. Similarly, the events with unsteady-state rainfall follow the same pattern, resulting in the same rainfall volumes but temporally distributed with Huff 1st quartile hyetographs.

Event	Rainfall boundary condition	Rainfall duration [min]	Rainfall volume [mm]	Gauge	NSE [-]	Vol Error [%]	r^2 [-]	PBIAS [%]
1	Steady state	135	24.3	Outlet	0.999	1.076	1.000	0.347
				Left	[-]	[-]	[-]	[-]
				Right	0.999	0.951	0.999	0.589
2	Steady state	45	8.1	Outlet	0.993	1.847	0.997	1.184
				Left	[-]	[-]	[-]	[-]
				Right	0.989	1.456	0.995	5.171
3	Steady state	135	48.6	Outlet	0.998	2.121	1.000	0.494
				Left	0.990	5.961	1.000	1.269
				Right	0.999	0.224	1.000	0.381
4	Steady state	45	36.45	Outlet	0.999	0.378	1.000	0.712
				Left	0.996	3.227	1.000	1.696
				Right	0.998	−0.483	0.999	1.119
5	Huff 1st Quartile	135	24.3	Outlet	0.997	−2.505	0.999	1.189
				Left	0.960	−20.399	0.997	4.777
				Right	0.997	0.226	0.999	1.401
6	Huff 1st Quartile	45	8.1	Outlet	0.983	−3.425	0.993	1.862
				Left	−0.061	[-]	[-]	59.408
				Right	0.895	6.872	0.946	15.927
7	Huff 1st Quartile	135	48.6	Outlet	0.999	1.310	1.000	0.684
				Left	0.996	4.049	1.000	1.358
				Right	0.998	0.059	0.999	0.972
8	Huff 1st Quartile	45	36.45	Outlet	0.999	0.708	1.000	1.022
				Left	0.992	3.125	0.998	2.573
				Right	0.995	−0.424	0.998	2.571

maximum flood depths, and (iii) outlet hydrograph. Therefore, using the same data as required in the calibration of lumped-parameter models, one can calibrate HydroPol2D and provide spatially varied outputs that can be used for water resources planning and management.

- A4-1: The equifinality problem is reduced by the addition of runoff-producing events and by choosing at least one gauge in hydrological unit regions. Using only the outlet gauge as the information might produce a feasible (i.e., within the parameter range) but wrong parameter set that explains the observed data. This set of parameters tends to produce small errors in peaks and hydrograph shapes and relatively larger errors in runoff volumes. Using runoff-producing events with different flow parameters typically produces better parameter estimation. The parameter estimation error is reduced by a more reduced parameter range that experts can attach or by GIS available worldwide datasets to reduce parameter ranges.
- A4-2: The model presented accurate results when calibrated only with the outlet gauge hydrograph as the sole information for calibration. Although the equifinality is observed by the compensation of the infiltration parameters, the model presented acceptable results in most cases of different rainfall volumes, intensities, and distributions. A reduced model performance is obtained for events with little or no observed runoff in the gauges. However, in general, the model presented great results in the gauges not used for the calibration under different rainfall durations, volumes, intensities, and rainfall distributions from the ones used in the calibration, even with the model calibrated with only the outlet gauge. This indicates an opportunity to move towards conceptual and simplified lumped models flood assessment to physically-based, fully-distributed analysis since both models can be calibrated with the same input data.

Therefore, the methods applied in this paper can be replicated in all catchments with observations of at least one gauge. The more gauges with runoff observations, the typically better the reduction of equifinality, assuming reliable observed data. Using only freely available datasets, this method can be applied for catchments with observations at gauging stations to extrapolate results in the catchment domain, moving from typically limited lumped-parameter models to fully-distributed physically-based analysis. However, the methodology strategy developed in this paper is only applicable if some constraints are satisfied, such as:

- The overland flow is predominantly hortonian.
- The effect of human made drainage systems such as reservoirs, dams, polders, or any other hydraulic structure operation does not govern the whole catchment hydrodynamics.
- The catchment can be modeled with space-invariant precipitation.
- The optimization cost function is relatively fast, allowing multiple evaluations in a reasonable time.

The requirements above are typically satisfied in relatively small to mid-size urban catchments. Advancing these limitations and developing a framework capable of adapting to whatever available data could help modelers use distributed models and improve flood and water quality spatial analysis. Future studies can investigate the effect of different spatial resolutions in the calibration of HydroPol2D and how that affects the equifinality problem.

