Revista Brasileira de Recursos Hídricos
Brazilian Journal of Water Resources
Versão On-line ISSN 2318-0331
RBRH, Porto Alegre, v. 30, e26, 2025
Scientific/Technical Article

https://doi.org/10.1590/2318-0331.302520240004

Impacts of spatial variability in urban hydrological modelling: a case study of the Canal do Mangue Basin in Rio de Janeiro, Brazil

Impactos da variabilidade espacial na modelagem hidrológica urbana: um estudo de caso da Bacia do Canal do Mangue no Rio de Janeiro, Brasil

James de Melo Sampaio Júnior[1], Marcelo de Miranda Reis[1] & Igor da Silva Rocha Paz[1]

[1]Instituto Militar de Engenharia, Rio de Janeiro, RJ, Brasil
E-mails: melo.james@ime.eb.br (JMSJ), marceloreis@ime.eb.br (MMR), igorpaz@ime.eb.br (ISRP)

Received: January 23, 2024 - Revised: February 05, 2025 - Accepted: March 03, 2025

ABSTRACT

This study analyzes spatial and rainfall variability of an urban basin using the Stormwater Management Model (SWMM), a semi-distributed and open-source hydrological tool. The Canal do Mangue Basin (CMB) in the city of Rio de Janeiro, Brazil, was modelled with three map resolutions, varying the numbers of sub-basins, and three rainfall spatialization methods: arithmetic mean, Thiessen polygons, and Inverse Distance Weighted (IDW). Simulated water levels were compared to observed data at control points, considering different rain gauge densities. Results showed good agreement between observed and simulated volumes. The arithmetic mean method performed best overall, with errors ranging from 0.00% to 18.60%, while IDW showed higher accuracy for specific nodes (10.82% to 15.18%). The study concludes that spatial resolution, interpolation methods, and rain gauge density significantly affect model performance, providing valuable insights for optimizing hydrological modeling in urban basins.

Keywords: Urban hydrological modelling; Spatial variability; Rain gauge; Land use and land cover; SWMM.

RESUMO

Este estudo analisa a variabilidade espacial e pluviométrica de uma bacia urbana utilizando o *SWMM* (*Stormwater Management Model*), um modelo hidrológico semi-distribuído e de código aberto. A Bacia do Canal do Mangue (BCM), localizada no Rio de Janeiro-RJ, Brasil, foi modelada empregando três resoluções de mapas, variando o número de sub-bacias, e três métodos de espacialização pluviométrica: média aritmética, polígonos de Thiessen e *Inverse Distance Weighted* (*IDW*). Os níveis d'água simulados foram comparados com dados observados em pontos de controle da bacia, considerando diferentes densidades de pluviômetros. Os resultados indicaram boa aproximação geral entre os volumes observados e simulados. O método da média aritmética apresentou o melhor desempenho geral, com erros variando de 0,00% a 18,60%, enquanto o IDW demonstrou maior precisão em nós específicos (10,82% a 15,18%). O estudo conclui que a resolução espacial, os métodos de interpolação e a densidade de pluviômetros afetam significativamente o desempenho do modelo, fornecendo informações valiosas para otimizar a modelagem hidrológica em bacias urbanas.

Palavras-chave: Modelagem hidrológica urbana; Variabilidade espacial; Pluviômetro; Uso e cobertura do solo; *SWMM*.

INTRODUCTION

In the context of flood prediction, hydrological modelling has become a significant tool, gaining increasing prominence in recent years (Perera et al., 2022; Yang et al., 2023; Kalakuntla & Umamahesh, 2025). However, in the course of model development, several challenges arise, including the lack of adequate land use and land cover data with resolutions below the required level, and the demand for more precise precipitation data. These factors directly impact the efficiency of the models (Santos, 2018; Hu et al., 2020; Guzder-Williams et al., 2023).

Regarding spatial relationships, hydrological models can be classified as lumped, semi-distributed, and distributed (Singh, 2018), with the latter being most suitable for representing interactions among processes and simulating effects resulting from land use changes, despite limitations in flood predictions (Beven, 1991). Considering the need for incorporating a vast amount of data into distributed models, coupled with restricted accessibility to most, semi-distributed models such as SWMM, HEC-HMS, TOPMODEL, etc., can be employed.

The Stormwater Management Model (SWMM) is a dynamic semi-distributed rainfall-runoff model developed in 1971 by the U.S. Environmental Protection Agency (EPA) (U.S. Environmental Protection Agency, 2022). It operates as an open-access and open-source model, with the most recent version, SWMM 5.2.4, released on August 7, 2023, used utilized in this study. Concerning the spatial variability of the study area, the software allows the division of the area into smaller, homogeneous water catchment areas called sub-basins. Each sub-basin consists of permeable and impermeable sub-areas, with surface runoff routed among these sub-basins or between points of entry in a drainage network. What distinguishes SWMM from other urban basin models is its emphasis on water conveyance systems for both stormwater and wastewater, incorporating considerations for the design and performance of sewage systems (Niazi et al., 2017). Additionally, the model exhibits significant flexibility in simulating hydrological events in urban areas compared to other models (Shinma, 2015).

In this context, this study aimed to conduct hydrological modelling of an urban area experiencing extreme hydrological events, intending to investigate the effects of spatial discretization scale and rain distribution variability on modelling outcomes. To achieve this, the SWMM model was applied to three maps with different resolutions (95, 48, and 27 sub-basins) of the Canal do Mangue Basin (CMB), located in the eastern region of the city of Rio de Janeiro, Brazil. Rainfall data from stations within the Alerta Rio System were spatially interpolated using three methods: arithmetic mean, Thiessen Polygons, and Inverse Distance Weighted (IDW). These methods were chosen for their complementary characteristics and relevance in spatial rainfall representation. The arithmetic mean method, widely used due to its simplicity, provided a baseline analysis of average precipitation without considering spatial variability. Thiessen Polygons were applied to account for the influence area of each station, a critical advantage in regions with irregular rainfall station distribution. IDW offered a more refined interpolation, weighting values based on inverse distance to capture spatial rainfall variation more accurately. By employing these three techniques, the study achieved a robust comparative analysis, enhancing the reliability and comprehensiveness of the spatial rainfall representation in the study area. Furthermore, the study compared the results based on the utilization of varying quantities of rain gauges, conducting this comparison concerning water levels observed at control points within the CMB.

Rainfall spatialization in hydrological modelling presents critical challenges, particularly in urban basins prone to extreme hydrological events. The choice of interpolation methods and the density of rain gauge networks are pivotal factors affecting model performance (Brocca et al., 2020; Paz et al., 2020). As highlighted by Suri & Azad (2024), optimizing rain gauge placement in complex terrains is essential for monitoring extreme rainfall events effectively. Similarly, Workneh et al. (2024) emphasize the significance of selecting appropriate interpolation techniques, such as IDW, for capturing spatial rainfall variability in diverse regimes. Thiessen Polygons and arithmetic mean methods, although simpler, continue to serve as foundational approaches in rainfall spatialization. Furthermore, spatial discretization remains a central concern, as it influences the model's ability to represent land use impacts and water flow pathways (Souza et al., 2018; Campos & Paz, 2020; Borges et al., 2021; Cao et al., 2020). Kalakuntla & Umamahesh (2025) demonstrate that refined calibration strategies are indispensable for improving flood prediction accuracy in large catchments. Building on this context, our study investigates the effects of spatial discretization, rainfall interpolation methods, and rain gauge density on hydrological modelling outcomes, aiming to bridge gaps in urban flood prediction and optimize model performance through a robust methodological framework.

The objectives of this study can be categorized as follows: (i) comparing the outcomes of hydrological modelling while varying the model's discretization (number of sub-basins); (ii) comparing the outcomes of hydrological modelling while varying the method of spatializing rainfall data collected from rain gauges; and (iii) comparing the outcomes of hydrological modelling while varying the quantity of rain gauges used in each spatialization method. This study presents the hydrological modelling of an urban basin, detailing the methodology employed. Subsequently, the results are discussed, focusing on comparisons among the generated model responses. Finally, conclusions are drawn and presented.

