Revista Brasileira de Recursos Hídricos
Brazilian Journal of Water Resources
Versão On-line ISSN 2318-0331
RBRH, Porto Alegre, v. 27, e31, 2022
Scientific/Technical Article

https://doi.org/10.1590/2318-0331.272220220046

Multi-objective calibration of Tank model using multiple genetic algorithms and stopping criteria

Calibração multi-objetivo do Tank Model utilizando diversos algoritmos genéticos e critérios de parada

Juan Carlos Ticona Gutierrez[1], Cassia Brocca Caballero[1], Sofia Melo Vasconcellos[1], Franciele Maria Vanelli[1] & Juan Martín Bravo[1]

[1]Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil
E-mails: juan.ticona@ufrgs.br (JCTG), cassiabrocca@gmail.com (CBC), sofimello@hotmail.com (SMV), franciele.vanelli@ufrgs.br (FMV), jumarbra@iph.ufrgs.br (JMB)

Received: June 03, 2022 - Revised: October 16, 2022 - Accepted: October 19, 2022

ABSTRACT

Calibration of hydrologic models estimates parameter values that cannot be measured and enable the rainfall-runoff processes simulation. Multi-objective evolutionary algorithms can make the calibration faster and more efficient through an iterative process. However, the standard stopping criterion used to stop the iterative process is to reach a pre-defined number of iterations defined by the modeller. Alternatively, the Ticona stopping criterion is based on the minimum number of iterations required to achieve a determined number of non-dominated solutions in the Pareto front, resulting in a reduction of the computational time without losing performance during the calibration processes. We evaluated the Ticona stopping criterion in the Tank Model calibration. The calibration processes were performed using data from two river basins, with three genetic algorithms and two objective functions. The Ticona stopping criterion required a computational time 27.4% to 44.1% lower than using the standard stopping criterion and were obtaining similar results in simulated streamflow time series and similar values of the best set of parameters.

Keywords: Multi-objective evolutionary algorithm; Tank model; Stopping criterion; NSGA-II; NSGA-III; SPEA-II.

RESUMO

A calibração de modelos hidrológicos estima os valores de parâmetros que não podem ser mensurados e permite a simulação dos processos chuva-vazão. Os algoritmos evolucionários multi-objetivos podem tornam a calibração mais rápida e eficiente por meio de processos iterativos. Contudo, o critério de parada padrão usado para encerrar o processo iterativo é baseado em um número de iterações pré-definido pelo usuário. Como alternativa, o critério de parada Ticona é baseado no número mínimo de iterações requerido para alcançar um determinado número de soluções não-dominadas na Frente de Pareto, resultando em um menor tempo computacional sem perda de desempenho durante a calibração. Neste estudo, foi avaliado o uso do critério de parada Ticona na calibração do Tank Model. A calibração foi realizada em duas bacias hidrográficas, usando três algoritmos genéticos e duas funções-objetivo. Os resultados indicaram um tempo computacional 27,4% a 44,1% menor quando utilizado o critério de parada Ticona em comparação com o critério de parada padrão, ao mesmo tempo que foram obtidos resultados similares quanto aos valores dos parâmetros calibrados e à série temporal de vazão simulada.

Palavras-chave: Algoritmo evolucionário multi-objetivo; Tank model; Critério de parada; NSGA-II; NSGA-III; SPEA-II.

INTRODUCTION

Hydrologic models are efficient tools for simulating rainfall-runoff processes, representing the processes of the hydrological cycle in a basin in a simplified way (Beven. 2019). These models are based on mathematical equations and assumptions that aim to simplify the real-world system (Gupta et al., 1998). Thus, measured data are used as input to the model and the parameter values that cannot be measured because they represent an abstraction of reality are estimated by the calibration process.

The calibration process aims to find parameter values that represent the hydrological behaviour of the basin as closely as possible (Madsen, 2000). Different combinations of parameters can represent the observed data because of the model's simplifications as well as the inherent data uncertainties (Beven, 1993, 2001; Beven & Smith, 2015). Single or multi-objective optimization algorithms can be coupled to hydrologic models during the calibration to make this process faster and more efficient (Madsen, 2000). Single-objective optimization seeks to find a solution that minimises or maximises the objective function. Differently, multi-objective optimization results in a set of solutions that are not inferior or not dominated, known as the Pareto front (Yapo et al., 1998).

For the optimization of multi-objective problems, such as the calibration of hydrologic models, evolutionary algorithms simulate the basic principles of the evolutionary process in a set (population) of candidate solutions (individuals). Through an iterative process, where an iteration represents a generation, the individuals are sorted by their fitness and used to generate new individuals by the so-called evolutionary operators such as crossover and mutation (Coello et al., 2007; Gutierrez et al., 2019a). The following elements must be defined to apply a Multi-Objective Evolutionary Algorithm (MOEA) in hydrologic-model calibration: (1) objective functions, (2) optimisation algorithm, and (3) stopping criterion (Gupta et al., 1998). As the size of the population grows and the number of objective functions increases, the optimization process may be time-consuming, requiring a prior indication of a stopping criterion to analyse when the solutions reached are acceptable and additional calculation is not justified (Gutierrez et al., 2019b). Several stopping criteria can be used to verify if the solution achieved is the best answer or an acceptable answer. These criteria check if solutions do not change in a determined number of iterations. If the algorithm does not stop, a limit to the maximum number of iterations should be defined. This last criterion is referred to as the standard stopping criterion.

Gutierrez et al. (2019a) developed a stopping criterion that ensures the quality of the solutions of the multi-objective optimization process and interrupts it when no significant improvement occurs, reducing computational resources considerably. The approach proposed by the authors aims to find a representative set of solutions, sufficiently close to that which would be obtained using the standard stopping criterion. Gutierrez et al. (2019a) stopping criterion have three parameters that can be defined with some previous calibration tests: the minimum number of generations (*Imin*) required to reach the population size of non-dominated solutions (*Np*) and the number of generations that this population size must be maintained consecutively during the iterative process ($Count_{Max}$). Gutierrez et al. (2019a) used the stopping criterion during the calibration of seven parameters of the IPH-II conceptual hydrologic model and found that the metric values showed results similar to those obtained using the standard stopping criterion. Despite both stopping criteria resulted in similar Pareto fronts for the three MOEAs tested (NSGA-II, NSGA-III and SPEA-II), an expressive computational gain was observed with Gutierrez et al. (2019a) stopping criterion, decreasing the computational time by 70% compared to the standard stopping criterion. Some models, such as the Tank Model, have a higher number of parameters and, consequently, more computational time is spent on running the model code and generating the target output (Yilmaz et al., 2010).

The Tank Model is a conceptual, lumped rainfall-runoff model based on the representation of the hydrological system by a succession of vertically aligned tanks, simulating the different soil layers with their respective water retention and transfer properties (Sugawara & Singh, 1995). The resulting flow is a function of the precipitation, evapotranspiration, and water storage in the previous time step (Suryoputro et al., 2017; Jaiswal et al., 2020). Different methods have already been applied to calibrate the Tank model. Some examples of single-objective methods used to calibrate the Tank model are the Uniform random search, Pattern search, Rosenbrock method, Generalized reduced gradient method, Generic GA, and SCE-UA algorithm (e.g.: Tanakamaru, 1995; Madsen, 2000; Song et al., 2017) and multi-objective methods, such NSGA-II (Vasconcellos, 2017).

Therefore, this study aims to investigate if the computational gain in using the Gutierrez et al. (2019a) stopping criterion in comparison to the standard stopping criterion still holds on when applied in the calibration of a hydrologic model with a more complex structure. The IPH II and Tank Model are conceptual hydrologic models, however, the IPH II structure is simpler and had a lesser number of parameters (7 parameters). The Tank Model is highly non-linear, and in this study, were tested two structures (4-tank: 16 parameters, and 3-tank: 12 parameters). Thus, the Gutierrez et al. (2019a) stopping criterion, called the Ticona stopping criterion in this study, was tested in the multi-objective calibration of the Tank Model, applied in two different river basins with a larger calibration dataset.

MATERIAL AND METHODS

Study area

Two basins with different morphological characteristics were selected to compare the performances of the Ticona stopping criterion and the standard stopping criterion, during the calibration of the Tank model, as analysed by Gutierrez et al. (2019b) using the IPH II model.