CRediT authorship contribution statement

Marcus Nóbrega Gomes Jr.: Writing - review & editing, Writing - original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marcio Hofheinz Giacomoni:** Writing - review & editing, Validation, Methodology, Formal analysis. **Fabricio Alonso Richmond Navarro:** Writing - review & editing, Resources, Data curation. **Eduardo Mario Mendiondo:** Writing - review & editing, Supervision, Resources, Funding acquisition, 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.

Data availability

Some or all data, models, or code generated or used during the study are available in a repository or online in accordance with funder data retention policies. All software, figures, and data can be freely downloaded in Gomes et al. (2023a).

Acknowledgment

The authors appreciate the support of the City of San Antonio, by the San Antonio River Authority, CAPES Ph.D Scholarship 88887.663281/2022-00 and by the PPGSHS PROEX Graduate Program.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.envsoft.2024.106128.

References

Abreu, F.G.d., 2019. Quantificação dos Prejuízos Econômicos à Atividade Comercial Derivados de Inundações Urbanas (Ph.D. thesis). Universidade de São Paulo.

Afshar, A., Kazemi, H., Saadatpour, M., 2011. Particle swarm optimization for automatic calibration of large scale water quality model (CE-QUAL-W2): Application to Karkheh reservoir, Iran. Water Resour. Manag. 25, 2613–2632.

Akan, A.O., Iyer, S.S., 2021. Open Channel Hydraulics. Butterworth-Heinemann.

Ardıçlıoğlu, M., Kuriqi, A., 2019. Calibration of channel roughness in intermittent rivers using HEC-RAS model: Case of Sarimsakli creek, Turkey. SN Appl. Sci. 1, 1–9.

Batalini de Macedo, M., Pereira de Oliveira, T.R., Halmenschlager Oliveira, T., Nóbrega Gomes Junior, M., Teixeira Brasil, J.A., Ambrogi Ferreira do Lago, C., Mendiondo, E.M., 2021. Evaluating low impact development practices potentials for increasing flood resilience and stormwater reuse through lab-controlled bioretention systems. Water Sci. Technol. 84, 1103–1124.

Bates, P.D., Horritt, M.S., Fewtrell, T.J., 2010. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. Journal of hydrology 387, 33–45.

Behrouz, M.S., Zhu, Z., Matott, L.S., Rabideau, A.J., 2020. A new tool for automatic calibration of the storm water management model (SWMM). J. Hydrol. 581, 124436.

Bellos, V., Tsakiris, G., 2015. Comparing various methods of building representation for 2D flood modelling in built-up areas. Water Resour. Manag. 29, 379–397.

Bermudez, A., Vazquez, M.E., 1994. Upwind methods for hyperbolic conservation laws with source terms. Comput. & Fluids 23, 1049–1071.

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, 11–29.

Blasone, R.-S., Madsen, H., Rosbjerg, D., 2008. Uncertainty assessment of integrated distributed hydrological models using GLUE with Markov chain Monte Carlo sampling. J. Hydrol. 353, 18–32.

Brath, A., Montanari, A., Toth, E., 2004. Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model. J. Hydrol. 291, 232–253.

Brunner, G.W., 2016. HEC-RAS river analysis system modeling user's manual US army corps of engineers hydrologic engineering center. Information on http://www.hec.usace.army.mil.

Cho, J.H., Lee, J.H., 2015. Watershed model calibration framework developed using an influence coefficient algorithm and a genetic algorithm and analysis of pollutant discharge characteristics and load reduction in a TMDL planning area. J. Environ. Manag. 163, 2–10.

Confesor, Jr., R.B., Whittaker, G.W., 2007. Automatic calibration of hydrologic models with multi-objective evolutionary algorithm and Pareto optimization 1. JAWRA J. Am. Water Resour. Assoc. 43, 981–989.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1. 4. Geosci. Model Dev. 8, 1991–2007.

Damodaram, C., Giacomoni, M.H., Prakash Khedun, C., Holmes, H., Ryan, A., Saour, W., Zechman, E.M., 2010. Simulation of combined best management practices and low impact development for sustainable stormwater management 1. JAWRA J. Am. Water Resour. Assoc. 46, 907–918.