MATERIALS AND METHODS

Study area

The study area employed in this research is the Canal do Mangue Basin (CMB), with UTM coordinates 675701E, 7458905S; 685816E, 7466951S – Zone 23S. It encompasses an approximate area of 45.4 km^2 and is located in an urban setting within the city of Rio de Janeiro, in the Southeast region of Brazil, as illustrated in Figure 1. The CMB has a history of flooding, notably in 2010, which resulted in 219 fatalities and over 11,000 displaced individuals (Carissimi, 2011; Maia & Sedrez, 2011; Ferreira & Cunha, 1996).

In the plains of Rio de Janeiro, precipitation is evenly distributed, ranging between 1000 to 1200 mm annually. Typically, the wettest period spans from November to January, while the least rainy period extends from June to August. Concerning the

climate, the city features a hot, superhumid, sub-dry climate with an average temperature of 23.6 °C, experiencing variations of approximately 7 °C throughout the year (Instituto Brasileiro de Geografia e Estatística, 2022).

The hydrography in the region flows from west to east, comprising eighteen main water bodies. The Maracanã River (10.1 km) stands as the primary watercourse with the highest number of tributaries, followed by the Joana River (5.5 km). The Canal do Mangue, a 3 km-long artificial canal, receives runoff from the CMB and discharges into the Guanabara Bay in the northeast region of the basin.

The CMB is bordered to the west and southwest by the Tijuca Massif, which encompasses the headwaters of the Maracanã, Joana, Trapicheiros, and Comprido rivers. The Digital Elevation Model (DEM) was obtained from the Brazilian National Institute for Space Research, called INPE, and has a scale of 12.5 meters, as presented in Figure 2. Notably, most of the basin lies in an urbanized and flat area, with elevations reaching up to 40 meters.

Regarding soil type, 63.1% of the area is classified as urbanized, followed by 24.4% of Class A soil, 9.5% of Class C soil, and 3.0% of Class B soil, based on the hydrological classification using the Curve Number (CN) infiltration model (Soil Conservation Service, 1956).

The land use and land cover maps, along with vegetation maps for the study area, were acquired from the Pereira Passos Institute (IPP) and were updated in 2019, provided in vector format (shapefile). Following the reconciliation of these maps, the categories obtained were water bodies (0.58 km^2), open soil (0.36 km^2), built-up areas (25.73 km^2), forested areas (16.83 km^2), and grasslands (1.90 km^2). The consolidated map is presented in Figure 3.

It is important to highlight that the urban drainage system of the CMB was unavailable and, as a result, was not considered in the models. Consequently, surface runoff occurred along slopes, and drainage took place into the rivers and channels present in the basin's hydrography. Additionally, bathymetries of these elements were not provided, leading to the utilization of intrinsic data from maps and visual inspection to define their widths.

Storm Water Management Model (SWMM)

According to the objectives of this study, three different resolutions were selected, and the maps were developed using GIS tools along with the collected DEM. For the most finely discretized map (R1), 95 sub-basins were obtained, each with an average area of 0.46 km^2. In map R2, sub-basins from R1 were aggregated based on topography and their respective outlets, resulting in 48 sub-basins with average areas of 0.92 km^2. Lastly, in map R3, an aggregation of sub-basins from R2 yielded 27 sub-basins with average areas of 1.63 km^2.

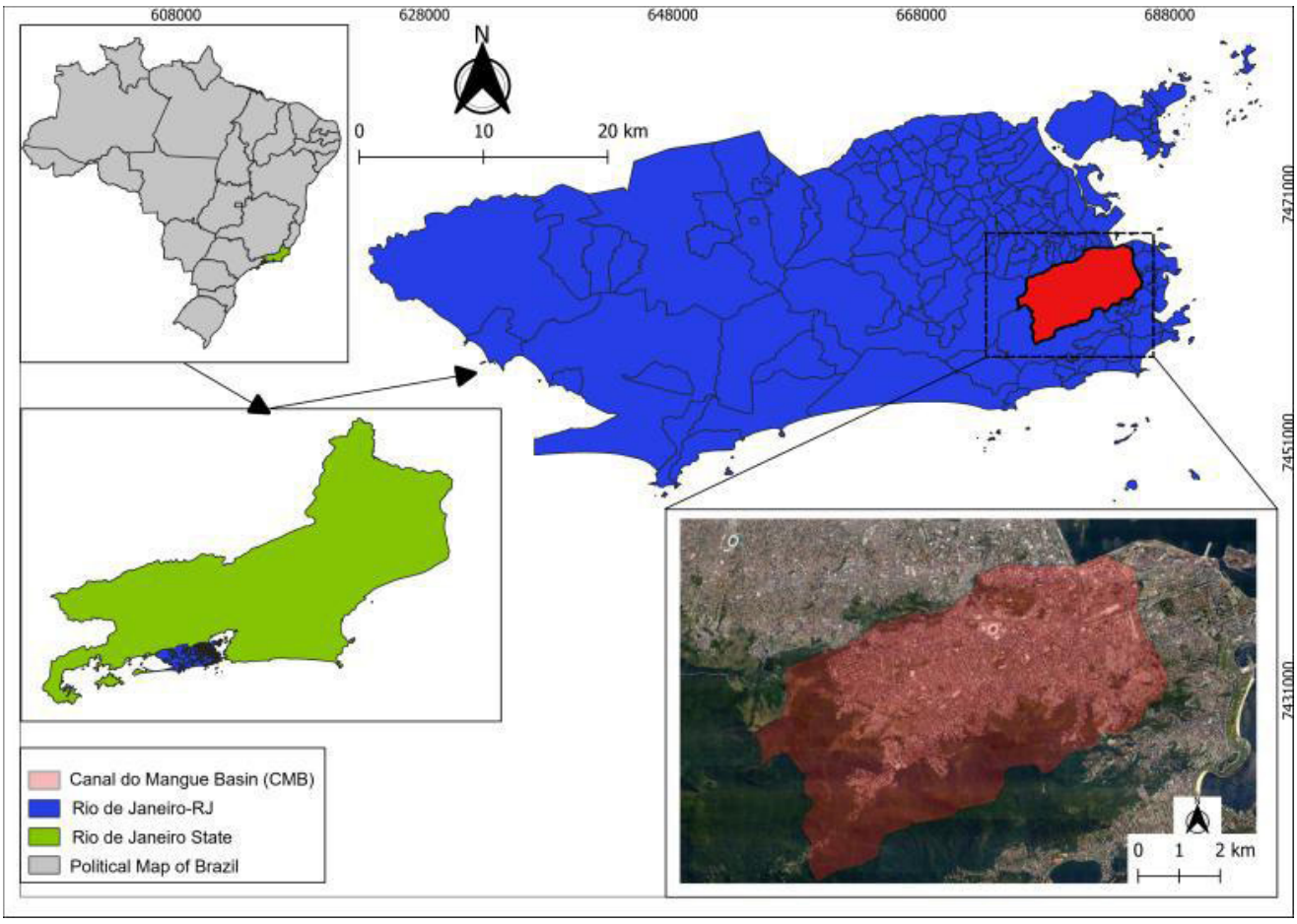

Figure 1. Location of the Canal do Mangue Basin (CMB).