The Ijuí River basin is in the northwest region of the state of Rio Grande do Sul (south of Brazil), the basin has a drainage area of 5,414 km^2. The terrain is composed of hills in regions of fields with smooth slopes that vary between 3 to 15%. The concentration time estimated by the Kirpich formula is 2 days. The daily hydrological data were obtained from the National Water Agency, Brazil (ANA - Agência Nacional de Águas-, Brazil) dataset in the period from 01/01/2003 to 12/31/2018. The seven gauging stations in the basin provided the data used in this study. Figure 1a shows the daily hyetograph

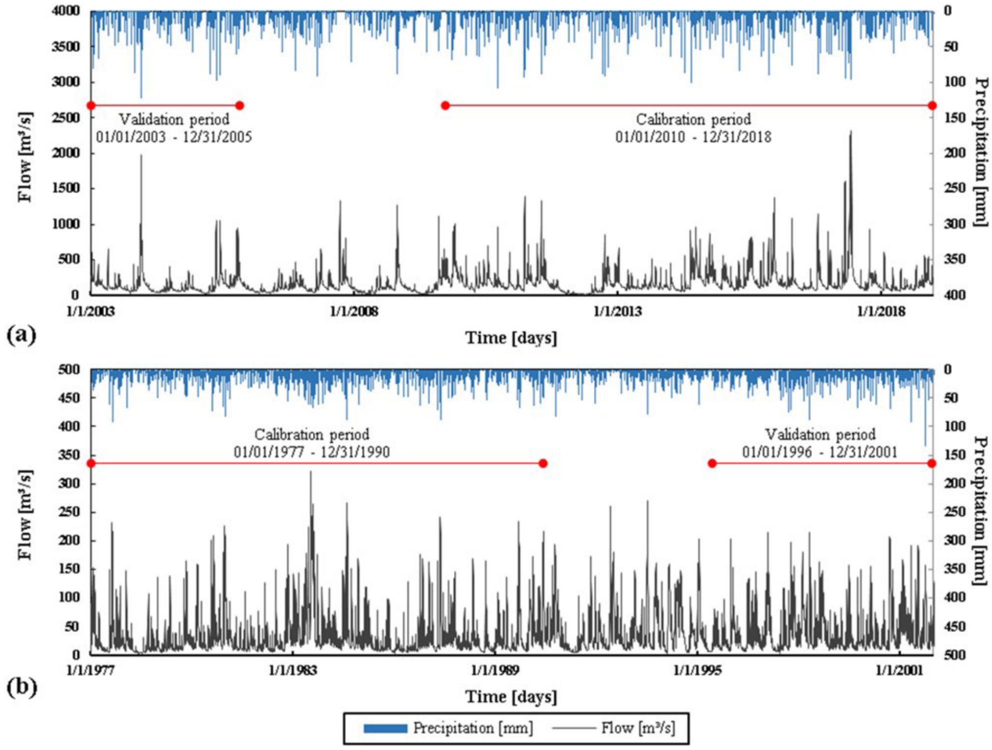

Figure 1. (a) Ijuí River basin – Precipitation and streamflow time series during calibration (01/01/2010 – 12/31/2018) and validation (01/01/2003 – 12/31/2005) of the Tank model. (b) Vila Canoas River basin – Precipitation and streamflow time series during validation (01/01/1996-12/31/2001) and calibration (01/01/1977-12/31/1990) of the Tank model.

and hydrograph used in the simulations, with the calibration period from 2010 to 2018 and the validation period from 2003 to 2005. The period between 2006 and 2009 was not used due to missing data.

The Vila Canoas River basin is in the state of Santa Catarina (south of Brazil) and it has a drainage area of 989 km². Unlike the Ijuí River basin, the Vila Canoas basin is in a mountainous region, with very shallow soils over basaltic rocks and sandstone, with a predominance of fields, sparse forests in areas of the greater slope, and some cultivated areas and reforestation. The concentration time was estimated at 2 days using the Kirpich formula. The hydrological data were also obtained from the ANA dataset. Five gauging stations were identified and used. Figure 1b shows the daily hyetograph and hydrograph for simulations, with the period of calibration from 1977 to 1990 and validation period from 1996 to 2001. The period from 1991 to 1995 was not used due to missing data.

Hydrologic model

The Tank model estimates runoff from precipitation data (Sugawara, 1961, 1972, 1979; Sugawara & Singh, 1995). It is a lumped hydrologic model that simulates the water balance of a basin using tanks or reservoirs arranged in a vertical series, where the storage of the first tank is determined by precipitation and the storage of the other tanks is determined by the infiltration from the upper tank (Sugawara and Singh 1995). For a better understanding of the Tank model structure, the reader could access Sugawara (1961, 1972, 1979), and Sugawara & Singh (1995).

The Tank model has never been applied in these basins; therefore, we tested a 3-tank (Figure 2b) and 4-tank (Figure 2a) structure aiming to evaluate the best number of tanks to represent the Ijuí River basin and Vila Canoas River basin.

The parameters that can be calibrated in the Tank model are the runoff and infiltration coefficients and the coefficients that determine the storage capacity (such as the height of the side outlets) of each reservoir in the model (Haan, 1989). They are related to the soil type and land use, and geological characteristics of the basin (Ishihara & Kobatake, 1979). Table 1 shows the parameters of the Tank model and the range values for each of the parameters. This range of values was applied in the calibration of the 3-tank and 4-tank structures of the Tank model used in this study.

Objective functions

Two objective functions were used in the multi-objective calibration of the Tank model: the $RMSE_{inv}$ (Equation 1) and the NS defined by Nash & Sutcliffe (1970) (Equation 2), also referred to in this study as OF1 (Objective Function 1) and OF2 (Objective Function 2), respectively.

$$RMSE_{inv} = \sqrt{\frac{1}{NT}\sum_{t=1}^{NT}\left(\frac{1}{Q_{Obs,t}} - \frac{1}{Q_{Calc,t}}\right)^2} \quad (1)$$

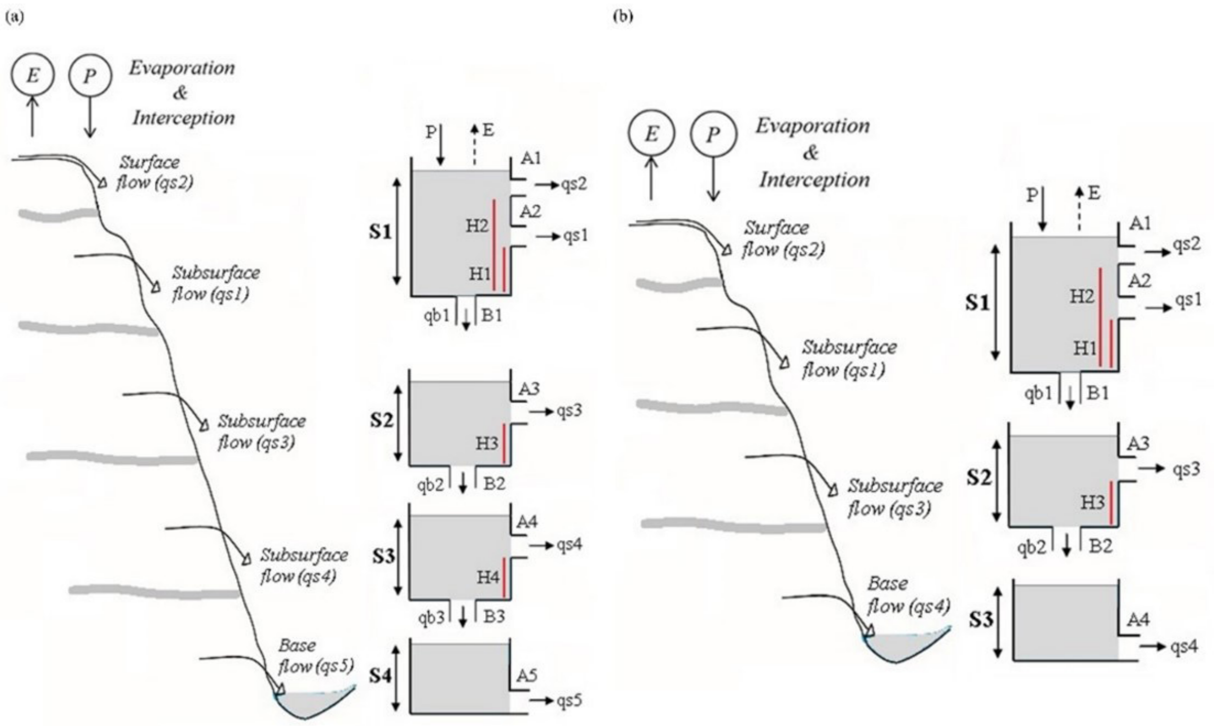

Figure 2. Representation of the Tank model with 4-tank (a) and 3-tank (b) structure (Adapted from Setiawan et al., 2003).

