Evaluation of hydrological parameters of the Goiana River basin in the State of Pernambuco using the automatic calibration tool of the hydrodynamic model PCSWMM in multiple fluviometric stations

Avaliação dos parâmetros hidrológicos da bacia do Rio Goiana no Estado de Pernambuco utilizando ferramenta de calibração automática do modelo hidrodinâmico PCSWMM em múltiplas estações fluviométricas

Evaluación de parámetros hidrológicos de la cuenca del Río Goiana en el estado de Pernambuco utilizando la herramienta de calibración automática del modelo hidrodinámico PCSWMM en múltiples estaciones fluviométricas

Edinislon de Castro Ferreira
ORCID: https://orcid.org/0000-0003-2217-2359
Universidade Federal de Pernambuco, Brasil
E-mail:edinilson.castro@ufpe.br
Alexson Caetano da Silva
ORCID: https://orcid.org/0000-0001-8253-7378
Universidade Federal de Pernambuco, Brasil
E-mail:alexsoncaetano@hotmail.com
Jaime Joaquim da Silva Pereira Cabral
ORCID: https://orcid.org/0000-0002-1348-8004
Universidade Federal de Pernambuco, Brasil
E-mail: jaime.cabral@ufpe.br
José Roberto Gonçalves de Azevedo
ORCID: https://orcid.org/0000-0002-2662-7341
Universidade Federal de Pernambuco, Brasil
E-mail: jose.azevedo@ufpe.br

Abstract
The objective of this study was to perform a sensitivity analysis and automatic calibration of the hydrological parameters of the Goiana river basin in the state of Pernambuco. A hydrological modeling structure for the basin was built using the PCSMMM (Personal Computer Storm Water Management Model) hydrodynamic model and the analyzes were carried out in 10 sub-basins. Based on the model developed for the basin, the resources available in the PCSWMM for sensitivity analysis and automatic calibration, known as SRTC (Sensitivity-based Radio Tuning Calibration), were explored initially to evaluate the most sensitive parameters and then to perform the calibration of basin parameters. The daily flow data were extracted in 5 fluviometrics stations present in the main water courses of the basin and the precipitation data in 10 pluviometric stations. The results of the simulations were analyzed for the following statistics: Integral Square Error (ISE), Nash-Sutcliffe Efficiency (NSE), Determination Coefficient (R^2), in addition to the qualitative analysis of the hydrograph modeled in relation to the observations. The most sensitive parameters were the Curve Number (CN) and Manning's roughness for the permeable ones. In three control points (Engenho Retiro, Caricé and Nazaré) the SES values were 0.919, 0.899 and 0.884, respectively. The results, in general, indicate a very satisfactory performance of the tools of automatic calibration and sensitivity analysis of the model, besides being effective in relation to the traditional approaches of calibration by trial and error.
Keywords: Sensitivity analysis; Automatic calibration; PCSWMM model; Statistical error analysis.

Resumo
Objetivou-se com este estudo realizar a análise de sensibilidade e calibração automática dos parâmetros hidrológicos da bacia hidrográfica do rio Goiana no estado de Pernambuco. Uma estruturação de modelagem hidrológica da bacia foi construída no modelo hidrodinâmico PCSMMM (Personal Computer Storm Water Management Model) e as análises foram realizadas em 10 sub-bacias da mesma. A partir do modelo desenvolvido para a bacia, os recursos disponíveis no PCSWMM de análise de sensibilidade e calibração automática, conhecido como SRTC (Sensitivity-

based Radio Tuning Calibration), foram explorados inicialmente para avaliar os parâmetros mais sensíveis e em seguida para realizar a calibração dos parâmetros da bacia. Os dados de vazão diária foram extraídos em 5 estações fluviométricas presentes nos principais cursos d'água da bacia e os dados de precipitação em 10 postos pluviométricos. Os resultados das simulações foram analisados quanto as seguintes estatísticas: Erro Quadrado Integral (ISE), Eficiência Nash-Sutcliffe (NSE), Coeficiente de Determinação (R^2), além da análise qualitativa dos hidrograma modelados em relação as observações. Os parâmetros com maior sensibilidade foram o *Curve Number* (CN) e a rugosidade de *Manning* para as permeáveis. Em três pontos de controle (Engenho Retiro, Caricé e Nazaré) os valores de NSE foram de 0,919, 0,899 e 0,884, respectivamente. Os resultados, de maneira geral, indicam um desempenho bastante satisfatório das ferramentas de calibração automática e análise de sensibilidade do modelo, além de serem eficazes em relação as abordagens tradicionais de calibração por tentativa e erro.
Palavras-chave: Análise de sensibilidade; Calibração automática; Modelo PCSWMM; Análise estatística do erro.

Resumen
El objetivo de este estudio fue realizar un análisis de sensibilidad y calibración automática de los parámetros hidrológicos de la cuenca del río Goiana en el estado de Pernambuco. Se construyó una estructura de modelado hidrológico de la cuenca en el modelo hidrodinámico PCSMMM (Personal Computer Storm Water Management Model) y los análisis se realizaron en 10 subcuencas de la misma. Con base en el modelo desarrollado para la cuenca, los recursos disponibles en el PCSWMM para análisis de sensibilidad y calibración automática, conocidos como SRTC (Calibración de sintonización de radio basada en sensibilidad), fueron inicialmente explorados para evaluar los parámetros más sensibles y luego realizar la calibración de parámetros de la cuenca. Los datos de caudal diario se extrajeron de 5 estaciones fluviométricas presentes en los principales cursos de agua de la cuenca y los datos de precipitación de 10 estaciones pluviométricas. Los resultados de la simulación se analizaron utilizando las siguientes estadísticas: Error Cuadrado Integral (ISE), Eficiencia Nash-Sutcliffe (NSE), Coeficiente de Determinación (R^2), además del análisis cualitativo del hidrograma modelado en relación a las observaciones. Los parámetros con mayor sensibilidad fueron el Número de Curva (CN) y la rugosidad de Manning para las permeables. En tres puntos de control (Engenho Retiro, Caricé y Nazaré) los valores de NSE fueron 0,919, 0,899 y 0,884, respectivamente. Los resultados, en general, indican un desempeño muy satisfactorio de las herramientas de calibración automática y análisis de sensibilidad del modelo, además de ser efectivas en relación a los enfoques tradicionales de calibración de prueba y error.
Palabras clave: Análisis de Sensibilidad; Calibración automática; Modelo PCSWMM; Análisis estadístico de errores.

1. Introduction

Conceptual runoff models, whether in urban or rural areas, require a large number of variables and parameters to adequately describe the complex relationships between rainfall, runoff and river basin characteristics. The difficulties in estimating these parameters often end up limiting the use of such models. Therefore, the performance of the models in predicting the hydrograph is highly dependent on the accuracy of its calibration (Loucks et al., 2005).

Due to the presence of several optimal solutions and the large number of parameters considered, the trial-and-error calibration approach often used is tedious and time-consuming. It is therefore necessary to develop robust and reliable automatic calibration procedures to obtain the ideal parameters for the model (Javaheri, 1998).

Traditional calibration methods are applied to create model results that are in accordance with the measured outputs, the agreement is reached by the appropriate choice of parameters. To improve the accuracy of the calibration and the reliability of the models, as well as streamline the calibration process, an automatic calibration method becomes desirable (James, 2005).

Automatic calibration algorithms coupled with a hydrodynamic flow model have been increasingly used by researchers in order to improve the accuracy and efficiency of parameter adjustment processes. Formiga et al. (2016) made modifications to the structure of the SWMM model in order to allow the coupling of the R-NSGA algorithm (Evolucionary Reference Point Based Non-Dominated Sorting Genetic Algorithm) to perform automatic calibration of model parameters. James, Wan e James (2002) implemented in PCSWMM a genetic algorithm for automatic calibration and design optimization. Gou et al. (2020) used an automatic calibration structure which combines a sensitivity analysis (SA) adapted to an optimization algorithm for calibrating a hydrological model. And Wang et al. (2020) developed a routine for automatic parameter calibration of the SWAT model based on a genetic algorithm (GA) and particle swarm optimization (PSO).

Typically, automatic calibration methods based on a sensitivity analysis involve four main steps: (1) sensitivity analysis; (2) determination of the limit or range of parameters (uncertainty analysis); (3) calibration and (4) performance evaluation of the calibrated model (James et al, 2002)

The objective of the sensitivity analyzes is to prepare the calibration of the basin parameter sets in order to illustrate the impact of the uncertainties of the input parameters on the model responses (Garambois et al., 2015). Therefore, the intention is to verify the rate of change in the response of a model in relation to changes in the input parameters of the same model (James, 2005). After the sensitivity analysis, it is necessary to determine the acceptable ranges of the parameters that will be adjusted in the model, also called the uncertainty analysis. In this step, an attempt is made to describe the entire set of possible results of the model parameters, together with their associated probabilities of occurrence, or simply to assign limits of physically acceptable minimum and maximum values (Loucks; Beek, 2017).