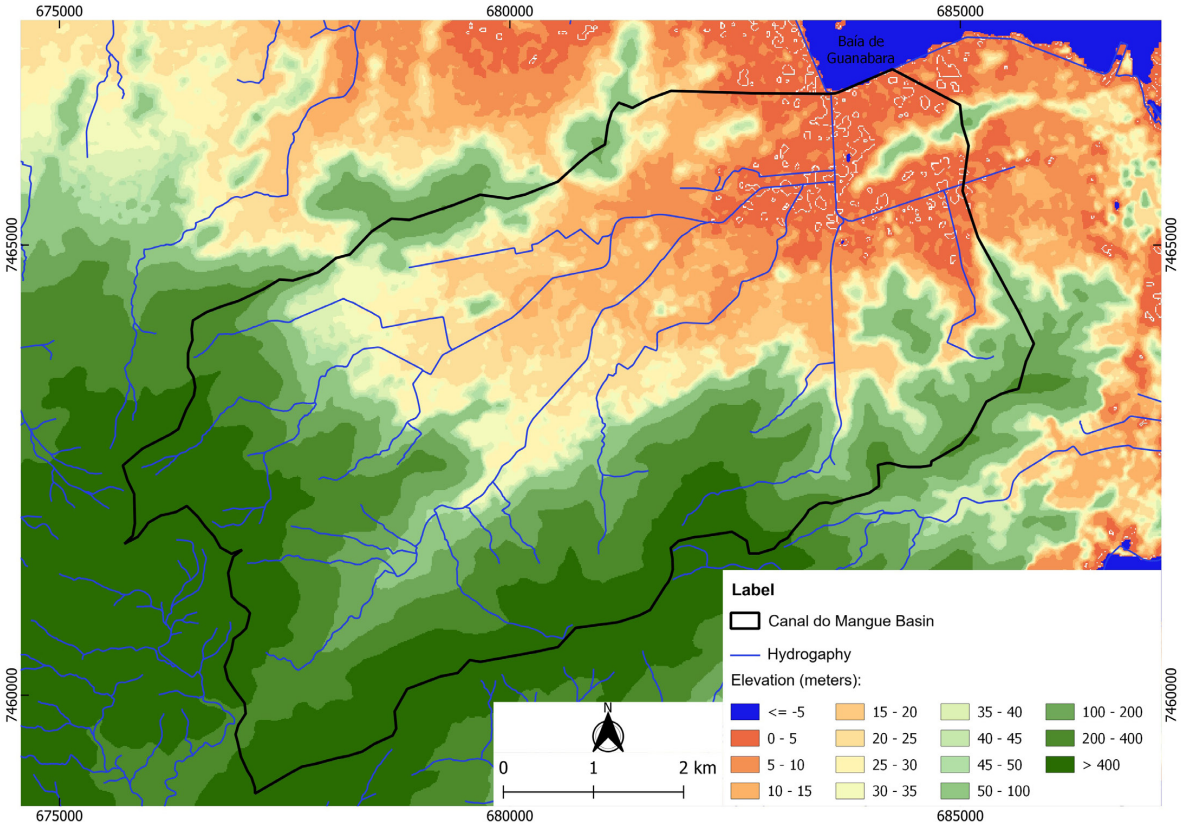

Figure 2. Topography of the CMB.

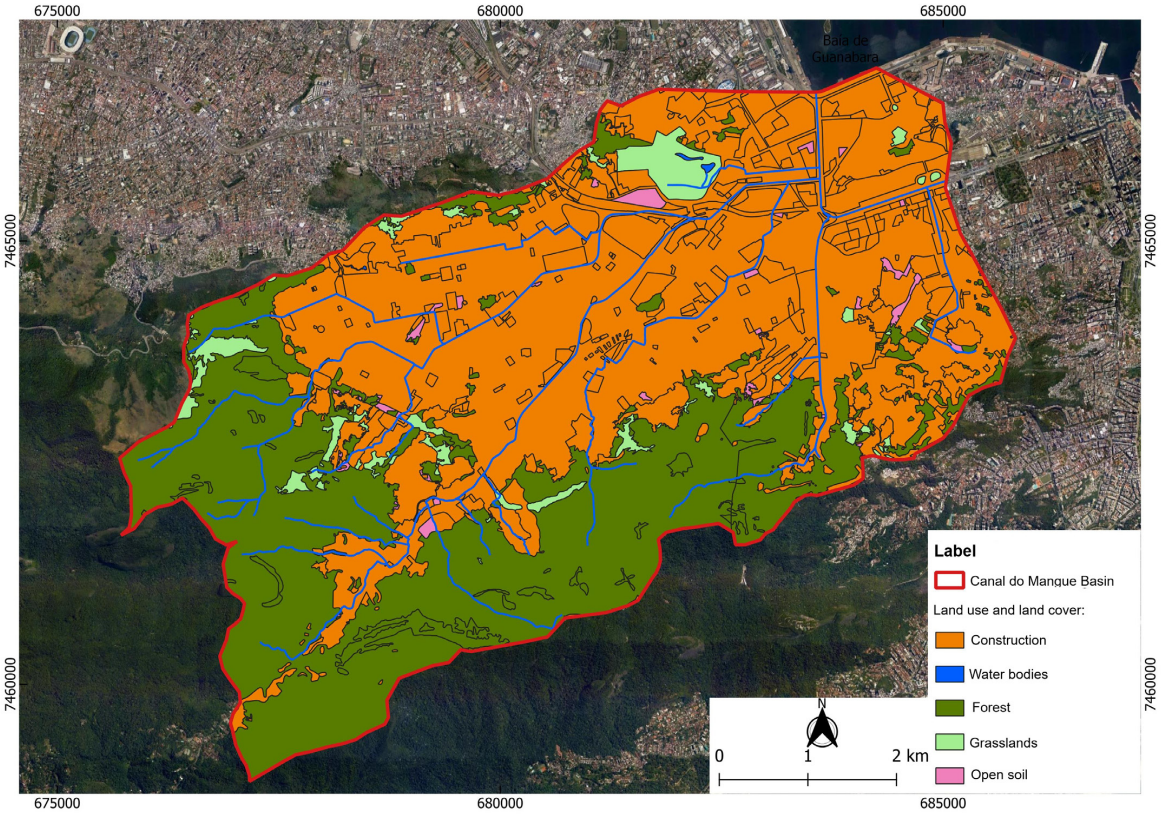

Figure 3. Land use and land cover map over the CMB area.

The nodes and conduits included in SWMM were initially plotted using QGIS, based on maps of watercourses and drainage direction generated from the DEM. Furthermore, outlets for each sub-basin were defined.

Finally, drawings were created for the following objects:

a. 95, 48, and 27 sub-basins for maps R1, R2, and R3, respectively;
b. 64 junction nodes and one outlet node for all three resolutions;
c. 64 conduits, comprising 48 with trapezoidal cross-sections and 16 with rounded rectangular cross-sections, for all three resolutions;
d. 102, 55, and 34 rain gauges for maps R1, R2, and R3, respectively. Among these, one virtual rain gauge was drawn for each sub-basin at its geometric center for use with the IDW method; six real rain gauges representing the utilized rainfall stations (PBV, PGR, PSD, PST, PTJ, and PTM); and one virtual rain gauge for use in the arithmetic mean method (P-MÉDIA), situated at the geometric center of the CMB.

For the sub-basins, the fixed (non-calibratable) physical parameters included area, maximum length, width, slope, sub-basin roughness, storage depth in depressions, and the percentage of impermeable area without storage. Slope determination utilized the direct method relying on extreme values (Paz & Collischonn, 2007).

For impermeable area roughness, a constant value of 0.013 was considered, corresponding to the "normal concrete" roughness from Appendix A.6 of the SWMM Manual (Universidade Federal da Paraíba, 2012). To calculate the permeable area roughness, a weighted average considering the area of each land use category was used. The roughness values for each category were constant: 0.60 for forest, 0.28 for grass, and 0.05 for open soil. These values were also derived from Appendix A.6 of LENHS/UFPB (Universidade Federal da Paraíba, 2012), considering the average of two forest categories (light brush - 0.40 and dense brush - 0.80) for the "forest" class; the average of minimum (0.15) and maximum (0.41) values for the "grass" class; and the value for fallow soil for the "open soil" class. Thus, the roughness for impermeable areas remained constant at 0.013, while the roughness for permeable areas reached 0.5936, 0.5788, and 0.5757 for maps R1, R2, and R3, respectively. It is important to note that in areas entirely covered by constructions, due to the calculation method and model insertion, the roughness value considered for permeable areas was 0.00, since forest, grass, and open soil areas were null.

For the infiltration model, the Curve Number (CN) method (Soil Conservation Service, 1956) was used, following the approach by Bae and Lee (2020). Their approach used the CN method to estimate infiltration in densely populated urban areas with significant impermeable surfaces. By integrating the method into a hydrological model, they simulated the effects of Low-Impact Development (LID) strategies, such as green roofs and permeable pavements, using rainfall data from 2010 and 2011 flood events in Seoul, South Korea.