Table 1. Range values of parameters in the Tank model.

Parameter	Description	Unit	Range	
			Minimum	Maximum
S1I[a]	Initial storage of tank 1	mm	5	50
S2I[a]	Initial storage of tank 2	mm	10	50
S3I[a]	Initial storage of tank 3	mm	20	100
S4I[b]	Initial storage of tank 4	mm	30	500
H1[a]	Height of side outlet 1 in tank 1	mm	10	70
H2[a]	Height of side outlet 2 in tank 1	mm	10	45
H3[a]	Height of side outlet of tank 2	mm	10	70
H4[b]	Height of side outlet of tank 3	mm	10	70
A1[a]	Runoff coefficient (tank 1)	min^{-1}	0.09	0.5
A2[a]	Sub-surface flow coefficient (tank 1)	min^{-1}	0.09	0.5
A3[a]	Intermediate flow coefficient (tank 2)	min^{-1}	0.09	0.5
A4[a]	Sub-base flow coefficient (tank 3)	min^{-1}	0.01	0.1
A5[b]	Base flow coefficient (tank 4)	min^{-1}	0.001	0.01
B1[a]	Infiltration coefficient in tank 1	min^{-1}	0.01	0.1
B2[a]	Infiltration coefficient in tank 2	min^{-1}	0.01	0.1
B3[b]	Infiltration coefficient in tank 3	min^{-1}	0.001	0.1

[a]Parameter common to both the 3-tank and 4-tank structures of the Tank model; [b]Parameter presents only in the 4-tank structure.

$$NS = 1 - \left[\frac{\sum_{t=1}^{NT} \left(Q_{Obs,t} - Q_{Calc,t} \right)^2}{\sum_{t=1}^{NT} \left(Q_{Obs,t} - \underline{Q}_{Obs} \right)^2} \right] \quad (2)$$

where $Q_{Calc,t}$ is the calculated flow in time interval t (daily), Q_{Obs} is the average of the observed flow values, $Q_{Obs,t}$ is the value of the observed flow in time interval t and NT is the number of time intervals. $RMSE_{inv}$ values vary from 0 to $+\infty$, with 0 being the ideal value, indicating full agreement between observations and simulations. NS values are in the range of $-\infty$ to 1.0, where 1.0 is the ideal value (Krause et al., 2005).

NS is a performance measure commonly used to assess the quality of the simulated hydrograph, that emphasizes the representation of high flows (Legates & McCabe Junior, 1999; Moussa & Chahinian, 2008; Reynolds et al., 2017). To reproduce the entire shape of the hydrograph and not just the high flows, the $RMSE_{inv}$ was also used because it emphasized the representation of low flows (Pushpalatha et al., 2012; Garcia et al., 2017). Thus, both objective functions were included in the multi-objective calibration simultaneously to consider the maximum and minimum flow representation by the model. Therefore, the optimisation goal was to maximise *NS* and minimise $RMSE_{inv}$.

Multi-objective calibration of hydrologic models using genetic algorithms

The multi-objective calibration of hydrologic models is an optimisation process in which evolutionary algorithms, such as genetic algorithms (GA), are commonly used. GAs are a particular class of evolutionary algorithms that apply techniques inspired by evolutionary biology, such as heredity, mutation and natural selection. During successive generations (iterations), the population (set of candidate solutions) converge towards an approximation of the Pareto front after a stopping criterion is satisfied (Fiben & Smith 2003; Chugh et al., 2019).

The standard stopping criterion used to stop the iterative process is to reach a pre-defined number of generations N_{Gen}. Meanwhile, in the Ticona stopping criterion (Gutierrez et al., 2019a), a counter of consecutive generations is initialised each time the generation number is higher than *Imin* and at least N_P non-dominated solution exists in the population. The counter adds *1* if in the next generation both conditions are satisfied, otherwise returns to a value equal to zero. If the counter reaches the $Count_{Max}$ value, the iterative process stops.

Several multi-objective GA as the NSGA-II, NSGA-III, MOEA/D, SPEA2 and others have been used in many applications and their performances tested. In the present study, some of the most representative algorithms for multi-objective optimization have been selected (Echevarría et al., 2016): NSGA-II (Deb et al., 2002), NSGA-III (Deb & Jain, 2014) and SPEA-II (Zitzler et al., 2001). These three MOEAs have been used in the calibration of different hydrologic models (e.g.: Rozos et al., 2004; Rouhani et al., 2007; Khu et al., 2008; Shafii & De Smedt 2009; Le et al., 2016; Mostafaie et al., 2018; Adeyeri et al., 2020).

The simulations were performed using MATLAB R2018a version 9.4.0.813654 in a computer with an Intel Core i5-3337U 1.80GHz processor operating with 8GB of RAM. During the calibration of the hydrologic model, $Np = 100$ and $Ngen = 500$ were adopted for the standard stopping criterion as suggested by Pushpalatha et al. (2012) and Yapo et al. (1998). For the Ticona stopping criterion, $Np = 100$, $Imin = 110$ and $CountMax = 10$ were used as suggested by Gutierrez et al. (2019a).

Selection of the calibrated parameters of the Tank model

The calibrated parameter values of the Tank model were defined as the best solution from the sets of non-dominated solutions that generated the Pareto Fronts (PFs). As it is possible to obtain different PFs for the same MOEA when the calibration process is repeated (i.e. using other initial population), selecting the best solution requires that first, the best PF is selected, and then the best solution among the PF is selected.

Therefore, to select the best PF in each calibration attempt (*w*), the criterion of the lowest average Euclidean distance ($D_{i,w}$) (Equation 3) was used, as adopted by Gutierrez et al. (2019a). According to this criterion, the best PF corresponds to the lowest value of $D_{i,w}$ when the stopping criterion is satisfied.

$$D_{i,w} = \left[\frac{1}{ND_i}\sum_{j=1}^{ND_i}\left(OF_{1(i,j)} - OF_{1rel}\right)^2 + \left(OF_{2(i,j)} - OF_{2rel}\right)^2\right]^{0.5} \tag{3}$$

From the best PF, the best solution was selected based on the minimum distance (D_{min}) to a reference result, derived from Equations 4 and 5.

$$R_s = \left[\left(OF_{1(s)} - OF_{1rel}\right)^2 + \left(OF_{2(s)} - OF_{2rel}\right)^2\right]^{0.5} \tag{4}$$

$$D_{min} = \left\{R_1, R_2, R_3, \ldots, R_{N_p}\right\} \tag{5}$$

where: $OF1_{(i,j)}$ and $OF2_{(i,j)}$ is the j^{TH} non-dominated solution of the PF in generation *i* of the calibration attempt; ND_i é is the number of non-dominated solutions for generation *i*; $OF1_{(s)}$ and $OF2_{(s)}$ are the s^{TH} non-dominated solution of the Pareto front in the final generation; $OF1_{rel}$ and $OF2_{rel}$ are a point relative optimum ($OF1_{rel} = 0$ and $OF2_{rel} = 1$); and *Np* is the number of solutions of the PF in the final generation.

MOEA Performance assessment based on the stopping criteria

Several metrics to assess the performance of MOEAs were already proposed (Ishibuchi et al., 2015; Yen & He, 2013; Araújo et al., 2011). In this study, we selected three well-known metrics to compare the MOEAs performance based on the stopping criterion used when applied to the Tank model calibration. The maximum spread (MS – Zitzler et al., 2000) and spacing (SP - Schott, 1995) metrics were used to evaluate the diversity of solutions across the PF. We also used the generational distance (GD - Van Veldhuizen, 1999) metric that measures the proximity of solutions to the true Pareto front (PF_{true}).

The PF_{true} of a multi-objective calibration process is generally unknown. Ishibuchi et al. (2014) showed how it could be achieved using multiple executions of one or more MOEAs. Using this approach and considering the three MOEAs described previously, we created a PF_{true} for the calibration process of each basin (Figure 3).