The performance analysis of the calibrated models, involves qualitative evaluation through visual comparisons between the time series of simulations and observations to identify how well the model captures the time, trends and magnitudes of the observations. And quantitative assessments, which assess the level of agreement between simulations and observations using statistics that measure the quality of the fit, which are used as criteria for accepting or rejecting the calibration results (Ahmadisharaf et al., 2019).

Complementing the analysis of the model's performance evaluations, Daggupati et al. (2015) argued that the use of quantitative statistics alone can be misleading and unreliable. Graphical comparisons together with quantitative statistics result in a better assessment criterion.

The present study uses the automatic calibration resources of the PCSWMM hydrodynamic model, to estimate the hydrological parameters of the Goiana river basin in the state of Pernambuco, using five flow observation points distributed in three of the main rivers in the basin. The performance of the model's response is analyzed by means of statistical measures of fit quality, such as the Nash-Sutcliffe Efficiency (NSE), the determination coefficient (R^2) and the Integral Square Error (ISE). In addition to the qualitative evaluations of the simulated hydrographs in relation to the observations, analyzing the maximum flows and the trends.

2. Materials and Methods

The methodological strategy to achieve the objectives of the study, involves the following steps: (1) Construction of the hydrological model in the PCSWMM; (2) Sensitivity and uncertainty analysis of the model parameters; (3) Automatic calibration through multiple fluviometric stations and sub-basins; (4) Performance analysis, quantitative and qualitative, of the model. Figure 1 shows the flowchart of the research steps.

**Figure 1.** Flowchart of the study steps.
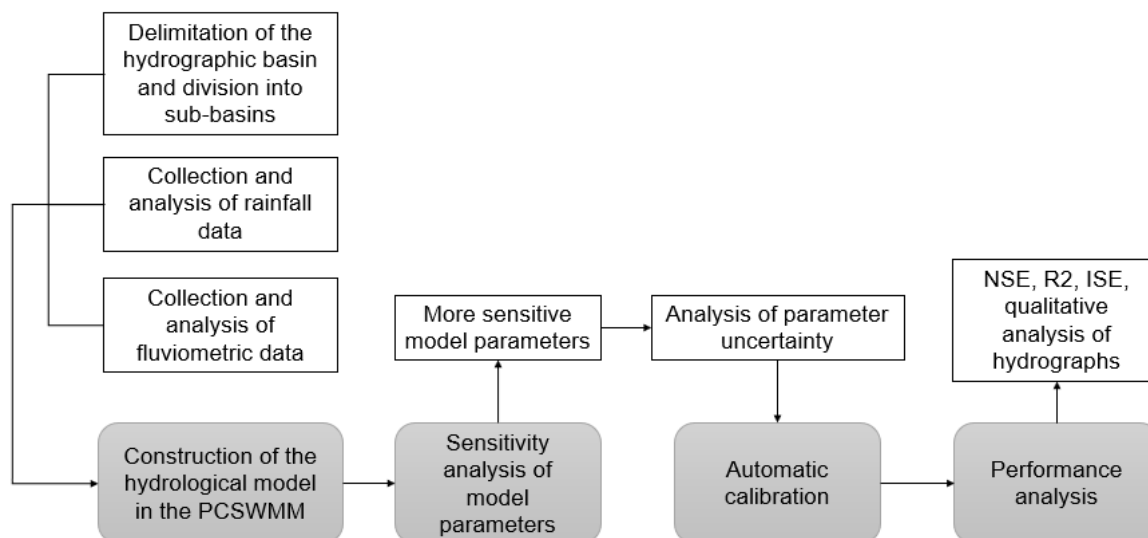
Source: Authors (2021).

2.1 Study area

Study area corresponds to the hydrographic basin of the Goiana River, located in the eastern northern portion of the State of Pernambuco, between the coordinates 07º22'20” and 07º54'47” south latitude, and 34º49'06” and 35º41'43” west longitude. The basin is limited to the north with the state of Paraíba and to the south with the Capibaribe River basin (Figure 2).

The Goiana River basin has a total area of 2,847.53 km², corresponding to 2.90% of the total area of the state. The main water courses in the basin are the Capibaribe-Mirim, Siriji, Tracunhaém and Goiana rivers. The Capibaribe-Mirim River is approximately 93 km long and drains most of the basin. The Siriji River and Tracunhaém have an approximate length of 74 km and 127 km, respectively. The river named after the basin, the Goiana River, arises from the confluence of the Tracunhaém and Capibaribe-Mirim rivers, with an extension of approximately 18 km to the mouth of the Atlantic Ocean.

Figure 2. Map of the location of the hydrographic basin of the Goiana River.
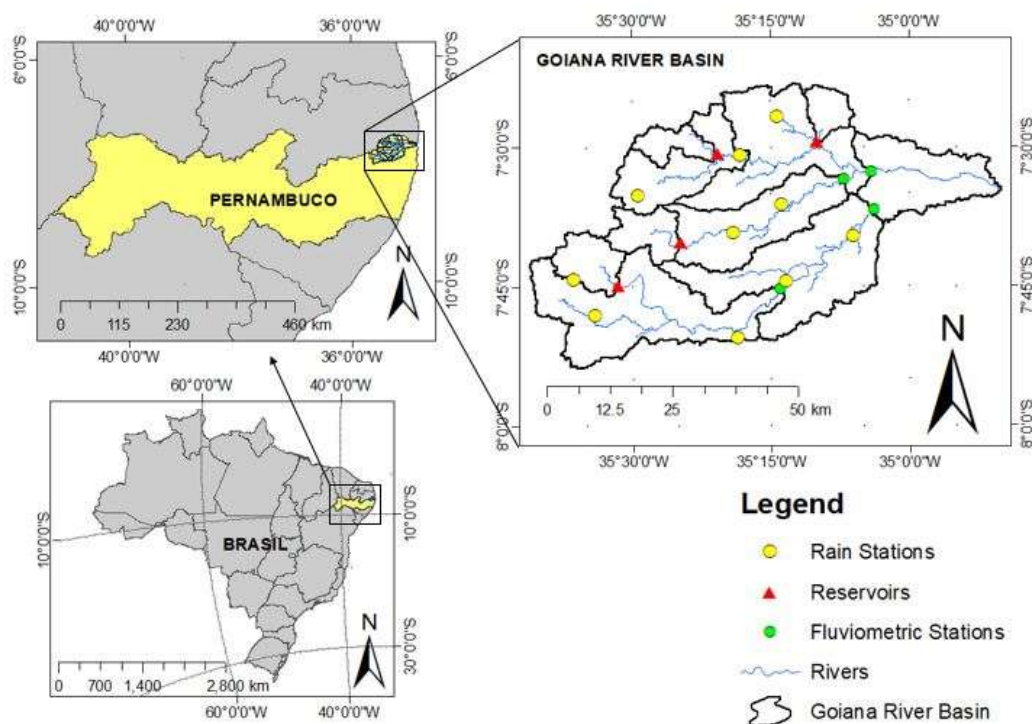
Source: Authors (2021).

The medium and lower stretches of the Goiana River basin are located in the Agreste region of the state. This region is considered to be intermediate between the humid climate (Zona da Mata) and dry climate (Sertão) areas, presenting characteristics of one, now of the other. Thus, in the areas closest to the Sertão the climate is hot and dry, and the rainiest period runs from February to June. In the middle section, the rainy season extends from March to July (CPRH; MMA, 2011).

The lower part of the basin has characteristics of hot and humid climate, with rainfall averages above 1.000 mm per year, reaching more than 2.000 mm in coastal areas. The rainy season lasts six months, from March to August (CONDEPE/FIDEM, 2005).

2.2 Pluviometry and Fluviometry of the region

In this study, the Goiana River basin was subdivided into 10 sub-basins using the following criteria for division of 9 of these: sub-basins with exudatory at the fluviometric stations (control points) and sub-basins with exutory in the basin's reservoirs. The last sub-basin, at the mouth, is the sub-basin corresponding to the Goiana River, which does not have a fluviometric station.

In total, the basin has 5 fluviometric stations all operated by ANA (National Water Agency) and 4 reservoirs, as shown in Tables 1 and 2. For the purposes of this study, the sub-basins were named after the fluviometric stations or reservoir in their exutory (Figure 3).

Figure 3. Division and name of sub-basins.

Source: Authors (2021).

Table 1. Fluviometric stations and corresponding basin.

Fluviometric Station	Latitude	Longitude	Drained area (km²)	River
Engenho Volta	-7,5219	-35,2883	250,17	Capibaribe-Mirim
Caricé	-7,5432	-35,0689	450,16	Capibaribe-Mirim
Engenho Retiro	-7,5578	-35,1175	359,24	Siriji
Itapissirica	-7,3161	-35,0647	542,70	Tracunhaém
Nazaré da Mata	-7,7542	-35,2331	570,52	Tracunhaém

Source: Authors (2021).

Table 2. Reservoirs and corresponding basin

Reservoirs	Latitude	Longitude	Drained area (km²)	River
Guararema	-7,4917	-35,1664	69,54	Capibaribe-Mirim
Tiúma	-7,5139	-35,3383	79,34	Capibaribe-Mirim
Siriji	-7,6389	-35,2839	113,68	Siriji
Palmeirinha / Pedra Fina	-7,7483	-35,5242	121,59	Tracunhaém

Source: Authors (2021).