In this study, CN values for sub-basins were calibratable. Initially, these values were defined based on soil type, its coverage, and according to Table A.4 of LENHS/UFPB (Universidade Federal da Paraíba, 2012). Table 1 presents the values considered in this initial phase. It is highlighted that the value for urban soil is independent of the hydrological class (A, B, or C), being defined as the highest value for residential areas from Table A.4 of the Manual.

For conduits, the fixed (non-calibratable) parameters considered were length, initial and final elevations, and cross-sectional shape. The conduit roughnesses were regarded as calibratable, and these values were initially defined based on the longitudinal arrangement of river and channel sections, following the approach by Rezende (2018). In this case, the following initial values were chosen: 0.015 for straight sections, 0.039 for sections with smooth curves, and 0.070 for sections with sharp curves. These values are also consistent with those presented in Appendix A.8 of LENHS/UFPB (Universidade Federal da Paraíba, 2012).

For the simulation, the dynamic wave flow propagation model was considered, accounting for node flooding, a minimum conduit slope of 0.2%, and a propagation time step of 30 seconds. Finally, Table 2 presents the ranges of values inserted for each variable.

Rainfall data

The rainfall data utilized in the hydrological modelling were collected from the Rainfall Alert System of the Rio de Janeiro City Hall (Rio de Janeiro, 2023), which comprises 33 rain gauge stations distributed across the municipality. These stations provide real-time data every 15 minutes. Currently, there are seven stations monitoring rainfall in the CMB region, with three of them (PBV, PSC, and PSD) located outside the CMB, as presented in Figure 4.

The calibration event utilized occurred from 8:00 PM on February 16, 2000, to 7:30 PM on February 17, 2000, spanning 23 hours and 30 minutes, as per the model calibration conducted by Rezende (2018). This event, referred to in this study as Event 1, utilized data from four Alerta Rio stations (PGR, PSD, PST, and PTJ), which were operational at that time. Figure 5 displays the rainfall series utilized.

The event selected for validation occurred between 5:30 PM and 11:30 PM on March 12, 2016, spanning 6 hours, as per the modelling conducted by Sousa et al. (2022). For this period, three events were defined using different rain gauge networks: Event 2, utilizing four rain gauges (PGR, PSD, PST, and PTJ); Event 3, utilizing five rain gauges (PGR, PST, PTJ, PBV, and PTM); and Event 4, utilizing six rain gauges (PGR, PST, PSD, PTJ, PBV, and PTM). The hyetographs used for each station are presented in Figure 6.

Table 1. Initial CN Parameters for each Soil Hydrological Class.

Hydrological Classes	Forest Area > 70%	Forest Area < 70%
A	25	45
B	55	66
C	70	77
Urban	92	92

Table 2. Physical Parameters inserted in the SWMM Model.

Property	Map R1	Map R2	Map R3
Sub-basin Area (ha)	0.83 to 168.29	7.61 to 231.81	61.39 to 394.89
Sub-basin Width (m)	50.61 to 1369.33	198.83 to 1854.62	515.19 to 1745.68
Sub-basin Slope (%)	0.24 to 59.09	0.50 to 48.21	0.68 to 48.14
Average Sub-basin Slope (%)	15.09	18.40	19.25
Impervious Area (%)	0.00 to 100.00	0.88 to 99.80	0.88 to 98.37
Watercourse Roughness	0.015 to 0.070	0.015 to 0.070	0.015 to 0.070
Average Watercourse Slope (%)	2.80	2.80	2.80
Curve Number (CN)	25 to 92	25 to 92	25 to 92

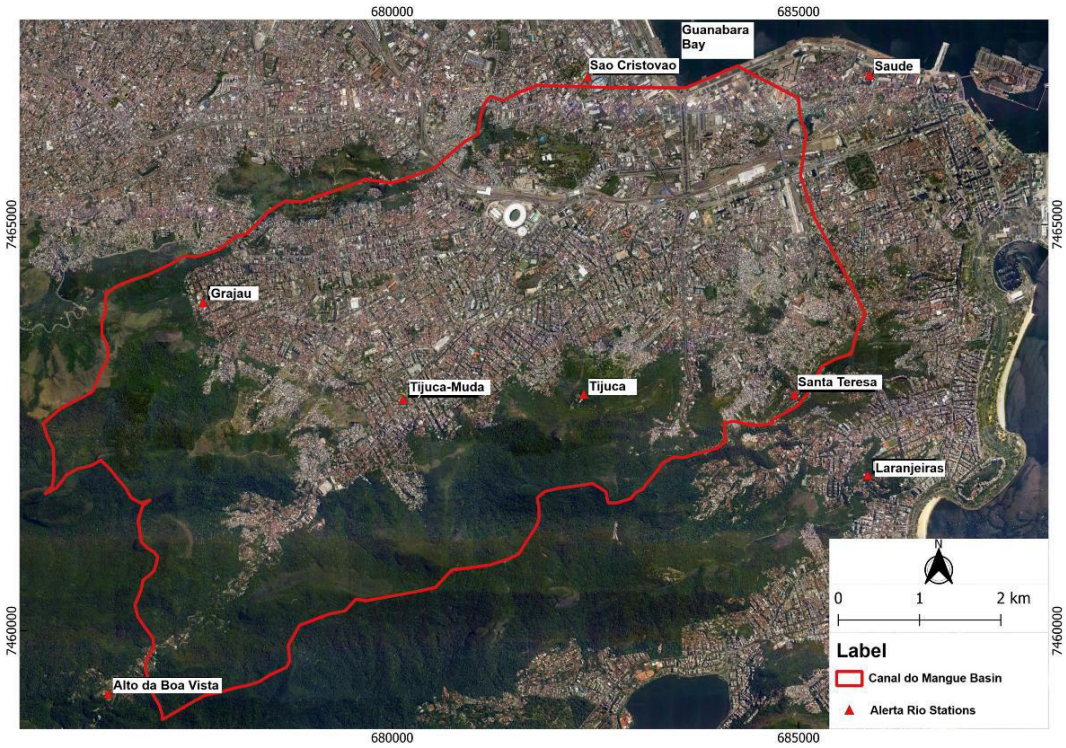

Figure 4. Alerta Rio stations located in Canal do Mangue basin.

For the calibration, the Inverse Distance Weighting (IDW) method was applied using QGIS, with a pixel size of 10 meters and the power parameter p=2, in accordance with the recommendations of Landim (2000) and Mello et al. (2003). The procedure was performed for each 15-minute interval, totaling 95 processes, resulting in the hyetographs for the event in each sub-basin. This procedure was conducted for the three maps R1, R2, and R3. Additionally, to compare water levels at control points, the calibrated models were run using spatialized rainfall data from Event 1 through the arithmetic mean and Thiessen polygons methods, generating six additional models.

For the validation, the methods employed were the arithmetic mean (uniformly distributed precipitation), Thiessen polygons, and the IDW method. For the IDW method, the interpolated hyetographs were associated with virtual rain gauges located at the geometric centers of each sub-basin.

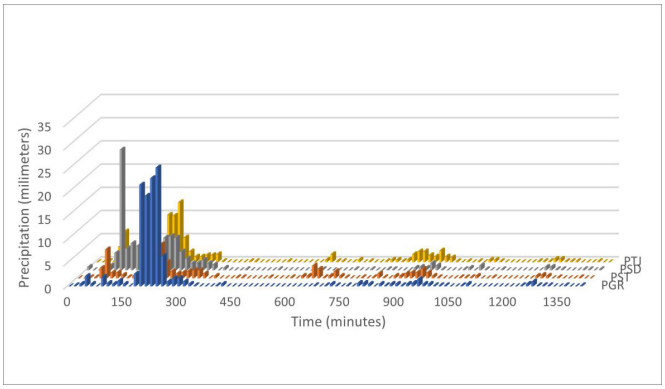

Figure 5. Hyetographs of Event 1 (calibration) from stations PGR, PST, PSD, and PTJ, occurring between 8:00 PM on 16/02/2000 and 7:30 PM on 17/02/2000.