EXPERIMENTAL DESIGN

The performance of the stopping criterion developed by Gutierrez et al. (2019a) was evaluated in three MOEAs (NSGA-II, NSGA-III and SPEA-II) in the calibration of the Tank model arranged in a 3-tank (12 parameters) and a 4-tank (16 parameters) structure, compared with the results using the standard stopping criterion. Two objective functions were considered: the maximisation of the NS and the minimisation of the $RMSE_{inv}$.

Twenty initial population sets were used for each MOEA test. These initial population sets (size equal to 100) were generated uniformly at random in the range specified in Table 1. The results with the standard and Ticona stopping criteria were obtained in the same calibration processes. When the MOEA reached the Ticona stopping criterion, the results were archived, and the computation continued until the maximum number of generations was reached according to the standard stopping criterion. The calibration procedure was performed twenty times with each MOEA using both stopping criteria, for a total of sixty results for each study basin.

The calibrated parameter values of the Tank model were defined as the non-dominated solution with D_{min} from the PF with lower $D_{i,w}$. Hydrographs were generated for each study basin by the Tank model using the calibrated parameter values obtained for each MOEA.

RESULTS AND DISCUSSION

Performance evaluation based on the number of tanks in the Tank model structure

The objective function NS presented slightly better accuracy for the 4-tank structure than for the 3-tank structure (Table 2).

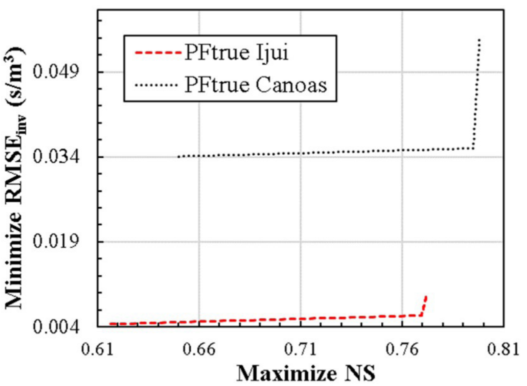

Figure 3. True Pareto Fronts to derive GD based on the calibration results of the Ijuí and Canoas River basins.

The 3-tank structure provided similar performance in representing high flows as the 4-tank structure of the Tank model. However, the 3-tank structure could not represent the low flows correctly, as presented in Figure 4 and Figure S1 (Supplementary Material), which show the comparison between the observed and simulated flow for the 3-tank structure using parameters calibrated by MOEAs for the Vila Canoas River basin and the Ijuí River basin, respectively. The lower performance of the 3-tank structure in representing low flows in both basins was also reflected by the $RMSE_{inv}$ values shown in Table 2. For this reason, the 4-tank structure of the Tank model was selected to generate the following results.

Standard stopping criterion vs Ticona stopping criterion: performance metrics and best PF results

The best PF with the lowest average Euclidean distance ($D_{i,w}$) was selected from twenty calibration attempts (w = 1, …, 20) performed with each MOEA (Table 3). The value of $D_{i,w}$ obtained using the Ticona stopping criterion was very similar to the value obtained when applying the standard stopping criterion ($D_{500,\,w}$).

The metrics (SP, MS, and GD) calculated for the best PF found, for each MOEA, are presented in Table 4 for the Vila Canoas River basin, and in Table S1 (Supplementary Material) for the Ijuí River basin. Results showed a small variation of the metrics values for the best PF found, in both basins, when using the Ticona stopping criterion and the standard stopping criterion.

The 20-run results analysis is presented hereafter. The metrics (SP, MS and GD) calculated and computational times during calibration, for all performed runs, are shown in Figure 5 and Figure 6 for the Ijuí and Vila Canoas River basins, respectively. Analysing the median of the metrics obtained in the Ijuí River basin (Figure 5) with both stopping criteria, the SP and MS metrics showed a better performance using NSGA-III. For the metric GD, the SPEA-II algorithm presented a higher approximation with the PF_{true}. For the Vila Canoas River basin (Figure 6), GD and SP metrics performed better when the NSGA-III and NSGA-II algorithms were applied, respectively, for both stopping criteria. However, based on the MS metric, the SPEA-II algorithm showed a better performance when the Ticona stopping criterion was used, while the NSGA-III algorithm performed better when used the standard stopping criterion.

Table 2. Objective Functions values during calibration of the Ijuí and Vila Canoas River basins.

Basin	MOEA	3-tank structure				4-tank structure			
		CALIBRATION[a]		CALIBRATION[b]		CALIBRATION[a]		CALIBRATION[b]	
		NS	$RMSE_{inv}$ *	NS	$RMSE_{inv}$ *	NS	$RMSE_{inv}$ *	NS	$RMSE_{inv}$ *
Ijuí	NSGA-II	0.6519	19963.5	0.6515	19887.1	0.7645	0.00626	0.7640	0.00628
	NSGA-III	0.6521	19887.1	0.6518	20190.7	0.7638	0.00623	0.7585	0.00620
	SPEA-II	0.6489	18947.0	0.6528	20039.5	0.7650	0.00626	0.7633	0.00632
Vila Canoas	NSGA-II	0.7002	23317.6	0.7004	23692.0	0.7914	0.03604	0.7913	0.03615
	NSGA-III	0.6934	20884.0	0.6921	20743.1	0.7925	0.03627	0.7900	0.03607
	SPEA-II	0.7001	23484.8	0.7004	23567.9	0.7885	0.03617	0.7902	0.03644

[a]Standard stopping criterion; [b]Ticona stopping criterion; *Unit: (s/m^3); NS: Nash-Sutcliffe efficiency; $RMSE_{inv}$: Root Mean Square Error for inverse flows.

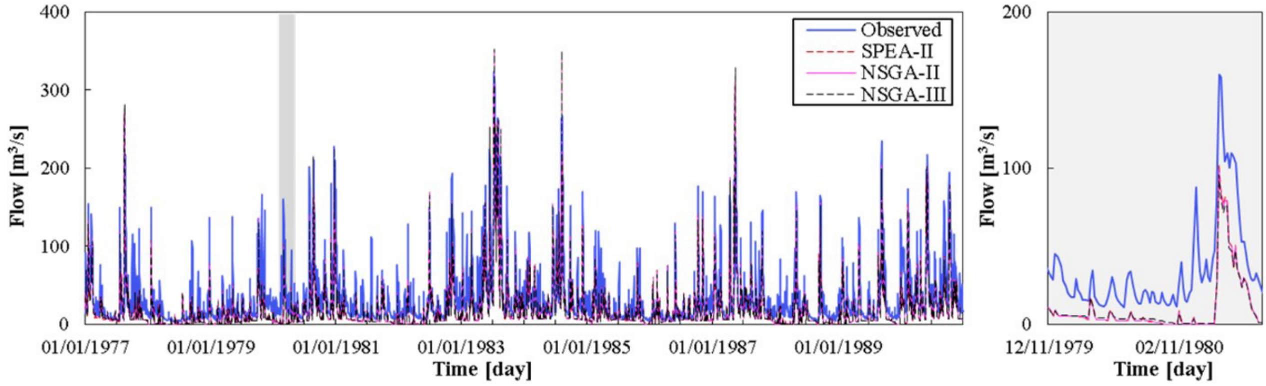

Figure 4. Hydrograph of the calibration period for the Vila Canoas River basin using the 3-tank structure.