For calibration, daily flow data were extracted in the period from 01 to 30 June of the year 2013 and for validation, the period between 01 to 30 September of the following year, 2014. These periods were chosen because they present data consistency, the absence failure and because it encompasses periods where intense rains have occurred in the region.

Daily precipitation data were obtained from 10 pluviometric stations (Table 3), all operated by ANA, with 2 stations within the limits of the Engenho Volta, Engenho Retiro and Itapissirica sub-basins, 3 stations within the Nazaré sub-basin and 1 station in the Caricé sub-basin.

The PCSWMM model associates only a series of precipitation data for each sub-basin. Therefore, the interpolation method weighted by the inverse of the squared distance was used to obtain rainfall in each sub-basin. This method considers

that the rain in a place (centroid of the sub-basins), can be calculated as a weighted average of the rain in pluviometers of the region. The weighting is done in such a way that the nearest pluviometric stations have a greater weight in the calculation of the average, according to equation 1 (Collischonn; Dornelles, 2013).

$$Pm_i = \frac{\sum_{j=1}^{NP} \frac{P_j}{d_{ij}^2}}{\sum_{j=1}^{NP} \frac{1}{d_{ij}^2}} \qquad (1)$$

Where NP is the number of rainfall stations with available data; P_j is the rain observed at station j; and d_{ij} is the distance between the centroid of sub-basin i and the pluviometric post j.

Table 3. Rain Stations in the Goiana River Basin.

Rain Stands	Latitude	Longitude
Nazaré da Mata	-7,7408	-35,2228
Vicência	-7,6536	-35,3183
Aliança	-7,6042	-35,2306
Ferreiros	-7,4461	-35,2386
Timbaúba	-7,5144	-35,3058
Lagoa do Carro	-7,8422	-35,3114
Bom Jardim	-7,8017	-35,5678
Orobó	-7,7372	-35,6069
Itaquitinga	7,6597	-35,1014
São Vicente Ferrer	-7,5875	-35,4889

Source: Authors (2021).

2.3 Hydrodynamic model PCSWMM

The PCSWMM (Personal Computer Storm Water Management Model) rainwater management model, based on the SWMM (version 5) of USEPA (United States Environmental Protection Agency), integrated with a geographic information system (GIS), is the model used in the present study. In this model, each sub-basin is represented as a rectangular surface that has a uniform slope (S) and a flow width (W) that converges the flow to a single point in the outlet channel. Runoff is generated by modeling the sub-basin as a non-linear reservoir (Rossman, 2015).

The main input data for the model are: the sub-basin area (ha), average slope (%), runoff width (m), percentage of impermeable area (%), Manning roughness for impermeable and permeable areas and depth of storage in the depressions of impermeable and permeable areas (mm). And the infiltration model, which was chosen the Curve Number (CN).

The SWMM model is traditionally used in basin modeling in small and medium-sized urban areas. Usually as a tool for managing and detecting flooding points. With the insertion of new features, such as a GIS environment and tools for sensitivity analysis and automatic calibration, in the PCSWMM version, motivated the use in other types of water systems. In this study the model will be tested for its performance, in a basin with rural predominance and large dimensions.

Munir, Ahmad and Hafeez (2019) proved the good performance of the PCSWMM model in large rural basins, using the model's automatic calibration tools, found Nash-Sutcliffe efficiency values (NSE) and determination coefficient, between 0,75-0,97 and 0,94-0,98, respectively. As well as Ghofrani, Sposito and Faggian (2019), who also used the PCSWMM model in a predominantly rural basin.

2.4 Sensitivity analysis and calibration

The PCSWMM uses the tool known as SRTC (*Sensitivity-based Radio Tuning Calibration*), calibration of radio tuning based on sensitivity, to perform the calibration of hydrological models in relation to the observed data or to test sensitivity attributes and parameters. The SRTC tool requires estimates of uncertainties for the attributes and parameters to be calibrated. SWMM (version 5) is executed using the limit values of the sensitivity range specified by the user. The PCSWMM provides the user with the option of considering intermediate values within the limit specified for each parameter in the sensitivity graphs, making it possible to check whether there is linearity or not in the sensitivity of a parameter in relation to its response (CHI WATER, 2021).

Therefore, the first step in calibrating the model was to select the most sensitive parameters and estimate a range of uncertainty for them. This uncertainty indicates the limits in which the user feels comfortable with the adjustment of the parameter, being expressed as a percentage deviation (more or less) from its current value (best estimate).

The values of the lower and upper parameters for a specified uncertainty are calculated by the PCSWW according to equations 2 and 3 (Chi Water, 2021):

Lower parameter values:

$$P_{inf} = P_{atual} \times \left(\frac{1}{1 + \frac{U}{100}} \right) \quad (2)$$

Higher parameter values:

$$P_{sup} = P_{atual} \times \left(1 + \frac{U}{100} \right) \quad (3)$$

Where P_{inf} is the parameter value for the lower limit of the uncertainty range; P_{sup} is the parameter value for the upper limit of the uncertainty range; P_{atual} is the current value of the estimated parameter; and $U(\%)$ is the uncertainty considered.

The strategy used to calibrate the model was to carry out the joint analysis of the sub-basins separated by the corresponding rivers, that is, the Engenho Volta, Caricé, Guararema and Tiúma sub-basins, for example, drain their waters into the Capibaribe River -Mirim, therefore, the calibration of the parameters of these sub-basins was analyzed together, considering the two control points existing in this river. The same procedure was also used for the other rivers, the Siriji river with the Engenho Retiro and Siriji sub-basins, and the Tracunhaém river with the Nazaré, Palmeirinha and Itapissirica sub-basins.

It is important that the calibration methods do not allow the parameters to exceed physically significant maximum and minimum limits. Carefully determining the limits of each calibration parameter increases the accuracy and efficiency of the process (James, 2005). The first estimates of the model parameters should be as close as possible to the calibrated value. Thus, a realistic, physically significant range of model parameters must be identified.

The parameters analyzed in this study were:

- Area (*A*) and slope (*I*) of the sub-basins – These parameters were obtained with GIS technology, but using a Digital Terrain Model (MDT) with a spatial resolution of 30 meters. Therefore, they are subject to errors, even if minimal. James (2005), suggests a limit of ±2% in relation to the data measured for the area and of -12% to 29% of the data measured for the slope. ±5% for area and ±30% for slope was used, which are existing and selectable values within the model's resources.

- Runoff width (W) – To systematically estimate the width of the sub-basin, the measured area of the sub-basin is divided by the length of the longest flow path. An uncertainty range of ±60% was assigned to this parameter, given the possibilities of error both in the area measurement and in the length of the largest watercourse.

- Percentage of impermeable area (%) – In relation to the impermeable areas of the basin, an initial value of 10% was assigned to all sub-basins with an uncertainty range of ±100%. All basins are predominantly rural, with small urban areas in percentage terms, no more than 20% of the sub-basin areas.

- Manning roughness coefficient of impermeable and permeable areas (n_i e n_p) – The initial values of this parameter were established according to the literature recommended by the PCSWMM. Due to the high variability of the Manning roughness coefficient for different types of soil cover, the limit and range of this parameter are very difficult to estimate. Therefore, a limit of ± 100% was adopted in relation to the initial value initially established, 0,030 for permeable areas and 0,015 for impermeable areas (Rossman, 2015).

- Depth of storage in the depressions of impermeable and permeable areas (D_i e D_p) - It is defined as the portion of water that is stored in depressions and exhausted only by evaporation. Its value varies according to the type of soil and the slope of the basin. The PCSWMM user manual recommends for impermeable areas a limit of 0,13 mm to 1,50 mm and for permeable areas 1,5 mm to 6,5 mm (Rossman, 2015). An uncertainty range of ±100% was also assigned and the initial value was the median of the values presented above.

- Curve Number (CN) – The initial value of the CN was attributed based on official data from the National Water Agency (ANA) on its "Metadata" portal. The range of uncertainty used was ±100%. For this parameter, the physically acceptable limits are from 0 to 100, where the model itself restricts its value in this range.

The observed time series files can be opened in the SRTC tool (after executing the sensitivity) or in the Graph panel (before starting the SRTC tool), in which case periods or ranges of values can be selected which will perform the calibration, decreasing thus computational time (Chi, Water, 2021).

In the SRTC calibration window, the radio tuning sliders (Figure 4) are displayed for each parameter, where it is possible to view the calculated effect of the adjustment automatically.