Dariane, A.B., Javadianzadeh, M.M., James, L.D., 2016. Developing an efficient auto-calibration algorithm for HEC-HMS program. Water Resour. Manag. 30, 1923–1937.

de Geografia e Estatística (IBGE), I., 2022. Panorama - população, território e ambiente. URL: https://cidades.ibge.gov.br/brasil/sp/sao-carlos/panorama.

de Meteorologia, I.I.N., 2022. Gráficos climatólogicos. URL: https://clima.inmet.gov.br/prec.

De Paiva, R.C.D., Buarque, D.C., Collischonn, W., Bonnet, M.-P., Frappart, F., Calmant, S., Bulhões Mendes, C.A., 2013. Large-scale hydrologic and hydrodynamic modeling of the Amazon river basin. Water Resour. Res. 49, 1226–1243.

Debele, B., Srinivasan, R., Parlange, J.-Y., 2008. Coupling upland watershed and downstream waterbody hydrodynamic and water quality models (SWAT and CE-QUAL-W2) for better water resources management in complex river basins. Environ. Model. Assess. 13, 135–153.

Deletic, A., Orr, D.W., 2005. Pollution buildup on road surfaces. J. Environ. Eng. 131, 49–59.

Domeneghetti, A., Castellarin, A., Brath, A., 2012. Assessing rating-curve uncertainty and its effects on hydraulic model calibration. Hydrol. Earth Syst. Sci. 16, 1191–1202.

Downer, C.W., Ogden, F.L., 2004. GSSHA: Model to simulate diverse stream flow producing processes. J. Hydrol. Eng. 9, 161–174.

Dung, N.V., Merz, B., Bárdossy, A., Thang, T.D., Apel, H., 2011. Multi-objective automatic calibration of hydrodynamic models utilizing inundation maps and gauge data. Hydrol. Earth Syst. Sci. 15, 1339–1354.

Fatichi, S., Vivoni, E.R., Ogden, F.L., Ivanov, V.Y., Mirus, B., Gochis, D., Downer, C.W., Camporese, M., Davison, J.H., Ebel, B., et al., 2016. An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. J. Hydrol. 537, 45–60.

Fava, M.C., Macedo, M.B.d., Buarque, A.C.S., Saraiva, A.M., Delbem, A.C.B., Mendiondo, E.M., 2022. Linking urban floods to citizen science and low impact development in poorly gauged basins under climate changes for dynamic resilience evaluation. Water 14, 1467.

Fava, M.C., Mazzoleni, M., Abe, N., Mendiondo, E.M., Solomatine, D.P., 2020. Improving flood forecasting using an input correction method in urban models in poorly gauged areas. Hydrol. Sci. J. 65, 1096–1111.

Fisher, R.A., et al., 1920. 012: A mathematical examination of the methods of determining the accuracy of an observation by the mean error, and by the mean square error. Mon. Not. R. Astron. Soc..

Giacomoni, M.H., Joseph, J., 2017. Multi-objective evolutionary optimization and Monte Carlo simulation for placement of low impact development in the catchment scale. J. Water Resour. Plan. Manag. 143, 04017053.

Gomes, M.N., do Lago, C.A.F., Rápalo, L.M.C., Oliveira, P.T.S., Giacomoni, M.H., Mendiondo, E.M., 2023a. HydroPol2D – Distributed hydrodynamic and water quality model: Challenges and opportunities in poorly-gauged catchments. J. Hydrol. 625, 129982. http://dx.doi.org/10.1016/j.jhydrol.2023.129982, URL: https://www.sciencedirect.com/science/article/pii/S0022169423009241.

Gomes, Jr., M.N., Giacomoni, M.H., de Macedo, M.B., do Lago, C.A.F., Brasil, J.A.T., de Oliveira, T.R.P., Mendiondo, E.M., 2023b. A modeling framework for bioretention analysis: Assessing the hydrologic performance under system uncertainty. J. Hydrol. Eng. 28, 04023025.

Gomes, Jr., N.M., Giacomoni, M.H., Papagiannakis, A.T., Mario Mendiondo, E., Dornelles, F., 2021. Spatial assessment of overland flow, pollutant concentration, and first flush using a 2D non-point source pollution and hydrological model for urban catchments. In: World Environmental and Water Resources Congress 2021. pp. 397–413.

Gomes, Jr., M.N., Rápalo, L.M.C., Oliveira, P.T.S., Giacomoni, M.H., do Lago, C.A.F., Mendiondo, E.M., 2023c. Modeling unsteady and steady 1D hydrodynamics under different hydraulic conceptualizations: Model/software development and case studies. Environ. Model. Softw. 105733.