Model calibration

During the calibration process, CN ranges were employed based on the inspection of water level versus time curves at control points. The primary objective was to match the maximum depth and the behavior of simulated curves with observed ones by adjusting the CN within the limits presented in Table 3, which were derived from Appendix A.4 of the SWMM Manual (Universidade Federal da Paraíba, 2012).

The conduit roughness values were also adjusted for map R1 (more discretized). Initially, these values were defined based on the longitudinal arrangement of river and canal sections, following Rezende (2018). Subsequently, the conduit roughness was adjusted within the range of 0.015 to 0.070 to align the simulated and observed curves at the control points. After calibrating map R1, the conduit roughness values were set for the other resolutions. Figure 7 presents a flowchart describing the calibration process.

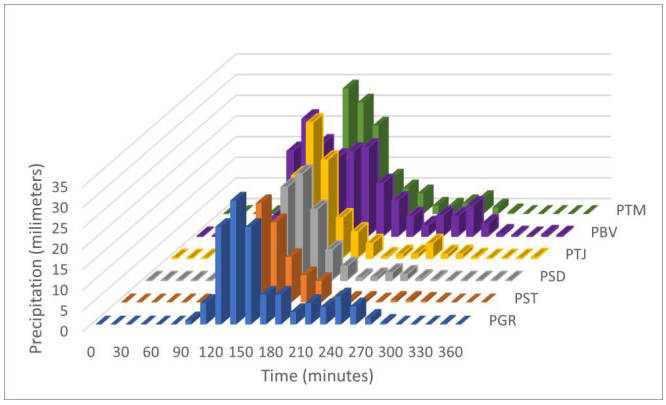

Figure 6. Hyetographs used in Events 2, 3, and 4 (validation) from stations PGR, PST, PSD, PTJ, PBV, and PTM, occurring between 5:30 PM and 11:30 PM on 12/03/2016.

During both the calibration and validation events, water levels resulting from the modelling in the rivers and channels of the CMB were compared with in-situ measurements at six stream gauge stations, taken in 2000 (Rezende, 2018) and 2016 (Sousa et al., 2022), respectively. Three out of the six control points from the calibration event were part of the set of points used in the validation events, namely N-19, N-35, and N-42. The sets of control points for both the calibration and validation events are presented in Figure 8.

Generated models

The generated models in SWMM are presented in Table 4. These models were created according to the resolutions (R1, R2, and R3), the methods of rainfall spatialization (arithmetic mean – M, Thiessen – T, and IDW – I), and the events used (1 - calibration; 2 - validation with four rain gauges; 3 - validation with five rain gauges; and 4 - validation with six rain gauges). The calibration models used were R1-1-I, R2-1-I, and R3-1-I, indicating that three different resolutions were applied for the calibration event (Event 1), all using the IDW spatialization method, involving changes to the CN of the sub-basins and the conduit roughness. After calibrating these, the arrangement was maintained, generating new models by only altering the input rainfall data.

Table 3. Range of CN values for sub-basins in the calibration process based on the SWMM model.

Hydrological Classes	Lower Limit	Upper Limit
A	25	72
B	55	81
C	70	88
Urban	51	92

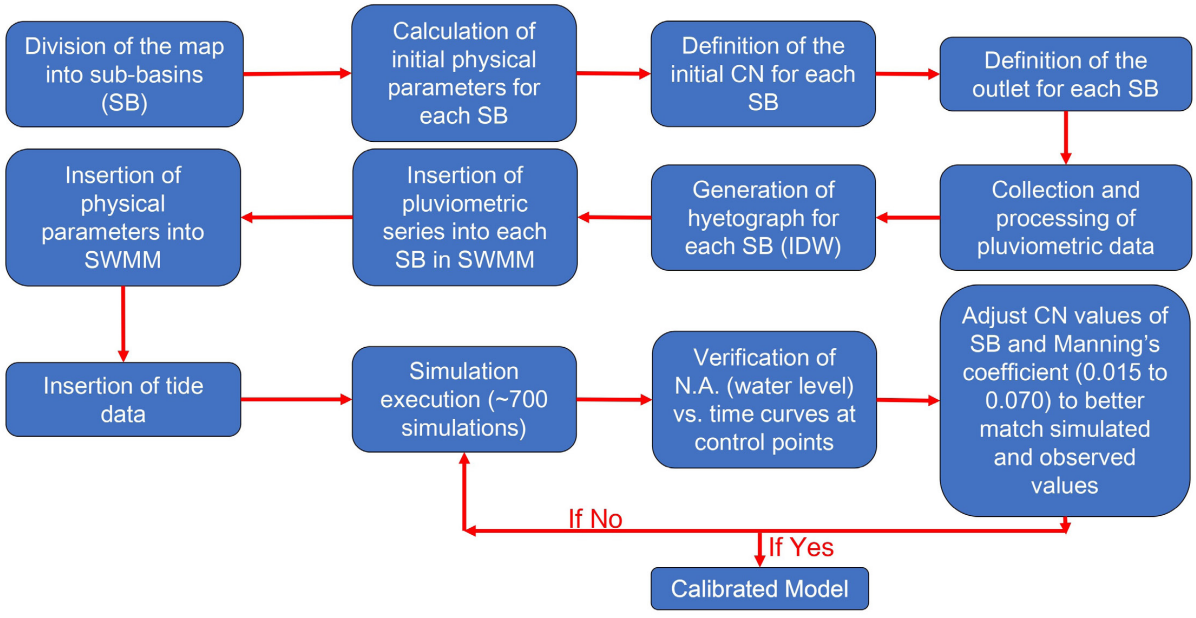

Figure 7. Calibration flowchart.

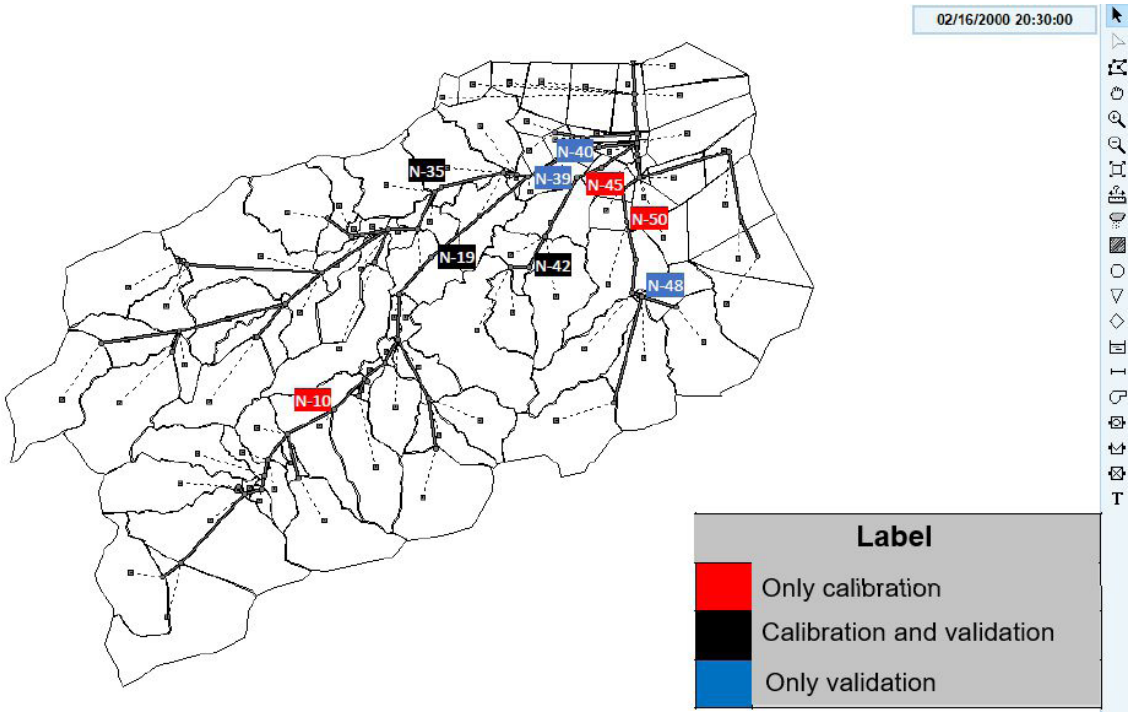

Figure 8. Control points utilized in the calibration and validation events represented on Map R1 in SWMM.