Figure 4. Radio tuning slider feature.
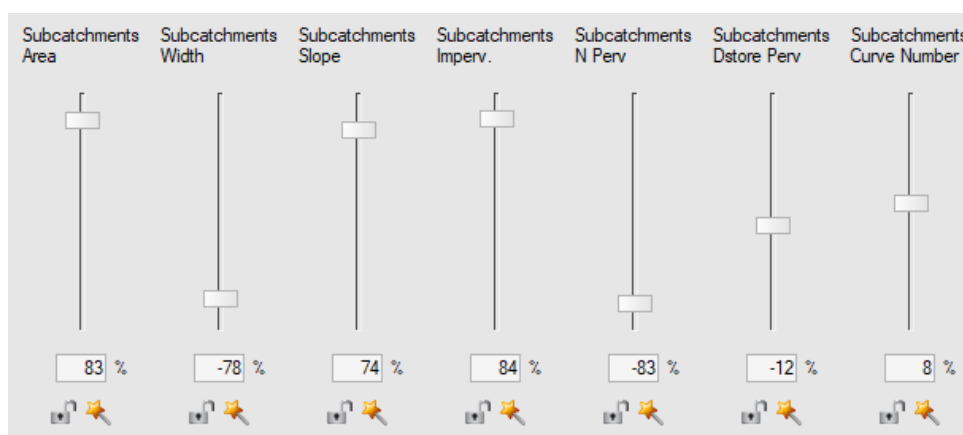
Source: PCSWMM (2021).

The "*Optimize*" button performs the automatic calibration of each parameter, which is optimized based on the ISE (integral square error) or NSE (Nash-Sutcliffe efficiency) adjustment measures.

2.5 Statistical error analysis

The SRTC calibration tool, calculates various error statistics between time series with the same units in real time. Statistics include, the classification of the integral square error, the value of the integral square error (ISE), Nash-Sutcliffe efficiency (NSE), coefficient of determination (R²), standard error of estimate (SEE), simple least squares (LSE), dimensionless simple least squares (LSE dim), root of the mean square error (RMSE) and root of the mean square error without dimension (dim RMSE).

In this study we will discuss and analyze the quantitative adjustments of the Integral Square Error (ISE), the Nash-Sutcliffe efficiency (NSE) and the determination coefficient (R²), in addition to the qualitative adjustment of the hydrographs observed and calculated by the model.

Integral Square Error (ISE)

The Integral Square error is a measure of adequacy between the observed and modeled responses. ISE integrates the square of the error over time. A feature of this index is that it provides large weights for large errors and small weights for small errors. The ISE criterion is defined mathematically according to equations 4 or 5 (Gujarati; Porter, 2009).

$$ISE = \int_0^T e^2(t)dt \qquad (4)$$

$$ISE = \frac{\left[\sum_{i=1}^N \left(Q_o(i) - Q_c(i)\right)^2\right]^{\frac{1}{2}}}{\sum_{i=1}^N Q_o(i)} \qquad (5)$$

Where $e(t)$ is the residual error; $Q_o(i)$ is the observed value for the i-th observation; $Q_c(i)$ is the computed value for observation i; and N the number of values observed or calculated. The same notations will be used in the equations presented below.

The lower the ISE value, the better the model matches the observed values. The ISE value is assessed according to the following classification (Sarma; Delleur; Rao, 1973):

$ISE < 3,0$ "Great".

$3,0 < ISE < 6,0$ "Very good".

$6,0 < ISE < 10,0$ "Well".

$10,0 < ISE < 25,0$ "Reasonable".

$ISE > 25$ "Poor"

Nash-Sutcliffe efficiency (NSE)

Nash-Sutcliffe efficiency (NSE) is a standardized statistic that determines the relative magnitude of the residual variance ("noise") compared to the variance of the measured data ("signal") (Nash; Sutcliffe, 1970). The equation is as follows:

$$NSE = 1 - \frac{\sum\left(Q_o(i) - Q_c(i)\right)^2}{\sum(Q_o(i) - \overline{Q_o})^2} \qquad (6)$$

$\overline{Q_o}$ is the average of the observed data.

The Nash-Sutcliffe efficiency values can vary from $-\infty$ to 1. Essentially, the closer to 1, the more accurate the model is. Moriasi et al (2015) proposed the following classification for the NSE.

$0,75 < NSE \leq 1$ "Very good".

$0,65 < NSE \leq 0,75$ "Well".

$0,50 < NSE \leq 0,65$ "Satisfactory".

$NSE \leq 0,50$ "Unsatisfactory".

Determination coefficient (R^2)

The coefficient of determination R² is a summary measure that tells how well the sample's regression line fits the data (Gujarati; Porter, 2009). It can be interpreted as the proportion of variance in the dependent variable that is predictable from the independent variable. An R² equal to 1 indicates that the regression line fits the data perfectly. Your equation is presented below.

$$R^2 = \left(\frac{\sum (Q_o(i) - \overline{Q_o})(Q_c(i) - \overline{Q_c})}{\sqrt{\sum (Q_o(i) - \overline{Q_o})^2 \sum (Q_c(i) - \overline{Q_c})^2}} \right)^2 \quad (7)$$

$\overline{Q_c}$ is the average of the calculated values.

Moriasi et al (2015), also proposed a classification for the R², being it:

$R^2 > 0,85$ "Very good".

$0,70 \leq R^2 \leq 0,85$ "Well".

$0,50 < R^2 < 0,70$ "satisfactory".

$R^2 \leq 0,50$ "Unsatisfactory".

3. Results and Discussion

The parameters of the PCSWMM model were analyzed for their sensitivity in relation to the peak flow in the sub-basins. The depth of storage in the depressions of the impermeable and permeable areas (D_i and D_p), and the Manning roughness of the impermeable areas (n_i) did not present any sensitivity to the model. Insensitivity to these parameters is mainly induced by the low proportion of impermeable areas and due to the predefined value limits for these, which are relatively narrow. Similar results of insensitivity of these parameters in the SWMM model were found by Krebs et al. (2013).

The CN (Curve Number) and the Manning roughness of the permeable areas (n_p) in the sub-basins, were the parameters with the highest sensitivity, followed by the runoff width (w) and the area (A) of the sub-basins, all analyzed within the specified uncertainty range. Formiga et al. (2016), performed sensitivity analysis of parameters of the SWMM model followed by automatic calibration using an R-NSGA optimization algorithm coupled to the model, the most sensitive parameters were also the runoff width (w) and the area (A), in addition to the slope (I) and the Manning roughness of the impermeable areas (n_i).

The CN even changes the maximum flow by up to 65%. Its minimum value, after adjustment, was found in the Itapissirica sub-basin, in the southern region of the Goiana river basin, with a value of 17.14, corroborating with some results found by Junior, Rodrigues and Oliveira (2019) for the CN after the calibration process. Still on the CN, its highest value was 54.51, in the Guararema sub-basin in the northern region of the basin, already on the border as the State of Paraíba.

Analyzing the CN together with the sub-basins of each river, there is a proximity in its values. The Caric, Engenho Volta, Guararema and Tiúma sub-basins that drain their waters into the Cabiparibe-Mirim River, located in the northernmost region of the basin, had an average CN value of 47,44. The Engenho Retiro and Siriji sub-basins, which contribute to the Siriji River, the central region of the basin, had an average CN of 23,10. And in the southernmost region of the basin, where the Tracunhaém river is located, which has the Itapissirica, Nazaré and Palmeirinha sub-basins, we obtained an average CN of

17,62. The values of all parameters, initial and calibrated, are shown in Table 07 and the sensitivity graphs in relation to the peak flow in Figures 05 to 09.

Regarding the quantitative adjustment measures, the Engenho Retiro, Caricé and Nazaré sub-basins, all with control points, presented the best values in relation to Nash-Sutcliffe Efficiency (NSE), with the respective values of 0,919, 0,899 and 0,884, these values have a performance classified as "very good" for Moriasi et al. (2015). The same sub-basins, in relation to the coefficient of determination R^2, also presented results considered "very good", with the following values: 0,919 for Engenho Retiro, 0,918 for Caricé and 0,89 for Nazaré.

In the validation process, the Engenho Retiro and Caricé sub-basins obtained an NSE of 0,649 and 0,711, respectively, results considered "satisfactory" and "good", respectively. The Nazaré sub-basin had an NSE of -5,03, therefore "unsatisfactory". Using a similar strategy, to perform the calibration locally by sub-basins and then do a global analysis through an optimization algorithm for automatic calibration, Gao et al. (2020) analyzed 57 sub-basins and in only one of them obtained a result considered unsatisfactory in terms of Nash Efficiency, with a negative value for it. In the validation analyzes, the R^2 maintained a satisfactory performance in the Engenho Retiro and Caricé sub-basins, and an unsatisfactory result for Nazaré, presenting an R^2 of 0,49.

The results presented in the Engenho Retiro and Caricé sub-basins, regarding the NSE and R^2, are consistent with the study prepared by Wang et al. (2020), which developed an automatic calibration model with genetic algorithm (GA) also applied in large basins.

The other sub-basins, Engenho Volta and Itapissirica, presented results, as for the NSE, "satisfactory" and "unsatisfactory", respectively. Regarding validation, the Itapissirica sub-basin obtained an NSE of 0,812, that is, better than the calibrated result. The Engenho Volta sub-basin had a "bad" result with the NSE value less than zero. The coefficient of determination for the same basins, were all above the limit considered satisfactory, both in calibration and in validation.