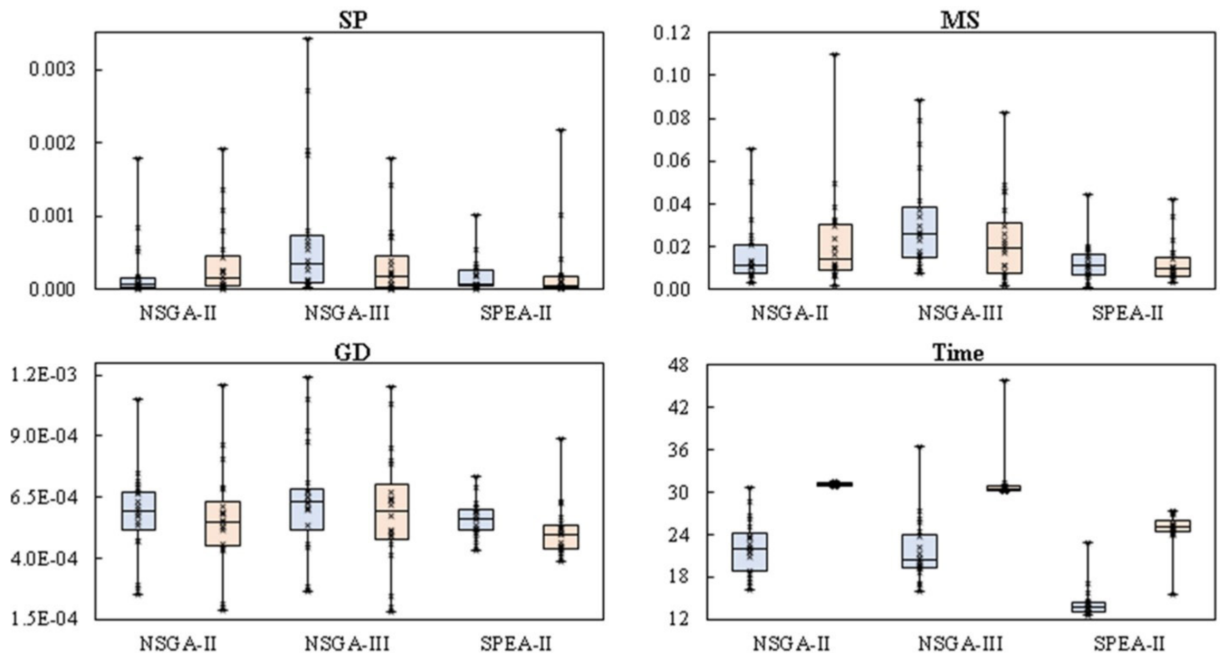

Figure 5. Boxplot of metrics and computational time for each MOEA (NSGA-II, NSGA-III, SPEA-II); using the Ticona (blue) and standard (red) stopping criterion in the Ijuí River basin.

Table 3. Values of the lowest average distance obtained for the Pareto front with each MOEA, with the standard stopping criterion ($D_{500,w}$) and Ticona stopping criterion ($D_{i,w}$), in both basins.

MOEA	Ijuí River basin			Vila Canoas River basin		
	$D_{i=500,w}$	$D_{i,w}$	i	$D_{i=500,w}$	$D_{i,w}$	i
NSGA-II	0.4576	0.4587	345	0.4623	0.4634	261
NSGA-III	0.4565	0.4576	324	0.4641	0.4632	283
SPEA-II	0.4572	0.4612	261	0.4649	0.4669	261

Table 4. Metrics of the best PF of each MOEA used; with standard and Ticona stopping criterion in the Vila Canoas River basin.

MOEA	Metrics[a]			Metrics[b]		
	SP	MS	GD	SP	MS	GD
NSGA-II	0.000055	0.009015	0.000248	0.000045	0.008048	0.000286
NSGA-III	0.000447	0.016029	0.000247	0.001193	0.031362	0.000331
SPEA-II	0.000633	0.032155	0.000441	0.000178	0.024078	0.000398

[a]Standard stopping criterion; [b]Ticona stopping criterion; SP: Spacing; MS: Maximum Spread; GD: Generational Distance.

Computational time is an important parameter when comparing the performance of different MOEAs, as it measures the efficiency of the optimisation process: a shorter time is better (Olazar, 2007). For both basins, MOEAs using the Ticona stopping criterion required less computational time than when they used the standard stopping criterion, as presented in Figure 5 and Figure 6. The computational time required using the Ticona stopping criterion for the Ijuí River basin was 28.8% (NSGA-II), 30.5% (NSGA-III) and 27.4% (SPEA 2) shorter than the standard stopping criterion. For the Vila Canoas River basin, computational time decreased by 42.7% (NSGA-II), 43.2% (NSGA-III) and 44.1% (SPEA 2) using the Ticona stopping criterion compared to the standard stopping criterion.

When the hydrologic model IPH-II was applied, the computational time gain from using the Ticona stopping criterion during calibration with the same MOEAs (as presented by Gutierrez et al., 2019b) was higher (70% to 77%) than the gains obtained during calibration of the Tank model (27.4% to 44.1%) in the same basins and periods. The present study suggests that the higher number of parameters to be calibrated (16 parameters for the 4-tank structure of the Tank model in comparison to 7 parameters for the IPH-II model) tends to increase the objective-functions complexity. Thus, the number of generations also increased before the MOEAs find a good representation of the PF, stopping the search.

Standard stopping criterion vs Ticona stopping criterion: results of the multi-objective Tank model calibration and validation

The results of analysing the hydrographs of the Vila Canoas and Ijuí River basins after applying the performance metrics are shown in Table 5. The *NS* showed values ranging

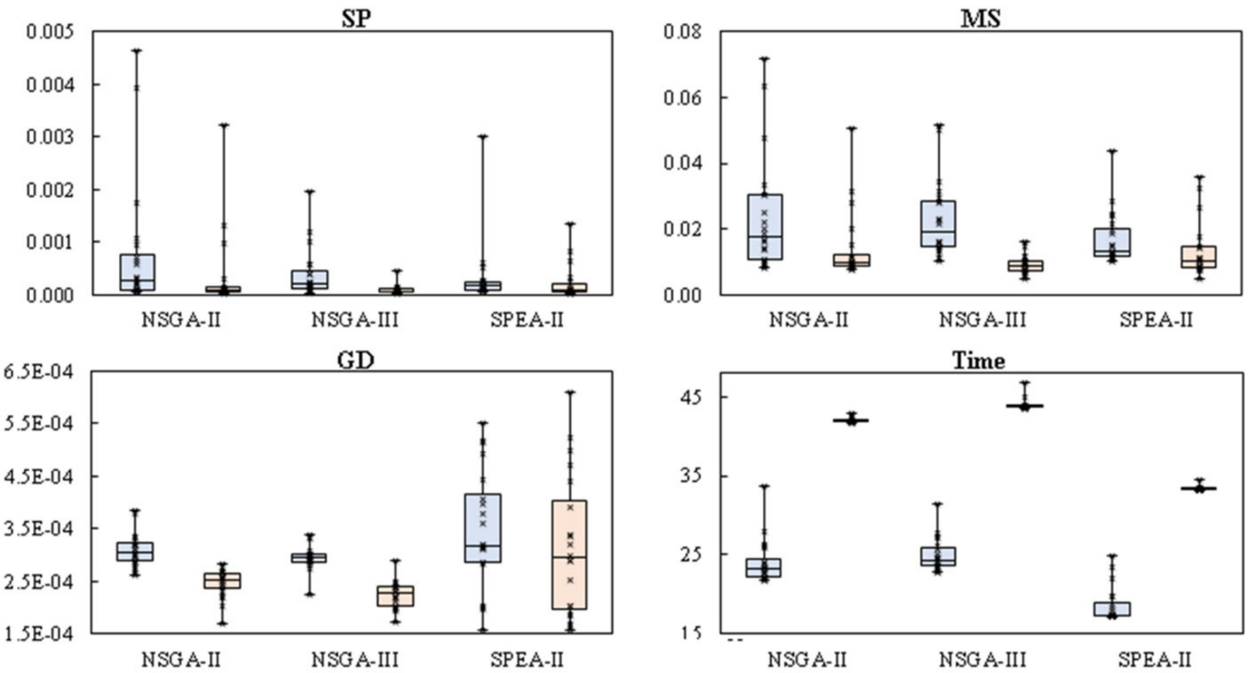

Figure 6. Boxplot of metrics and computational time for each MOEA (NSGA-II, NSGA-III, SPEA-II); using the Ticona (blue) and standard (red) stopping criterion in the Vila Canoas River basin.

Table 5. Objective Functions values for each MOEA during calibration and validation periods in the Vila Canoas and Ijuí River basins.