Table 4. Models generated in SWMM.

Event	Spatialization Method	Resolution 1	Resolution 2	Resolution 3
Event 1	Arithmetic Mean	R1-1-M	R2-1-M	R3-1-M
	Thiessen Polygons	R1-1-T	R2-1-T	R3-1-T
	IDW	R1-1-I	R2-1-I	R3-1-I
Event 2	Arithmetic Mean	R1-2-M	R2-2-M	R3-2-M
	Thiessen Polygons	R1-2-T	R2-2-T	R3-2-T
	IDW	R1-2-I	R2-2-I	R3-2-I
Event 3	Arithmetic Mean	R1-3-M	R2-3-M	R3-3-M
	Thiessen Polygons	R1-3-T	R2-3-T	R3-3-T
	IDW	R1-3-I	R2-3-I	R3-3-I
Event 4	Arithmetic Mean	R1-4-M	R2-4-M	R3-4-M
	Thiessen Polygons	R1-4-T	R2-4-T	R3-4-T
	IDW	R1-4-I	R2-4-I	R3-4-I

RESULTS AND DISCUSSIONS

Calibration results

Table 5 presents the comparison between the maximum water levels observed in situ at their control points and the simulation values for these models. A 15% threshold for the percentage differences between the maximum water levels was set to define nodes with satisfactory results, showing outcomes for nodes N-19, N-35, N-45, and N-50.

It is worth noting that the results for nodes N-10 and N-42 were not presented due to significantly high percentage differences. This may be attributed to N-10's more upstream location in the basin and N-42's upstream position in its section, along with gaps

Table 5. Maximum water levels in the calibrated nodes observed in situ and simulated in maps R1, R2, and R3 for Event 1.

Nodes		N-19	N-35	N-45	N-50
Observed Max Water Level (m)		9.00	7.90	2.30	7.10
R1-1-I (280 times)	Simulated Max (m)	9.04	7.79	2.03	6.30
	Abs. Diff. (m)	0.04	-0.11	-0.27	-0.80
	Percentage Diff. (%)	0.44%	-1.39%	-11.74%	-11.27%
R2-1-I (222 times)	Simulated Max (m)	9.03	7.78	2.10	6.27
	Abs. Diff. (m)	0.03	-0.12	-0.20	-0.83
	Percentage Diff. (%)	0.33%	-1.52%	-8.70%	-11.69%
R3-1-I (186 times.)	Simulated Max (m)	8.99	7.82	2.01	6.27
	Abs. Diff. (m)	-0.01	-0.08	-0.29	-0.83
	Percentage Diff. (%)	-0.11%	-1.01%	-12.61%	-11.69%

in the drainage system. Furthermore, the calibration was focused on peak levels, considering that calibration based on discharged volumes could potentially be a more suitable option.

Validation results

Event 1

For the calibration's well-performing control points in Event 1 (N-19, N-35, N-45, and N-50), the percentage differences between the simulated and observed maximum water levels are shown in Table 6.

The simulated volumes matched the observed volumes for nodes N-19, N-35, and N-45. However, for node N-50, the simulated volumes are lower than the observed volumes, indicating the use of a CN lower than the actual, leading to greater infiltration in the models. In all nodes, peak anticipations were observed, with variations from 20 to 80 minutes, demonstrating that a higher volume of water propagated in defined channels than in slopes.

In comparison to the results from the fully distributed model presented by Rezende (2018), there was a better fit of the simulated curves in shape and peak for nodes N-19, N-35, and N-45, indicating that the calibration in the semi-distributed model yielded satisfactory results. In the calibration process of this research, sub-basins with the same classes of land use and land cover might present different CN values, which doesn't occur in fully distributed models. This is justified because CN alterations were made based on the predominant land use and land cover category for each sub-basin and in accordance with the values presented in the Manual's tables (Universidade Federal da Paraíba, 2012). However, the calibration process in this format was time-consuming, involving nearly 700 simulations for the three models. Additionally, the SCS CN infiltration model appeared subjective, relying on visual inspection and the operator's perception.

Table 6. Range of percentage differences between simulated and observed maximum water levels for nodes N-19, N-35, N-45, and N-50 for Event 1.

Node	Lower difference	Upper difference
N-19	1.25% (R3-1-M)	4.09% (R2-1-T)
N-35	0.89% (R1-1-T)	7.72% (R1-1-M)
N-45	4.35% (R2-1-T)	12.61% (R3-1-I)
N-50	9.15% (R1-1-M)	12.11% (R2-1-T)

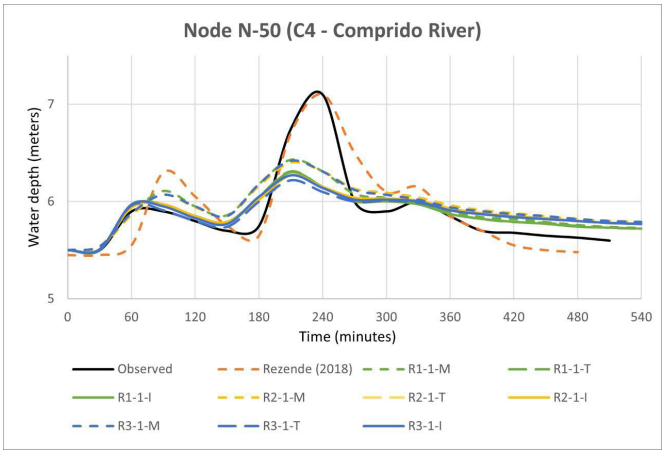

Figure 9. Simulated and observed water levels at node N-50.

For node N-50, the models using the arithmetic mean achieved a higher maximum peak compared to the other models. Nevertheless, there was an overestimation of the first peak and a delay, as shown in Figure 9. Thus, it is evident that solely comparing the simulated maximum peaks with the observed ones does not

conclusively define which spatialization method would be most suitable. It is essential to also consider the attainment of other peaks and the curve's overall adherence in shape and time.

Events 2, 3 and 4

For the validation control points (N-19, N-35, N-39, N-40, N-42, and N-48), the percentage differences between the simulated and observed maximum water levels are presented in Table 7.

The simulated and observed volumes were in agreement for nodes N-35, N-39, and N-40. However, for node N-19, located more upstream, the simulated volumes were below the observed volume, indicating greater infiltration in the sub-basins that discharge at this node. On the other hand, at nodes N-42 and N-48, located near the outlet in predominantly urban areas, the simulated volumes were greater than the observed volumes, suggesting lower infiltration in the sub-basins discharging at these nodes.

In all control points, anticipations in peaks were observed, with variations of 40 to 80 minutes, indicating a greater volume of water propagated in defined channels than on slopes.

Comparing the results to the fully distributed model presented by Sousa et al. (2022), the simulated curves in SWMM had less adherence in both shape and peak, despite both presenting anticipations of the maximum peak for nodes N-19, N-39, and N-40. For node N-35, the curves in SWMM had less adherence in shape but reached peaks similar to the observed values. For node N-42, the simulated curves had less adherence in volume and peak, despite good approximation in shape. This node is located upstream of its stretch and is in a predominantly urban area, potentially affected by the gap in the urban drainage system. For node N-48, compared to the curve obtained by the fully distributed model, the simulated curves had less adherence in volume, shape, and peak. The results in this node may have been affected by the gap in the drainage system since the location is in a predominantly urban area.

In summary, there was similarity in the curve results when varying the discretization, with few differences in the maximum peaks obtained. Nevertheless, comparing these values, the more discretized models (R1) presented smaller errors, corroborating the hypothesis that greater discretization generates smaller differences compared to the observed. However, it should be noted that greater discretization would lead to greater differences in the comparison between models.

Table 7. Range of percentage differences between simulated and observed maximum water levels for nodes N-19, N-35, N-39, N-40, N-42, and N-48 for Events 2, 3, and 4.