The behavior of the model when analyzed by the Integral Square Error (ISE) was considered good for the Engenho Retiro, Caricé and Nazaré sub-basins, the same ones that obtained good results regarding the NSE and the R^2, in the calibration process. Still in calibration, the Itapissirica sub-basin presented a "reasonable" result and only the Engenho Volta sub-basin obtained a "bad" result for this index. All NSE and R^2 results are presented graphically in Figures 10 and 11 and in table 05. The validation being analyzed by this index, obtained a "good" result only in one sub-basin, Itapissirica, in the others the results were considered from reasonable to bad. This measure was also analyzed by Jaiswal et al. (2014), evaluating the performance of the automatic calibration and validation of 41 sub-basins, obtaining results at a level considered "Great" for this index in all sub-basins.

The correspondence between the hydrographs observed and calculated by the PSWMM model, can be seen in Figures 12 to 16, and in the validations of Figures 17 to 21. It is possible to notice that the calibrated flows, in general, are better adjusted to the observations (the model fits best in both the maximum and minimum points). The performance of the model for the series of validated flows is, in general terms, considered satisfactory, similar performance was found by Gao et al. (2020). The best correspondence between peak flows is observed in the Nazaré sub-basin with a variation of only 5,11%, after the calibration process. Regarding the volume of total runoff, the Caricé sub-basin showed a 1,83% variation in relation to the same observed data. In all sub-basins, after calibration, the variation between the observed and calculated peak flows did not exceed 25%. The total volume of runoff, in 4 of the 5 controlled sub-basins, the variation did not exceed 6%. The other results regarding these criteria are presented in Tables 05 and 06.

The results, after the calibrations, regarding the variations of the maximum flows and total flow volume, in 90% of the cases were within the limits recommended by James (2005). However, the author also argues that the calibration tolerances need to consider the inaccuracies inherent in the observed data and the objectives proposed in the hydrological modeling.
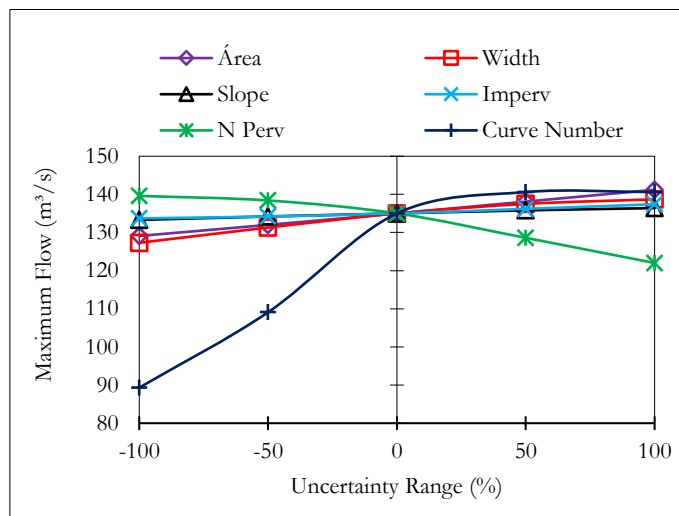
Figure 5. Sensitivity Analysis (Carice Sub-basin).
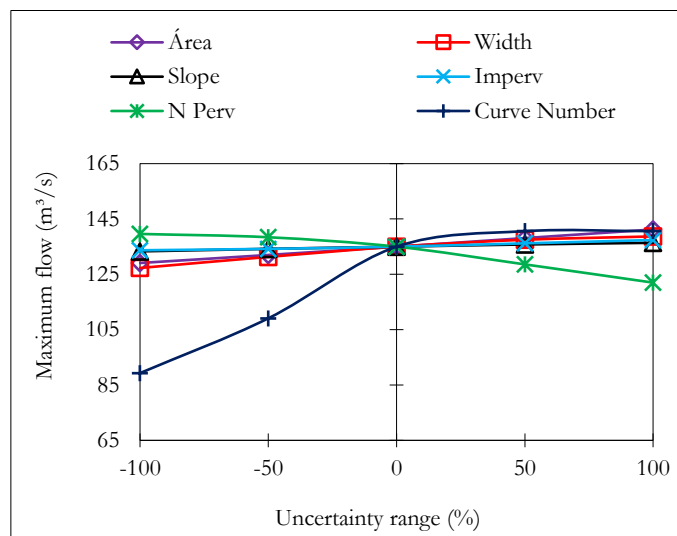
Source: Authors.

Figure 6. Sensitivity Analysis (Engenho Retiro Sub-basin).
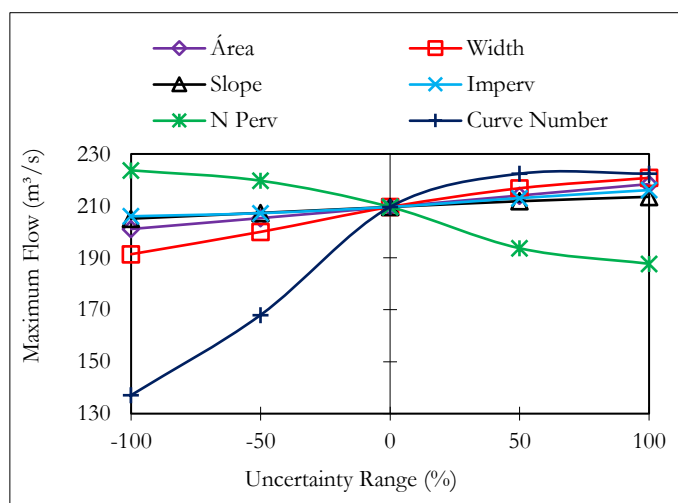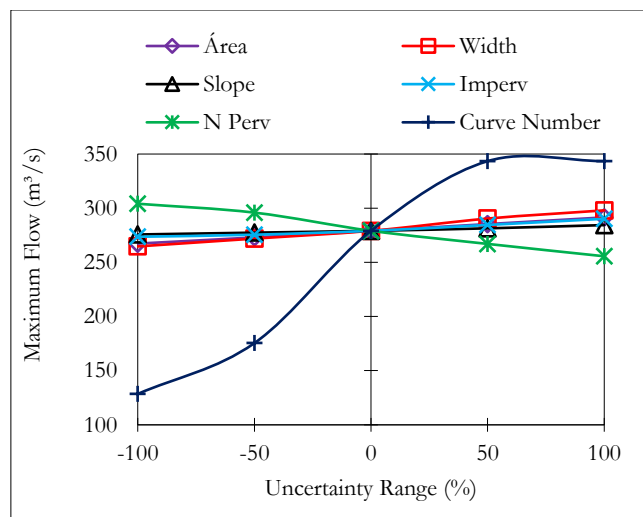

Source: Authors.

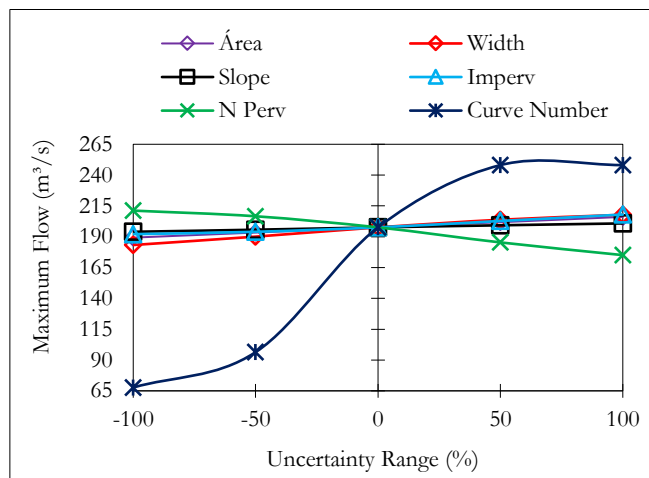**Figure 7.** Sensitivity Analysis (Engenho Volta Sub-basin).

Source: Authors.

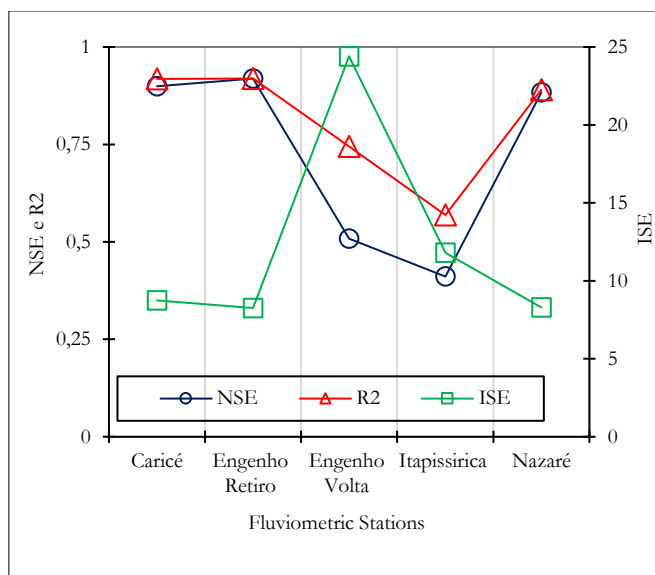**Figure 8.** Sensitivity Analysis (Engenho Itapissirica Sub-basin).