Basin	MOEA	CALIBRATION[a]		CALIBRATION[b]		VALIDATION[a]		VALIDATION[b]	
		NS	$RMSE_{inv}$ *	NS	$RMSE_{inv}$ *	NS	$RMSE_{inv}$ *	NS	$RMSE_{inv}$ *
Ijuí	NSGA-II	0.7645	0.0063	0.7640	0.0063	0.6192	0.0056	0.6131	0.0057
	NSGA-III	0.7638	0.0062	0.7585	0.0062	0.6325	0.0056	0.6178	0.0057
	SPEA-II	0.7650	0.0063	0.7633	0.0063	0.6334	0.0056	0.6295	0.0055
Vila Canoas	NSGA-II	0.7914	0.0360	0.7913	0.0361	0.8107	0.0293	0.8110	0.0292
	NSGA-III	0.7925	0.0363	0.7900	0.0361	0.8118	0.0294	0.8094	0.0297
	SPEA-II	0.7885	0.0362	0.7902	0.0364	0.8072	0.0326	0.8093	0.0318

[a]Standard stopping criterion; [b]Ticona stopping criterion; *Unit: (s/m^3); NS: Nash-Sutcliffe efficiency; $RMSE_{inv}$: Root Mean Square Error for inverse flows.

from 0.7585 (0.7885) to 0.7650 (0.7925) for the calibration period and from 0.6131 (0.8072) to 0.6334 (0.8118) for the validation periods in the Ijuí River basin (Vila Canoas River basin). $RMSE_{inv}$ presented values very close to zero for all analysed MOEAs in both the calibration and the validation periods, and both basins, indicating a good representation of low flows. The performance metric values were very similar using any of the MOEA with both stopping criteria, and in both periods (calibration and validation). Also, the performance of the Tank model during calibration and validation periods based on the parameter values found by several MOEAs using the Ticona stopping criterion was like the performance observed in previous studies (e.g. Tanakamaru, 1995; Madsen, 2000; Song et al., 2017; Vasconcellos, 2017).

The values of calibrated parameters adopted considering the best solution (the one with D_{min}) of the best PF obtained in each MOEA are presented for the Ijuí River basin and the Vila Canoas River basin in Figure 7. The set of parameters of the Tank model found for each MOEA showed similar values for both stopping criteria.

Hydrographs were generated in the calibration and validation periods, in each basin, based on these parameter values. In the Vila Canoas River basin, Figure 8a and Figure 8b show the hydrographs during the calibration period using the parameter values defined by the standard stopping criterion and Ticona stopping criterion, respectively; while the hydrographs from the validation period are shown in Figure 9a and Figure 9b, respectively. As observed in these figures, the standard stopping criterion and the Ticona stopping criterion resulted in very similar hydrographs for both study basins, showing a good representation of the high and low flows.

Similar results were found in the calibration/validation of the Tank model in the Ijuí River basin, using the standard stopping criterion and the Ticona stopping criterion, as presented in Figure S2 and Figure S3 (Supplementary Material).

Therefore, the Ticona stopping criterion yields good and similar results to those obtained by the standard stopping criterion, while reducing the computational processing time. Gutierrez et al. (2019b), performing the same analysis as this study in the calibration of the hydrologic model IPH-II, obtained very similar parameter values and hydrographs during the calibration and validation periods, with both stopping criteria. As already stated, the IPH-II is conceptually simpler than the Tank model. From this analysis, we can affirm that the Ticona stopping criterion is also advantageous when applied to a more complex model (4-tank structure Tank Model) in different study areas.

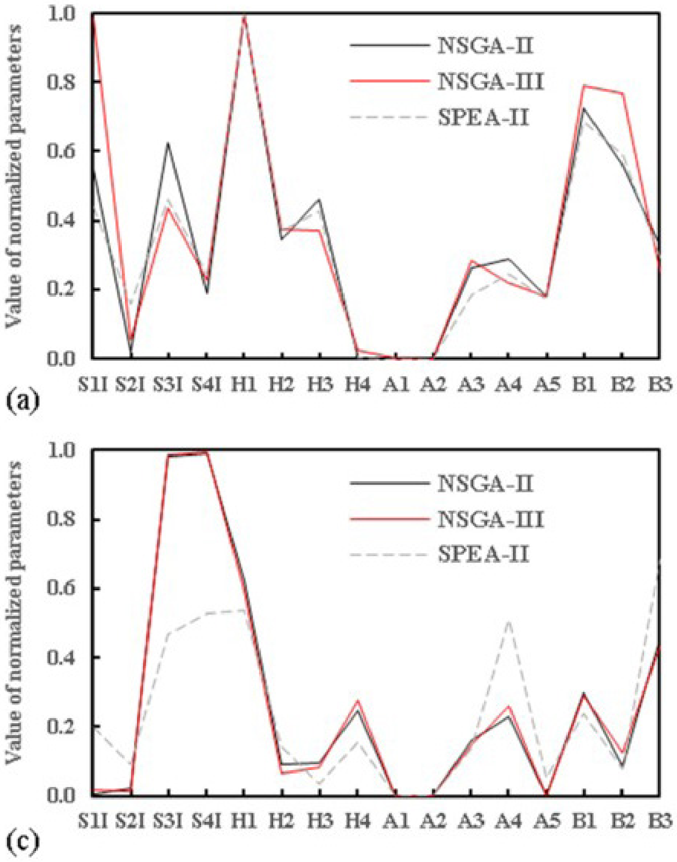

Figure 7. Calibrated parameter values of the best solution from the best PF found by each MOEA with both stopping criteria, (a, c - Standard stopping criterion; b, d - Ticona stopping criterion), in the Ijuí River basin (a, b) and the Vila Canoas River basin (c, d).

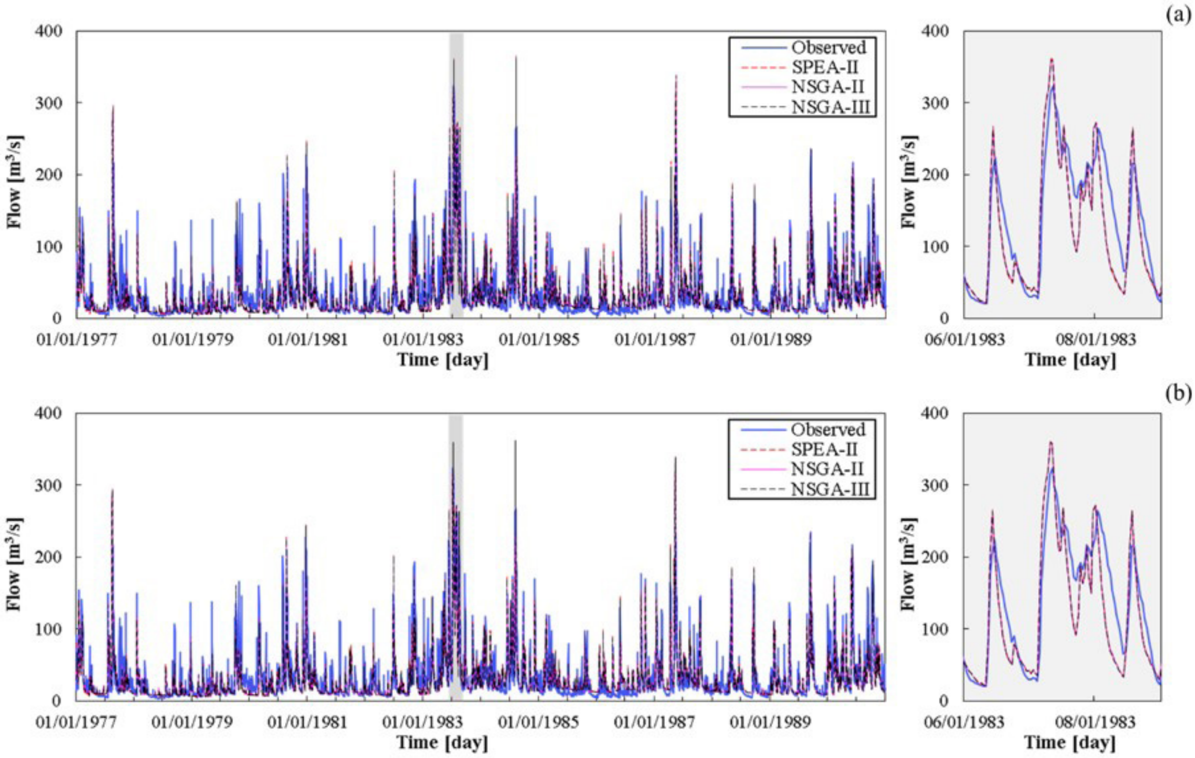

Figure 8. Hydrograph of the calibration period in the Vila Canoas River basin with parameter values defined by (a) the standard stopping criterion and (b) the Ticona stopping criterion. The event of the highest flows was highlighted in the time series and an enlargement view of the event is presented on the right side.