Node	Lower difference	Upper difference
N-19	10.82% (R1-3-I and R2-3-I)	13.73% (R3-2-M)
N-35	0.00% (R3-2-T, R3-3-I and R3-4-I)	2.07% (R1-3-I and R1-4-I)
N-39	9.12% (R2-2-M and R3-2-M)	22.35% (R3-1-I)
N-40	6.20% (R1-2-M, R2-2-M and R3-2-M)	20.80% (R3-3-I)
N-42	18.60% (R1-2-M)	27.29% (R2-3-I)
N-48	15.18% (R1-3-M)	16.96% (R3-2-T, R3-3-T and R3-4-T)

Regarding the hypothesis that there is an optimal resolution from which the differences between the models are minimal (below 5%, for example), it was not possible to confirm such resolution with the discretization used in this study. Future studies are suggested to develop models with a greater number of sub-basins, considering maximum areas of 0.10 km^2, which would generate approximately 450 sub-basins.

The hypotheses that changes in rainfall spatialization methods and the number of rain gauges impact on model outputs were confirmed, as variations in simulated water levels were observed at control points. This finding aligns with previous studies, which demonstrate that interpolation methods (Chen et al., 2017; Workneh et al., 2024), rain gauge density (Paula et al., 2018; Suri & Azad, 2024), and spatial rainfall resolution (Huang et al., 2019; Paz et al., 2019) significantly affect hydrological simulations. However, in this study, these differences were small due to similar hyetographs among stations. To better assess the effects of spatial rainfall variability, using an extreme event with greater hyetograph differences is recommended. Therefore, it is not possible to definitively determine which spatialization method presents lower errors or the ideal amount of rain gauges. Nevertheless, when only comparing the simulated maximum peaks with the observed ones, it was found that uniformly distributed precipitation (arithmetic mean) presented the lowest percentage differences for nodes N-35, N-39, N-40, N-42, and N-48. Next, the IDW method presented the lowest differences in nodes N-19 and N-35. Finally, it was observed that Event 2 (use of 4 rain gauges) presented the lowest percentage differences in relation to the maximum water level for nodes N-35, N-39, N-40, and N-42. Next, Event 3 (use of 5 rain gauges) presented the lowest differences in nodes N-19 and N-48.

It is reiterated that the differences between the simulated results in this work and those modeled by Rezende (2018) and Sousa et al. (2022) can be explained by the use of different models and the greater discretization of the mesh in these works, which more coherently encompass the heterogeneity of the basin and internal flows. Semi-distributed models, such as SWMM, treat the hydrographic basin in a more aggregated way, which can lead to the loss of crucial details in modelling and, consequently, to lower calibration.

Although the calibration process in the semi-distributed model is more complete, considering different CN values, for example, the fully distributed model presented curves closer to those observed, indicating better calibration. This can be justified using more precise discretization and implementation of the urban drainage system, allowing the model to adjust more closely to the observed data during the calibration process, improving its ability to replicate the real behavior of the basin.

CONCLUSIONS

The main objective of this work was to model an urban area experiencing extreme hydrological events, with a focus on investigating the effects of spatial discretization scale and rainfall

distribution variability on modelling results. The secondary objectives included comparing hydrological modelling results by varying the resolution of the land use and cover map, assessing the impact of different rainfall spatialization methods, and, finally, examining the influence of varying the number of rain gauges used.

Calibration was performed for three resolutions, employing the IDW method for spatialization during- an approximately 24-hour event. The calibration process involved adjusting the CN of the sub-basins and the roughness of the conduits within the limits specified by the SWMM Manual. It was observed that, in terms of volumes and shapes, the simulated curves for Event 1 (calibration) closely approximated the observed curves, demonstrating comparable adherence to those obtained by fully distributed models. However, when simulating the models generated for Events 2, 3, and 4 (validation), the obtained curves displayed limited adherence. This could be attributed to fixed calibration values (roughness and width of the sub-basins, for example), the gap in the urban drainage system data, and the decision to calibrate based on the maximum water level rather than the drained volumes.

The results confirmed that more discretized models exhibited lower errors and underscored the hydrological impacts on model outputs when changing the method of spatialization of pluviometric data and the number of rain gauges used. It should be noted that the modelling focused on the surface runoff on slopes and discharge into the hydrographic network due to data gaps related to the urban drainage network and cross sections of the rivers and channels of the Canal do Mangue Basin, influencing the obtained results. Furthermore, the study highlighted limitations such as the low discretization of the semi-distributed model compared to the fully distributed model, the assumption of null interception and evapotranspiration, and the use of events with similar station hyetographs, which contributed to smaller differences between the models.

For future studies, it is recommended to employ greater discretization in the models (e.g., average sub-basin areas of 0.20 km^2, 0.15 km^2, and 0.10 km^2), integrate the urban drainage network, include interception parcels, utilize less subjective infiltration models, select extreme rain events with greater variation in measured rainfall, and incorporate radar and/or satellite data in conjunction with rain gauges.

Despite the existence of a wide range of hydrological models, no standard model has been universally accepted due to the complexity and incomplete understanding of the urban hydrological system. While additional testing is necessary, including the implementation of more varied real rains, the insertion of the drainage system, and refinement of calibration, the methodology developed in this research has proven effective in aiding the selection of pluviometric data and spatialization methods, as well as in the elaboration of land use and land cover maps for hydrological models. Consequently, the study successfully investigated the effects of spatial discretization scale and rainfall distribution variability on the modelling results of an urban basin hydrological modelling.

ACKNOWLEDGEMENTS

This research was funded by Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), grant number SEI-260003/000537/2023.

DATA AVAILABILITY

Research data is only available upon request

REFERENCES

Bae, C., & Lee, D. K. (2020). Effects of low-impact development practices for flood events at the catchment scale in a highly developed urban area. *International Journal of Disaster Risk Reduction*, *44*, 101412.

Beven, K. (1991). Infiltration, soil moisture, and unsaturated flow. In D. S. Bowles, & P. E. O'Connell (Eds.), *Recent advances in the modelling of hydrologic systems* (pp. 137-151). Dordrecht: Springer.

Borges, E. C., Paz, I., Leite Neto, A. D., Willinger, B., Ichiba, A., Gires, A., Campos, P. C. O., Monier, L., Cardinal, H., & Amorim, J. C. C. (2021). Evaluation of the spatial variability of ecosystem services and natural capital: the urban land cover change impacts on carbon stocks. *International Journal of Sustainable Development and World Ecology*, *28*(4), 339-349. http://doi.org/10.1080/13504509.2020.1817810.

Brocca, L., Massari, C., Pellarin, T., Filippucci, P., Ciabatta, L., Camici, S., Kerr, Y. H., & Fernández-Prieto, D. (2020). River flow prediction in data scarce regions: soil moisture integrated satellite rainfall products outperform rain gauge observations in West Africa. *Scientific Reports*, *10*(1), 12517. http://doi.org/10.1038/s41598-020-69343-x.

Campos, P. C. D. O., & Paz, I. (2020). Spatial diagnosis of Rain Gauges' distribution and flood impacts: case study in Itaperuna, Rio de Janeiro – Brazil. *Water*, *12*(4), 1120. http://doi.org/10.3390/w12041120.

Cao, X., Lyu, H., Ni, G., Tian, F., Ma, Y., & Grimmond, C. (2020). Spatial scale effect of surface routing and its parameter upscaling for urban flood simulation using a grid-based model. *Water Resources Research*, *56*(2), e2019WR025468. http://doi.org/10.1029/2019WR025468.

Carissimi, J. (2011). Comunicação comparada: um estudo sobre a enchente Rio de Janeiro (2010) com base na análise das revistas Veja, Época e IstoÉ. *Revista Mediação*, *12*(3), 1-22.

Chen, T., Ren, L., Yuan, F., Yang, X., Jiang, S., Tang, T., Liu, Y., Zhao, C., & Zhang, L. (2017). Comparison of spatial interpolation schemes for rainfall data and application in hydrological modeling. *Water*, *9*(5), 342. http://doi.org/10.3390/w9050342.