Source: Authors.
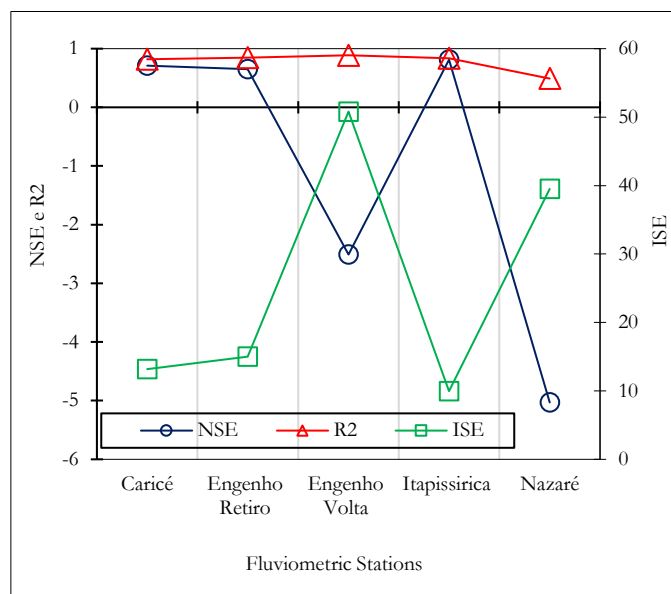
Figure 9. Sensitivity Analysis (Nazaré Sub-basin).

Source: Authors.

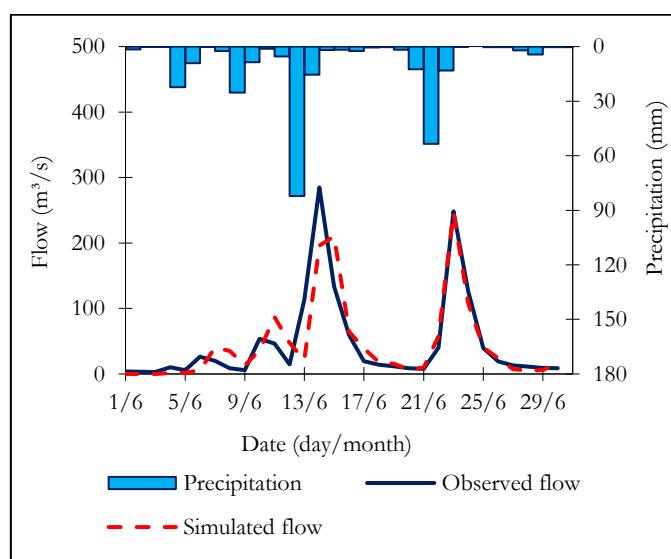**Figure 10.** NSE, R2 and ISE results (Calibration).

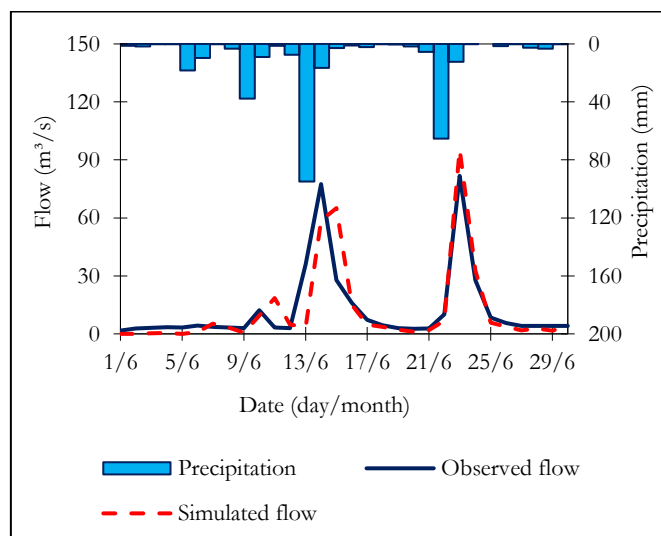Source: Authors.

Figure 11. Resultados NSE, R2 e ISE (Validação).

Source: Authors.

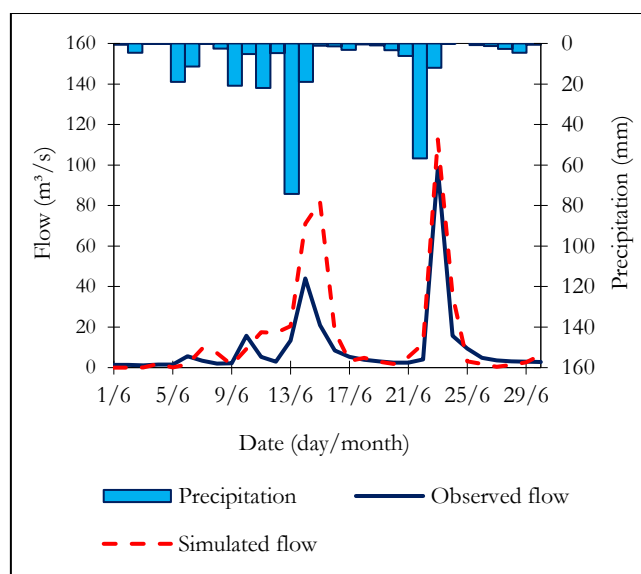**Figure 12.** Observed and calibrated hydrograph (Carice sub-basin).

Source: Authors.

16

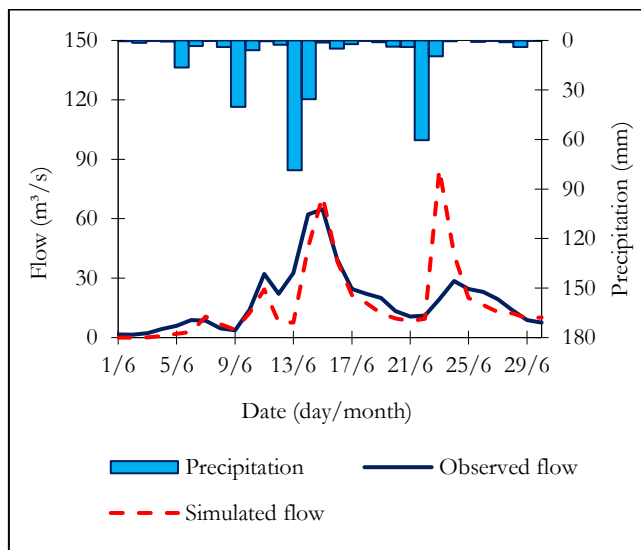**Figure 13.** Observed and calibrated hydrograph (Engenho Retiro sub-basin).

Source: Authors.

Figure 14. Observed and calibrated hydrograph (Engenho Volta sub-basin).
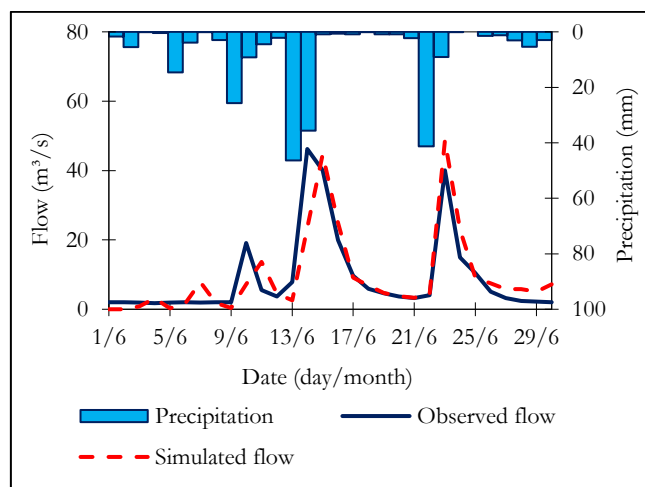

Source: Authors.

Figure 15. Observed and calibrated hydrograph (Engenho Itapissirica sub-basin).

Source: Authors.

Figure 16. Observed and calibrated hydrograph (Nazaré sub-basin).

Source: Authors.

Figure 17. Observed and validated hydrograph (Caricé sub-basin).

Source: Authors.

Figure 18. Observed and validated hydrograph (Engenho Retiro sub-basin).

Source: Authors.

Figure 19. Observed and validated hydrograph (Engenho Volta sub-basin).

Source: Authors.