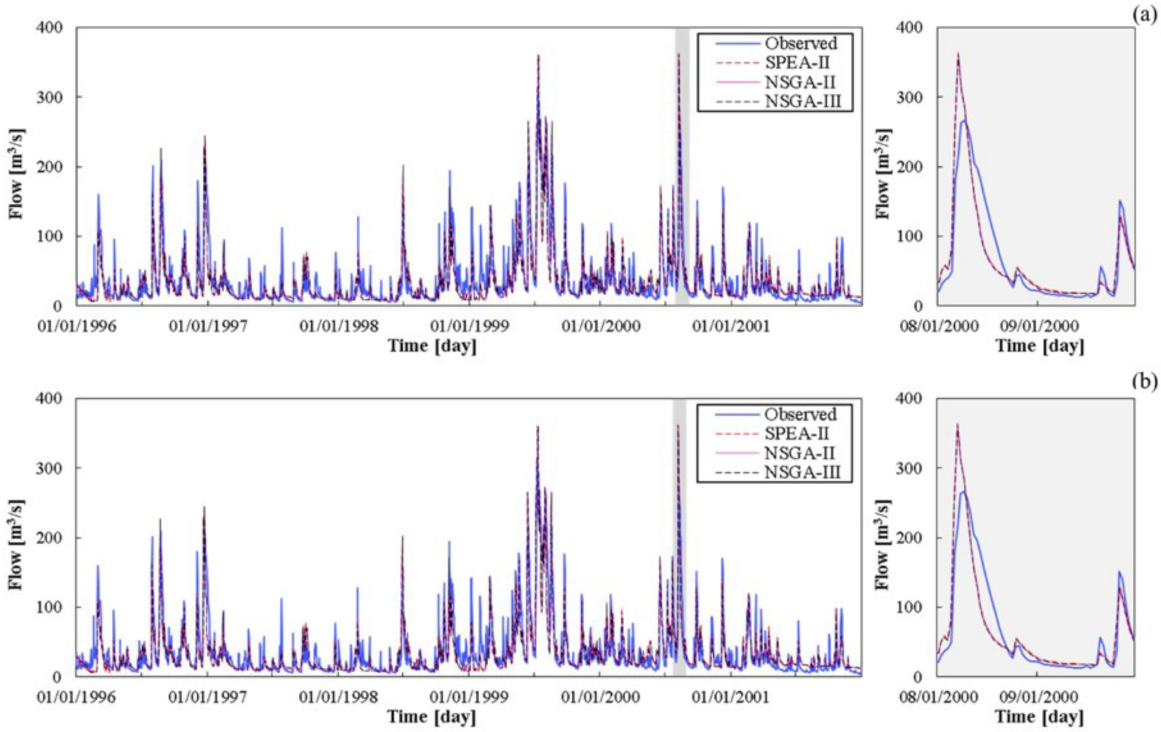

Figure 9. Hydrograph of the validation period in the Vila Canoas River basin with parameter values defined by (a) the standard stopping criterion and (b) the Ticona stopping criterion. The event of the highest flows was highlighted in the time series and an enlargement view of the event is presented on the right side.

CONCLUSIONS

In this study, we performed the multi-objective calibration of the Tank model for two basins using three MOEAs (NSGA-II, NSGA-III, and SPEA-II) and we compared two stopping criteria: the standard stopping criterion (a maximum number of generations) and the Ticona stopping criterion (Gutierrez et al., 2019a). Firstly, two Tank model structures (3-tank and 4-tank) were tested to identify which would be the most suitable to represent the hydrological processes of the two basins. The 4-tank structure was chosen for the study because it presented a better performance in representing the high and low flows in both basins. In the Tank model calibration, the three MOEAs resulted in better performance in terms of computational time using the Ticona stopping criterion in comparison to the standard stopping criterion in both evaluated basins. Using the Ticona stopping criterion, the calibration of the Tank model required a computational time that was 27.4% to 44.1% lower, compared to the calibration of the Tank model using the standard stopping criterion.

The Tank model calibration using both stopping criteria resulted in similar parameter values. A good fit between the simulated and observed streamflow values was obtained during calibration and validation in both basins. We have shown that applying the Ticona stopping criterion reduces computational time while maintaining the quality of the obtained results. For future studies, it would be interesting to assess the Ticona stopping criterion against another stopping criterion, and in the calibration of hydrologic models with a higher number of parameters.

DATA AVAILABILITY STATEMENT

Some or all data, models, and codes generated or used during the study are available from the corresponding author upon request.

REFERENCES

Adeyeri, O. E., Laux, P., Arnault, J., Lawin, A. E., & Kunstmann, H. (2020). Conceptual hydrological model calibration using multi-objective optimization techniques over the transboundary Komadugu-Yobe basin, Lake Chad Area, West Africa. *Journal of Hydrology: Regional Studies*, *27*, pp. 100655.

Araújo, D. R. B., Bastos-Filho, C. J. A., Barboza, E. A., Chaves, D. A. R., & Martins-Filho, J. F. (2011). A performance comparison of multi-objective optimization evolutionary algorithms for all-optical networks design. In *Proceedings of 2011 IEEE Symposium on Computational Intelligence in Multicriteria Decision-Making (MDCM)* (pp. 89-96). USA: IEEE.

Beven, K. (1993). Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, *16*, 41-51.

Beven, K. (2001). *Rainfall-Runoff modelling: the primer.* Hoboken: John Wiley & Sons.

Beven, K. J. (2019). How to make advances in hydrological modelling. *Hydrology Research*, *50*(6), 1481-1494. http://dx.doi.org/10.2166/nh.2019.134.

Beven, K. J., & Smith, P. J. (2015). Concepts of information content and likelihood in parameter calibration for hydrological simulation models. *Journal of Hydrologic Engineering*, *20*(1), 1-15. http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000991.

Chugh, T., Sindhya, K., Hakanen, J., & Miettinen, K. (2019). A survey on handling computationally expensive multiobjective optimization problems with evolutionary algorithms. *Soft Computing*, *23*(9), 3137-3166.

Coello, C. A. C., Lamont, G. B., & Van Veldhuizen, D. A. (2007). *Evolutionary Algorithms for Solving Multi-Objective Problems.* Springer, New York.

Deb, K., & Jain, H. (2014). An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. *IEEE Transactions on Evolutionary Computation*, *18*(4), 602-622.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. M. T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, *6*(2), 182-197.

Echevarría, Y., Sánchez, L., & Blanco, C. (2016). Assessment of multi-objective optimization algorithms for parametric identification of a li-ion battery model. In *Proceedings of International Conference on Hybrid Artificial Intelligence Systems* (pp. 250-260). http://dx.doi.org/10.1007/978-3-319-32034-2_21

Fiben, A. E., & Smith, J. E. (2003). *Introduction to Evolutionary Computing.* Berlin: Springer.

Garcia, F., Folton, N., & Oudin, L. (2017). Which objective function to calibrate rainfall-runoff models for low-flow index simulations? *Hydrological Sciences Journal*, *62*(7), 1149-1166.

Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1998). Toward improved calibration of hydrological models: multiple and noncommensurable measures of information. *Water Resources Research*, *34*(4), 751-763.

Gutierrez, J. C. T., Adamatti, D. S., & Bravo, J. M. (2019a). A new stopping criterion for multi-objective evolutionary algorithms: application in the calibration of a hydrologic model. *Computational Geosciences*, *23*, 1219-1235.

Gutierrez, J. C. T., Vanelli, F. M., & Bravo, J. M. (2019b). Aplicação de um novo critério de parada para algoritmos evolucionários de otimização multi-objetivo na calibração automática de modelos hidrológicos. In *23° Simpósio Brasileiro de Recursos Hídricos*. Foz do Iguaçu, PR.

Haan, C. T. (1989). Parametric uncertainty in hydrologic modeling. *Transactions of the ASAE*, *32*(1), 0137-0146.

Ishibuchi, H., Masuda, H., Tanigaki, Y., & Nojima, Y. (2014). Difficulties in specifying reference points to calculate the inverted generational distance for many-objective optimization problems. In *Proceedings of Symposium on Computational Intelligence in Multi-Criteria Decision-Making (MCDM)* (pp. 170-177). USA: IEEE.