Gomes, Jr., M.N., Taha, A.F., Rápallo, L.M.C., Mendiondo, E.M., Giacomoni, M.H., 2024. Real-time regulation of detention ponds via feedback control: Balancing flood mitigation and water quality. arXiv preprint arXiv:2403.04675.

Gomes Jr., 2024. HydroPol2D model v.114. https://github.com/marcusnobrega-eng/HydroPol2D.

Gomes Júnior, M.N., Giacomoni, M.H., Taha, A.F., Mendiondo, E.M., 2022. Flood risk mitigation and valve control in stormwater systems: State-space modeling, control algorithms, and case studies. J. Water Resour. Plan. Manag. 148, 04022067.

Green, W.H., Ampt, G.A., 1911. Studies on soil physics. J. Agric. Sci. 4, 1–24.

Guidolin, M., Chen, A.S., Ghimire, B., Keedwell, E.C., Djordjević, S., Savić, D.A., 2016. A weighted cellular automata 2D inundation model for rapid flood analysis. Environ. Model. Softw. 84, 378–394.

Gupta, S., Lehmann, P., Bonetti, S., Papritz, A., Or, D., 2021. Global prediction of soil saturated hydraulic conductivity using random forest in a covariate-based geotransfer function (CoGTF) framework. J. Adv. Modelling Earth Syst. 13, e2020MS002242.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. J. Hydrol. Eng. 4, 135–143.

Hawker, L., Bates, P., Neal, J., Rougier, J., 2018. Perspectives on digital elevation model (DEM) simulation for flood modeling in the absence of a high-accuracy open access global DEM. Front. Earth Sci. 6, 233.

Her, Y., Yoo, S.-H., Cho, J., Hwang, S., Jeong, J., Seong, C., 2019. Uncertainty in hydrological analysis of climate change: Multi-parameter vs. multi-GCM ensemble predictions. Sci. Rep. 9, 4974.

Higham, D.J., Higham, N.J., 2016. MATLAB Guide. SIAM.

Hong, Y., Liao, Q., Bonhomme, C., Chebbo, G., 2019. Physically-based urban stormwater quality modelling: An efficient approach for calibration and sensitivity analysis. J. Environ. Manag. 246, 462–471.

Huff, F.A., 1967. Time distribution of rainfall in heavy storms. Water Resour. Res. 3, 1007–1019.

Kollet, S.J., Maxwell, R.M., 2006. Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model. Adv. Water Resour. 29, 945–958.

do Lago, C.A.F., Giacomoni, M.H., Bentivoglio, R., Taormina, R., Junior, M.N.G., Mendiondo, E.M., 2023. Generalizing rapid flood predictions to unseen urban catchments with conditional generative adversarial networks. J. Hydrol. 618, 129276.

Li, W., Kiaghadi, A., Dawson, C., 2021. High temporal resolution rainfall-runoff modeling using long-short-term-memory (LSTM) networks. Neural Comput. Appl. 33, 1261–1278.

Lima, G.d., Boldrin, R.S., Mendiondo, E.M., Mauad, F.F., Ohnuma, Jr., A.A., 2007. Análise de incertezas de observações hidrológicas e sua influência na modelagem de pequenas bacias urbanas. Rev. Bras. Recur. Hídr. 12, 107–116.

Lindström, G., 1997. A simple automatic calibration routine for the HBV model. Hydrol. Res. 28, 153–168.

McClymont, K., Cunha, D.G.F., Maidment, C., Ashagre, B., Vasconcelos, A.F., de Macedo, M.B., Dos Santos, M.F.N., Júnior, M.N.G., Mendiondo, E.M., Barbassa, A.P., et al., 2020. Towards urban resilience through sustainable drainage systems: A multi-objective optimisation problem. J. Environ. Manag. 275, 111173.

Milledge, D.G., Lane, S.N., Warburton, J., 2009. The potential of digital filtering of generic topographic data for geomorphological research. Earth Surf. Process. Landf. 34, 63–74.

Moriasi, D.N., Gitau, M.W., Pai, N., Daggupati, P., 2015. Hydrologic and water quality models: Performance measures and evaluation criteria. Trans. ASABE 58, 1763–1785.