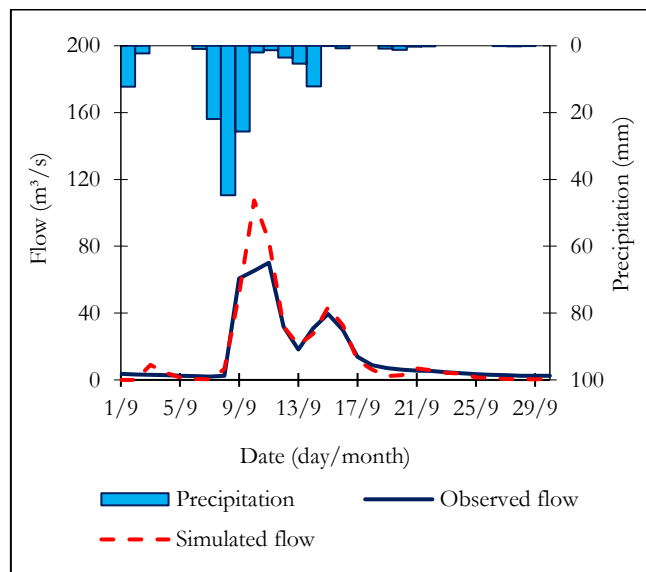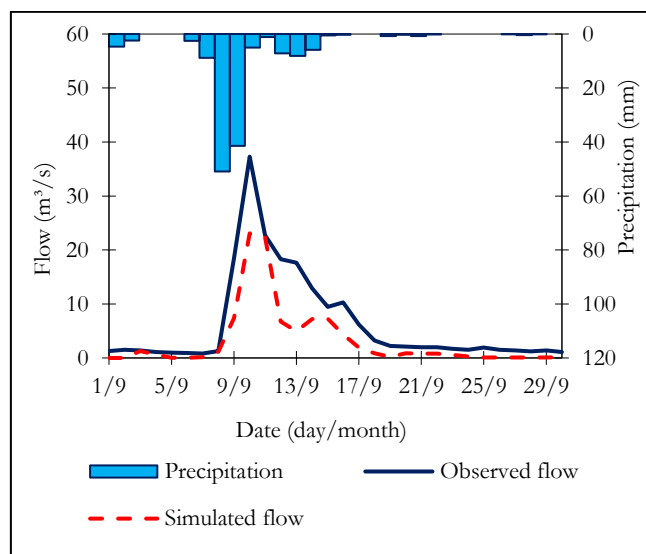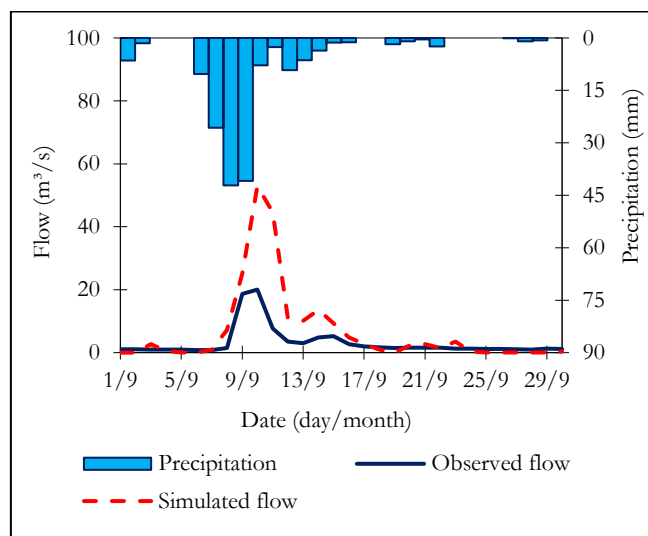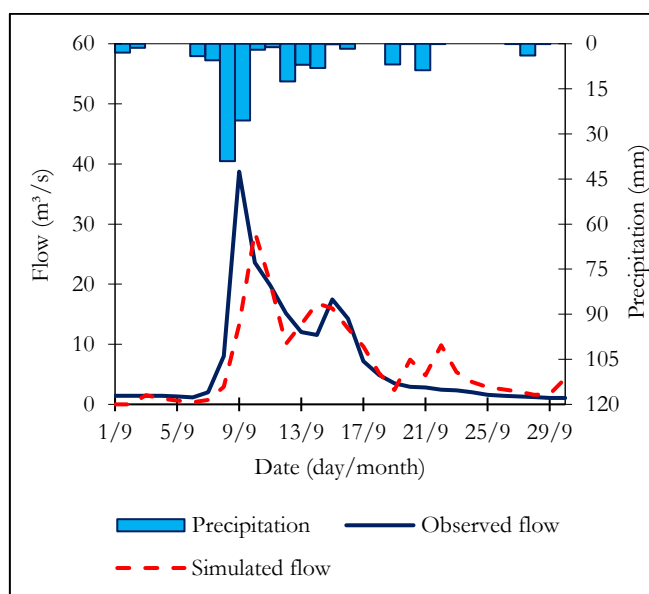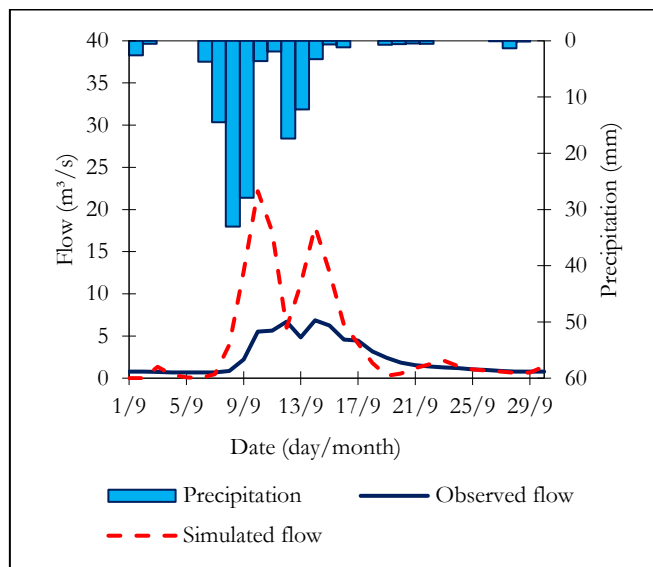**Figure 20.** Observed and validated hydrograph (Itapissirica sub-basin).

Source: Authors.

Figure 21. Observed and validated hydrograph (Nazaré sub-basin).

Source: Authors.

Table 4. ISE, NSE and R² statistical results.

	CARICE		ENGENHO RETIRO		ENGENHO VOLTA		ENGENHO ITAPISSIRICA		NAZARÉ	
	Observed	Calibrated	Observed	Calibrated	Observed	Calibrated	Observed	Calibrated	Observed	Calibrated
Maximum flow (m³/s)	284,8	246,8	81,77	95,43	98,53	112,7	64,79	85,07	46,21	48,7
Minimum flow (m³/s)	3,011	0,1452	1,695	0,02871	1,121	0,09377	1,493	0,07024	1,75	0,2285
Average flow (m³/s)	47,22	51,48	12,92	13,49	10,08	17,26	19,16	19,44	9,365	10,53
Total runoff volume (m³)	118300000	120500000	32360000	31570000	25260000	40400000	48010000	45500000	23470000	24640000

Source: Authors.

Table 5. Result of flow calibration.

	SUB-BASIN CARICÉ		SUB-BASIN ENGENHO RETIRO		SUB-BASIN ENGENHO VOLTA		SUB-BASIN ITAPISSIRICA		SUB-BASIN NAZARÉ	
	Calibration	Validation	Calibration	Validation	Calibration	Validation	Calibration	Validation	Calibration	Validation
Classification of the Integral Square Error	Good	Reasonable	Good	Reasonable	Reasonable	Poor	Reasonable	Good	Good	Poor
Integral Square Error (ISE)	8.74	13.2	8.26	15	24.4	50.8	11.8	9.96	8.29	39.5
Nash-Sutcliffe efficiency (NSE)	0.899	0.711	0.919	0.649	0.508	-2.51	0.411	0.812	0.884	-5.03
Determination coefficient (R²)	0.918	0.82	0.919	0.849	0.745	0.887	0.569	0.836	0.89	0.49

Source: Authors.

Table 6. Result of flow validation.

	CARICE		ENGENHO RETIRO		ENGENHO VOLTA		ENGENHO ITAPISSIRICA		NAZARÉ	
	Observed	Validated	Observed	Validated	Observed	Validated	Observed	Validated	Observed	Validated
Maximum flow (m³/s)	70,16	107,4	37,28	23,07	20,03	52,93	38,72	28,67	6,859	22,17
Minimum flow (m³/s)	1,988	0,2866	0,818	0,004431	0,804	0,002234	1,069	0,3483	0,679	0,03046
Average flow (m³/s)	15,19	18	6,408	3,586	3,208	7,598	7,12	7,656	2,444	5,002
Total runoff volume (m³)	38050000	42120000	16060000	8390000	8037000	17780000	17840000	17910000	6125000	11700000

Source: Authors.

Table 7. Results of parameters after calibration.