Ishibuchi, H., Masuda, H., Tanigaki, Y., & Nojima, Y. (2015). Modified distance calculation in generational distance and inverted generational distance. In A. Gaspar-Cunha, C. Henggeler Antunes, & C. Coello (Eds.), *Evolutionary Multi-Criterion Optimization. EMO 2015.* (vol 9019, Lecture Notes in Computer Science). Cham: Springer. https://doi.org/10.1007/978-3-319-15892-1_8

Ishihara, Y., & Kobatake, S. (1979). Runoff model for flood forecasting. *Bulletin of Disaster Prevention Research Institute*, *29*(1), 27-43.

Jaiswal, R. K., Ali, S., & Bharti, B. (2020). Comparative evaluation of conceptual and physical rainfall-runoff models. *Applied Water Science*, *10*(1), 1-14.

Khu, S.-T., Madsen, H., & Di Pierro, F. (2008). Incorporating multiple observations for distributed hydrologic model calibration: an approach using a multi-objective evolutionary algorithm and clustering. *Advances in Water Resources*, *31*, 1387-1398.

Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, *5*, 89-97.

Le, V. T., Kuo, C.-M., & Yang, T.-C. (2016, November 6-10). Application of non-dominated sorting genetic algorithm in calibration of HBV Rainfall-Runoff Model: a case study of tsengwen reservoir catchment in Southern Taiwan. In *Proceedings of 12th International Conference on Hydroscience & Engineering Hydro-Science & Engineering for Environmental Resilience*. Tainan, Taiwan.

Legates, D. R., & McCabe Junior, G. J. (1999). Evaluating the use of "goodness of- fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, *35*(1), 233-241.

Madsen, H. (2000). Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. *Journal of Hydrology*, *235*, 276-288.

Mostafaie, A., Forootan, E., Safari, A., & Schumacher, M. (2018). Comparing multi-objective optimization techniques to calibrate a conceptual hydrological model using in situ runoff and daily GRACE data. *Computational Geosciences*, *22*(3), 789-814.

Moussa, R., & Chahinian, N. (2008). Comparison of different multi-objective calibration criteria using a conceptual rainfall-runoff model of flood events. *Hydrology and Earth System Sciences*, *13*(4), 519-535.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I-A discussion of principles. *Journal of Hydrology*, *10*(3), 282-290.

Olazar, M. R. Z. (2007). *Algoritmos evolucionários multiobjetivo para alinhamento múltiplo de sequencias biológicas* (Dissertação de mestrado). Universidade Federal do Rio de Janeiro, Rio de Janeiro.

Pushpalatha, R., Perrin, C., Le Moine, N., & Andréassian, V. (2012). A review of efficiency criteria suitable for evaluating low-flow simulations. *Journal of Hydrology*, *420*, 171-182.

Reynolds, J. E., Halldin, S., Xu, C. Y., Seibert, J., & Kauffeldt, A. (2017). Subdaily runoff predictions using parameters calibrated on the basis of data with a daily temporal resolution. *Journal of Hydrology*, *550*, 399-411.

Rouhani, H., Willems, P., Wyseure, G., & Feyen, J. (2007). Parameter estimation in semi-distributed hydrological catchment modelling using a multi-criteria objective function. *Hydrological Processes*, *21*(22), 2998-3008.

Rozos, E., Efstratiadis, A., Nalbantis, I., & Koutsoyiannis, D. (2004). Calibration of a semi-distributed model for conjunctive simulation of surface and groundwater flows. *Hydrological Sciences Journal*, *49*(5), 819-842.

Schott, J. R. (1995). *Fault tolerant design using single and multicriteria genetic algorithm optimization* (M.S. thesis). Massachusetts Institute of Technology, Cambridge, Massachusetts.

Setiawan, B.I., Fukuda, T., & Nakano, Y. (2003). Developing procedures for optimization of tank model's parameters. *Agricultural Engineering International: the CIGR Journal of Scientific Research and Development*, 1-13. Manuscript LW 01 006.

Shafii, M., & De Smedt, F. (2009). Multi-objective calibration of a distributed hydrological model (WetSpa) using a genetic algorithm. *Hydrology and Earth System Sciences*, *13*(11), 2137-2149.

Song, J.-H., Her, Y., Park, J., Lee, K.-D., & Kang, M.-S. (2017). Simulink implementation of a hydrologic model: a tank model case study. *Water (Basel)*, *9*(639)

Sugawara, M. (1961). On the analysis of runoff structure about several japanese rivers. *Japanese Journal of Geophysics*, *2*(4), 1-76.

Sugawara, M. (1972). *Runoff analysis* (275 p.). Kyoritsu-shuppan.

Sugawara, M. (1979). Automatic calibration of the tank model. *Hydrological Sciences Bulletin*, *24*(3), 375-388.

Sugawara, M., & Singh, V. P. (1995). *Computer models of watershed hydrology* (pp. 165-214). USA: Water Resources Publications.

Suryoputro, N., Suhardjono, Soetopo, W., & Suhartanto, E. (2017). Calibration of infiltration parameters on hydrological tank model using runoff coefficient of rational method. *AIP Conference Proceedings*, 1887, 020056.

Tanakamaru, H. (1995). Parameter estimation for the tank model using global optimization. *Transactions of The Japanese Society of Irrigation, Drainage and Reclamation Engineering*, 178, 503-512.

Van Veldhuizen, D. A. (1999). *Multiobjective evolutionary algorithms: classifications, analyses, and new innovations* (Ph.D. Thesis). Graduate

School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, USA.

Vasconcellos, S. M. (2017). *Desenvolvimento de um Índice de Umidade do Solo Derivado da Versão Distribuída do Tank Model* (Dissertação de mestrado). Universidade Federal do Rio Grande do Sul, Porto Alegre.

Yapo, P. O., Gupta, H. V., & Sorooshian, S. (1998). Multiobjective global optimization for hydrologic models. *Journal of Hydrology*, *204*, 83-97.

Yen, G. G., & He, Z. (2013). Performance metric ensemble for multiobjective evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, *18*(1), 131-144.

Yilmaz, K., Vrugt, J., Gupta, H., & Sorooshian, S. (2010). Model calibration in watershed hydrology. In *Proceedings of Advances in Data-Based Approaches for Hydrologic Modeling and Forecasting* (pp. 53-105). World Scientific Publishing Co.

Zitzler, E., Deb, K., & Thiele, L. (2000). Comparison of multiobjective evolutionary algorithms: empirical results. *Evolutionary Computation*, *8*(2), 173-195.

Zitzler, E., Laumanns, M., & Thiele, L. (2001). *SPEA2: Improving the Strength Pareto Evolutionary Algorithm* (No. 103. TIK-report). Zürich, Suíça: ETH Zurich.

Authors contributions

Juan Carlos Ticona Gutierrez: Data Curation, Formal Analysis, Software, Validation, Visualization, Writing – First Redaction, Writing – Review and Editing.

Cassia Brocca Caballero: Data Curation, Formal Analysis, Visualization, Writing – Review and Editing.

Sofia Melo Vasconcellos: Formal Analysis, Software, Validation, Visualization, Writing – First Redaction, Writing – Review and Editing.

Franciele Maria Vanelli: Formal Analysis, Validation, Visualization, Writing – First Redaction, Writing – Review and Editing.

Juan Martín Bravo: Conceptualization, Formal Analysis, Software, Validation, Visualization, Writing – Review and Editing.

Editor in-Chief: Adilson Pinheiro

Associated Editor: Fábio Veríssimo Gonçalves

SUPPLEMENTARY MATERIAL

Supplementary material accompanies this paper.

Table S1. Metrics of the best PF of each MOEA used; with standard and Ticona stopping criterion in the Ijuí River basin.

Figure S1. Hydrograph of the calibration period for the Ijuí River basin using the 3-tank structure.

Figure S2. Hydrograph of the calibration period in the Ijuí River basin with parameter values defined by (a) the standard stopping criterion and (b) the Ticona stopping criterion. The event of the highest flows was highlighted in the time series and an enlargement view of the event is presented on the right side.

Figure S3. Hydrograph of the validation period in the Ijuí River basin with parameter values defined by (a) the standard stopping criterion and (b) the Ticona stopping criterion. The event of the highest flows was highlighted in the time series and an enlargement view of the event is presented on the right side.

This material is available as part of the online article from 10.1590/2318-0331.272220220046