Naeini, M.R., Analui, B., Gupta, H.V., Duan, Q., Sorooshian, S., 2019. Three decades of the shuffled complex evolution (SCE-UA) optimization algorithm: Review and applications. Sci. Iran. 26, 2015–2031.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I–A discussion of principles. J. Hydrol. 10, 282–290.

Phillips, B.C., Yu, S., Thompson, G.R., De Silva, N., 2005. 1D and 2D modelling of urban drainage systems using XP-SWMM and TUFLOW. In: 10th International Conference on Urban Drainage. Citeseer, pp. 21–26.

Rossman, L.A., et al., 2010. Storm Water Management Model User's Manual, Version 5.0. National Risk Management Research Laboratory, Office of Research and ....

Rotava, J., Mendiondo, E.M., Souza, V.C.B., 2013. Simulação de instabilidade humana em inundações: Primeiras considerações. In: XX Simpósio Brasileiro de Recursos Hídricos. pp. 1–8.

Sarmento Buarque, A.C., Bhattacharya-Mis, N., Fava, M.C., Souza, F.A.A.d., Mendiondo, E.M., 2020. Using historical source data to understand urban flood risk: A socio-hydrological modelling application at Gregório Creek, Brazil. Hydrol. Sci. J. 65, 1075–1083.

Schwanghart, W., Scherler, D., 2014. TopoToolbox 2-MATLAB-based software for topographic analysis and modeling in Earth surface sciences. Earth Surf. Dyn. 2, 1–7.

SCS, U., 1986. Urban Hydrology for Small Watersheds. Technical Release No. 55 (TR-55), US Department of Agriculture, US Government Printing Office, Washington, DC.

Senarath, S.U.S., Ogden, F.L., Downer, C.W., Sharif, H.O., 2000. On the calibration and verification of distributed, physically-based, continuous, hortonian hydrologic models. Water Resour. Res. 36, 1495–1510.

Shafii, M., De Smedt, F., 2009. Multi-objective calibration of a distributed hydrological model (WetSpa) using a genetic algorithm. Hydrol. Earth Syst. Sci. 13, 2137–2149.

Shen, H., Tolson, B.A., Mai, J., 2022. Time to update the split-sample approach in hydrological model calibration. Water Resour. Res. 58, e2021WR031523.

Soliman, M., Morsy, M.M., Radwan, H.G., 2022. Assessment of implementing land use/land cover LULC 2020-ESRI global maps in 2D flood modeling application. Water 14, 3963.

Souza, T.F.d., 2008. Drenagem Urbana Sob Cenários de Longo Prazo Visando Incentivos Ambientais (Ph.D. thesis). Universidade de São Paulo.

Swathi, V., Srinivasa Raju, K., Varma, M.R.R., Sai Veena, S., 2019. Automatic calibration of SWMM using NSGA-III and the effects of delineation scale on an urban catchment. J. Hydroinform. 21, 781–797.

Tigkas, D., Christelis, V., Tsakiris, G., 2016. Comparative study of evolutionary algorithms for the automatic calibration of the Medbasin-D conceptual hydrological model. Environ. Process. 3, 629–644.

Torres, M.A., Chávez-Cifuentes, J.F., Reinoso, E., 2022. A conceptual flood model based on cellular automata for probabilistic risk applications. Environ. Model. Softw. 157, 105530.

Van Der Knijff, J.M., Younis, J., De Roo, A.P.J., 2010. LISFLOOD: A GIS-based distributed model for river basin scale water balance and flood simulation. Int. J. Geogr. Inf. Sci. 24, 189–212.

Wang, S., Taha, A.F., Sela, L., Giacomoni, M.H., Gatsis, N., 2020. A new derivative-free linear approximation for solving the network water flow problem with convergence guarantees. Water Resour. Res. 56, no–no.

Yang, C., Raskin, R., Goodchild, M., Gahegan, M., 2010. Geospatial cyberinfrastructure: Past, present and future. Comput. Environ. Urban Syst. 34, 264–277.

Young, I.T., Van Vliet, L.J., 1995. Recursive implementation of the Gaussian filter. Signal Process. 44, 139–151.

Zhang, H., Haan, C.T., Nofziger, D.L., 1990. Hydrologic modeling with GIS: An overview. Appl. Eng. Agric. 6, 453–458.

Zhang, T., Xiao, Y., Liang, D., Tang, H., Xu, J., Yuan, S., Luan, B., 2020. A physically-based model for dissolved pollutant transport over impervious surfaces. J. Hydrol. 590, 125478.