Sub-basins	AREA (ha)		WIDTH (m)		SLOPE (%)		IMPERV. (%)		N PERV		DSTORE PERV (mm)		CURVE NUMBER	
	Initial	Calibrated	Initial	Calibrated	Initial	Calibrated	Initial	Calibrated	Initial	Calibrated	Initial	Calibrated	Initial	Calibrated
Guararema	6954,00	7273,88	4336,31	5403,04	6,12	7,35	10,0	14,0	0,030	0,015	1,00	1,72	81,76	54,51
Engenho Volta	25017,07	26167,85	5556,85	6923,84	14,13	16,97	10,0	14,0	0,030	0,015	1,00	1,72	56,52	45,22
Caricé	45016,14	47086,89	7471,09	9308,98	10,00	12,01	10,0	14,0	0,030	0,015	1,00	1,72	79,34	52,90
Goiana	30956,32	32380,31	7368,60	9181,28	5,59	6,71	10,0	14,0	0,030	0,015	1,00	1,72	78,20	52,14
Engenho Retiro	35924,56	35924,56	6166,54	6166,54	10,59	10,59	10,0	6,8	0,030	0,015	1,00	1,00	51,88	19,95
Itapissirica	54269,84	51685,56	8576,45	5360,28	8,67	6,67	10,0	10,0	0,030	0,015	1,00	1,28	51,41	17,14
Nazaré	57052,47	54335,69	7237,94	4523,71	10,80	8,31	10,0	10,0	0,030	0,015	1,00	1,28	53,81	17,94
Palmeirinha	12158,99	11579,99	5356,61	3347,88	14,08	10,83	10,0	10,0	0,030	0,015	1,00	1,28	53,36	17,79
Siriji	11368,12	11368,12	4608,74	4608,74	19,17	19,17	10,0	6,8	0,030	0,015	1,00	1,00	68,24	26,25
Tiúma	7934,44	8299,42	1095,52	1365,02	17,71	21,27	10,0	14,0	0,030	0,015	1,00	1,72	55,70	37,13

Source: Authors.

4. Conclusions

This study presented the sensitivity analysis tools and automatic calibration existing in the PCSWMM hydrodynamic model. The analyzes were performed in the hydrographic basin of the Goiana river in the State of Pernambuco. Observed flow data were used in five fluviometric stations present in the main streams of the basin.

The sensitivity analysis features of the PCSWMM model allow the hydrological parameters of the watershed to be analyzed for various model responses - maximum, minimum, average flow, total runoff volume - enabling an accurate assessment of the most sensitive parameters of the basin.

In relation to the statistical results (ISE, R² and NSE) and the qualitative assessment between the observed and simulated hydrographs, most of the sub-basins showed quite satisfactory results. However, in two sub-basins (Engenho Volta and Nazaré), there was a discrepancy between the calibrated and validated results, with results classified as "satisfactory" and "good" in the calibration (NSE = 0,508 in Engenho Volta and NSE = 0,884 in Nazaré) and "bad" in the validation process (NSE = - 2,51 in Engenho Volta and NSE = -5,03).

The automatic calibration tool available in the PCSWMM model, was quite effective when compared to the trial-and-error calibration approach, which is usually quite time-consuming. In addition to allowing immediate assessment of the

displacement of the hydrograph, from the change of the parameter to be calibrated within the range specified by the user (SRTC resource - Sensitivity-based Radio Tuning Calibration).

References

Agência estadual de planejamento e pesquisa de pernambuco – condepe/fidem. Bacia hidrográfica do rio goiana e sexto grupo de pequenos rios litorâneos - GL 6. *Governo do Estado de Pernambuco Secretaria de Planejamento* – SEPLAN, Recife, PE, 2005.

Ahmadisharaf, E., Camacho, R. A., Zhang, H. X., Hantush, M. M., Mohamoud, Y. M. (2019). Calibration and Validation of Watershed Models and Advances in Uncertainty Analysis in TMDL Studies. *Journal of Hydrologic Engineering*, V. 24 (7) 03119001-1. DOI: 10.1061/(ASCE)HE.1943-5584.0001794.

Chi Water. PCSWMM support. Ontario: CHI Water, 2017. <https://support.chiwater.com/>.

Collischonn, W., Dornelles, F. (2015). Hidrologia para engenharia e ciências ambientais. 2º Edição revisada e ampliada, Associação Brasileira de Recursos Hídricos (ABRH), Porto Alegre, RS.

CPRH – agência estadual de meio ambiente de pernambuco; mma – ministério do meio ambiente. (2011). Diagnóstico Socioambiental - Litoral Norte/Recursos Hídricos Superficiais. *Programa Nacional do Meio Ambiente*. Recife, PE.

Daggupati, P., Yen, H., White, M. J., Srinivasan, R., Arnold, J. G., Keitzer, C. S., Sowa, S. P. (2015). Impact of model development, calibration and validation decisions on hydrological simulations in West Lake Erie Basin. *Hydrological Processes*, V. 29, p. 5307-5320. https://doi.org/10.1002/hyp.10536.

Formiga, K. T. M., Carvalho, M.; Silva, K. A., Soares, A. K., Calibração do Storm Water Management Model (SWMM) utilizando algoritmos evolucionários multiobjetivo. (2016). *Engenharia Sanitária e Ambiental*, v.21, n.4, p. 697-707. https://doi.org/10.1590/s1413-41522016131862.

Garambois, P. A., Roux, H., Larnier, K., Labat, D., Dartus, D. (2015). Characterization of catchment behaviour and rainfall selection for flash flood hydrological model calibration: catchments of the eastern Pyrenees. *Hydrological Sciences Journal*, V.60, p. 424-447. http://dx.doi.org/10.1080/02626667.2014.909596.

Ghofrani, Z., Sposito, V., Faggian, R. (2019). Designing a Pond and Evaluating its Impact upon Storm-Water Quality and Flow: A Case Study in Rural Australia. *Ecological Chemistry and Engineering* S. V. 26(3), p. 475-491. https://doi.org/10.1515/eces-2019-0036.

Gao, J., Miao, C., Duan, Q., Tang, Q., Di, Z.; Liao, W., Wu, J., Zhou, R. (2020). Sensitivity Analysis-Based Automatic Parameter Calibration of the VIC Model for Streamflow Simulations Over China. *Water Resources Research*, V.56, p. 1-19. https://doi.org/ 10.1029/2019WR025968.

Gujarati, D. N., Porter, D. C. (2009) Basic Econometrics 5th ed. *Published by McGraw-Hill/Irwin*, New York, NY.

Jaiswal, R. K., Thomas, T., Galkate, R. V., Glosh, N. C., Lohani, A. K., Kumar, R. (2014). Development of Geomorphology Based Regional Nash Model for Data Scares Central India Region. *Water Resources Management*, V. 28 (2), p. 351-371. https://doi.org/10.1007/s11269-013-0486-x.

James, W. R. C., Wan, B., James, W. (2002). Implementation in PCSWMM using Genetic Algorithms for auto calibration and design- optimization. Global Solutions for Urban Drainage. *Proceedings of the Ninth International Conference on: Urban Drainage*; Portland, OR; United States.

James, W. (2005). Rules for Responsible Modeling. *Compiled and published by CHI (Computational Hydraulics International)*. Guelph, Ontario, Canadá. 4º ed.

Javaheri, H. (1998) Automatic Calibration of Urban Run-off Models Using Global Optimization Techniques. *Department of Civil Engineering and Applied Mechanic McGill University*. Montreal, Canada.

Junior, L. C. G. V., Rodrigues, D. B. B., Oliveira, P. T. S. (2019). Initial abstraction ratio and Curve Number estimation using rainfall and runoff data from a tropical watershed. *Revista Brasileira de Recursos Hídricos*, V. 24, p. 1-9. https://doi.org/10.1590/2318-0331.241920170199.

Krebs, G., Kokkonen, T., Valtanen, M., Koivusalo, H., Setala, H. (2013). A high resolution application of a stormwater management model (SWMM) using genetic parameter optimization. *Urban Water Journal*, V. 10 (6), p. 394-410. http://dx.doi.org/10.1080/1573062X.2012.739631.

Loukcs, D. P., Beek, E. V. (2005). Water Resources Systems Planning and Management and Applications: An Introduction to Methods, Models and Applications. *United Nations Educational, Scientific and Cultural Organization* – UNESCO. Paris.

Loukcs, D. P., Beek, E. V. (2017). Water Resources Systems Planning and Management and Applications: An Introduction to Methods, Models and Applications II. United Nations Educational, Scientific and Cultural Organization – UNESCO. Paris.

Moriasi, D. N., Gitau, M. W., Pai, N., Daggupati, P. (2015). Hydrologic and water quality models: Performance measures and evaluation criteria. *Transactions of the ASABE*. V. 58 (6), p. 1763-1785. doi: 10.13031/trans.58.10715.

Munir, B. A., Ahmad, S. R., Hafeez, S. (2020). Integrated Hazard Modeling for Simulating Torrential Stream Response to Flash Flood Events. *ISPRS International Journal of Geo-Information*, V. 9(1):1, p. 2-21. https://doi.org/10.3390/ijgi9010001

Nash, J. E., Sutcliffe, J. V. (1970). River Flow Forecasting Through Conceptual Models, Part I - A Discussion of Principles. *Journal of Hydrology*, V. 10 (3), p. 282-290. https://doi.org/10.1016/0022-1694(70)90255-6.

Rossman, L. A. (2015). Storm Water Management Model User's Manual Version 5.0. U. S. *Environmental Protection Agency (EPA)*, CINCINNATI, OH.

Sarma, P. B. S., Delleur, J. W., Rao, A. R. (1973). Comparison of rainfall-runoff models for urban areas. *Journal of Hydrology*, V. 18 (3-4), p. 329-347. https://doi.org/10.1016/0022-1694(73)90056-5.

Wang, Y., Zhu, S., Yuan, L., Deng, R. (2020). An automatic parameter calibration method for the SWAT model in runoff simulation. *River Research and Applications*, V.36, p. 1321-1333. https://doi.org/10.1002/rra.3655.