Ferreira, F. P. M., & Cunha, S. B. (1996). Enchentes no Rio de Janeiro: efeitos da urbanização no Rio Grande (Arroio Fundo) - Jacarepaguá. *Anuário do Instituto de Geociências*, *19*, 79-92. http://doi.org/10.11137/1996_0_79-92.

Guzder-Williams, B., Mackres, E., Angel, S., Blei, A. M., & Lamson-Hall, P. (2023). Intra-urban land use maps for a global sample of cities from sentinel-2 satellite imagery and computer vision. *Computers, Environment and Urban Systems*, *100*, 101917. http://doi.org/10.1016/j.compenvurbsys.2022.101917.

Hu, C., Liu, C., Yao, Y., Wu, Q., Ma, B., & Jian, S. (2020). Evaluation of the impact of rainfall inputs on urban rainfall models: a systematic review. *Water*, *12*(9), 2484. http://doi.org/10.3390/w12092484.

Huang, Y., Bárdossy, A., & Zhang, K. (2019). Sensitivity of hydrological models to temporal and spatial resolutions of rainfall data. *Hydrology and Earth System Sciences*, *23*(6), 2647-2663. http://doi.org/10.5194/hess-23-2647-2019.

Instituto Brasileiro de Geografia e Estatística – IBGE. (2022). *Mapa de clima do Brasil*. Retrieved in 2023, May 22, from https://portaldemapas.ibge.gov.br

Kalakuntla, N., & Umamahesh, N. V. (2025). A comprehensive evaluation of calibration strategies for flood prediction in a large catchment using HEC-HMS. *Modeling Earth Systems and Environment*, *11*(1), 32. http://doi.org/10.1007/s40808-024-02276-w.

Landim, P. M. B. (2000). *Introdução aos métodos de estimação espacial para confecção de mapas* (18 p.). Rio Claro: UNESP.

Maia, A. C. N., & Sedrez, L. (2011). Narrativas de um dilúvio carioca: memória e natureza na grande enchente de 1966. *História Oral*, *14*(2), 221-254.

Mello, C., Lima, J., Silva, A., Mello, J., & Oliveira, M. (2003). Krigagem e inverso do quadrado da distância para interpolação dos parâmetros da equação de chuvas intensas. *Revista Brasileira de Ciência do Solo*, *27*, 925-933.

Niazi, M., Nietch, C., Maghrebi, M., Jackson, N., Bennett, B. R., Tryby, M., & Massoudieh, A. (2017). Storm water management model: performance review and gap analysis. *Journal of Sustainable Water in the Built Environment*, *3*(2), 04017002. http://doi.org/10.1061/JSWBAY.0000817.

Paula, S. C. D., Tassi, R., Piccilli, D. G. A., & Lorenzini Neto, F. (2018). Influence of the rain gauge network on the performance of a hydrological lumped model applied at different basin scales. *Revista Brasileira de Recursos Hídricos*, *23*, e45. http://doi.org/10.1590/2318-0331.231820180018.

Paz, A. R., & Collischonn, W. (2007). Rede de drenagem para modelagem hidrológica distribuída. In *Anais do XVII Simpósio Brasileiro de Recursos Hídricos*. Porto Alegre: Associação Brasileira de Recursos Hídricos.

Paz, I., Willinger, B., Gires, A., Alves de Souza, B., Monier, L., Cardinal, H., Tisserand, B., Tchiguirinskaia, I., & Schertzer, D. (2019). Small-scale rainfall variability impacts analyzed by fully-distributed model using C-band and X-band radar data. *Water*, *11*(6), 1273. http://doi.org/10.3390/w11061273.

Paz, I., Tchiguirinskaia, I., & Schertzer, D. (2020). Rain gauge networks' limitations and the implications to hydrological modelling highlighted with a X-band radar. *Journal of Hydrology*, *583*, 124615. http://doi.org/10.1016/j.jhydrol.2020.124615.

Perera, T., Piyadasa, R. U., Gunarathne, M., & Kumar, D. N. (2022). Hydrological modelling (soil water analysis tool) in water-related disasters and climate crisis: a review. In *"Youth" in the Forefront: Before and After World Water Forum. Online Youth Water Congress: "Emerging Water Challenges Since COVID-19*. Madrid: IAHR.

Rezende, O. M. (2018). *Análise quantitativa da resiliência a inundações para o planejamento urbano: caso da Bacia do Canal do Mangue do Rio de Janeiro* (Tese de doutorado). Universidade Federal do Rio de Janeiro, Rio de Janeiro.

Rio de Janeiro. Prefeitura. (2023). *Sistema Alerta Rio*. Retrieved in 2023, May 22, from https://www.sistema-alerta-rio.com.br/

Santos, F. M. (2018). *Modelagem concentrada e semi-distribuída para simulação de vazão, produção de sedimentos e de contaminantes em bacias hidrográficas do interior de São Paulo* (Tese de doutorado). Universidade de São Paulo, São Carlos.

Soil Conservation Service – SCS. (1956). Hydrology. In United States Department of Agriculture (Ed.), *National engineering handbook*. Washington, D.C.: USDA.

Shinma, T. A. (2015*). Avaliação de incertezas na calibração automática do modelo SWMM* (Tese de doutorado). Universidade de São Paulo, São Carlos.

Singh, A. (2018). A concise review on introduction to hydrological models. *Global Research and Development Journal for Engineering*, *3*(10), 14-19.

Sousa, M. M. d., Oliveira, A. K. B., Rezende, O. M., Magalhães, P. M. C., Jacob, A. C. P., Magalhães, P. C., & Miguez, M. G. (2022). Highlighting the role of the model user and physical interpretation in urban flooding simulation. *Journal of Hydroinformatics*, *24*(5), 976-991.

Souza, B. A., Paz, I. S. R., Ichiba, A., Willinger, B., Gires, A., Amorim, J. C. C., Reis, M. M., Tisserand, B., Tchiguirinskaia, I., & Schertzer, D. (2018). Multi-hydro hydrological modelling of a complex peri-urban catchment with storage basins comparing c-band and x-band radar rainfall data. *Hydrological Sciences Journal*, *63*(11), 1619-1635. http://doi.org/10.1080/02626667.2018.1520390.

Suri, A., & Azad, S. (2024). Optimal placement of rain gauge networks in complex terrains for monitoring extreme rainfall events: a review. *Theoretical and Applied Climatology*, *155*(4), 2511-2521. http://doi.org/10.1007/s00704-024-04856-3.

U.S. Environmental Protection Agency – EPA. (2022). *Storm Water Management Model (SWMM)*. Retrieved in 2022, December 5, from https://www.epa.gov/water-research/storm-water-management-model-swmm

Universidade Federal da Paraíba – UFPB. Laboratório de Eficiência Energética e Hidráulica em Saneamento – LENHS. (2012). *Manual do usuário do Storm Water Management Model - SWMM/EPA* (279 p.). João Pessoa: UFPB.

Workneh, H. T., Chen, X., Ma, Y., Bayable, E., & Dash, A. (2024). Comparison of IDW, Kriging and orographic based linear interpolations of rainfall in six rainfall regimes of Ethiopia. *Journal of Hydrology. Regional Studies*, *52*, 101696. http://doi.org/10.1016/j.ejrh.2024.101696.

Yang, W., Zhang, J., & Krebs, P. (2023). Investigating flood exposure induced socioeconomic risk and mitigation strategy under climate change and urbanization at a city scale. *Journal of Cleaner Production*, *387*, 135929. http://doi.org/10.1016/j.jclepro.2023.135929.

Authors contributions

James de Melo Sampaio Júnior: Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft, writing—review and editing, visualization.

Marcelo de Miranda Reis: Conceptualization, methodology, formal analysis, investigation, writing—original draft, visualization, supervision.

Igor da Silva Rocha Paz: Conceptualization, methodology, formal analysis, investigation, writing—original draft, writing—review and editing, visualization, supervision, funding acquisition.

Editor-in-Chief: Adilson Pinheiro

Associated Editor: Priscilla Macedo Moura