Okkan, U., Gedik, N. (2018), Evaluation of population based evolutionary optimization algorithms in the Conceptual Hydrological Model Calibration. pp. 130-136. In Gastescu, P., Bretcan, P. (edit, 2018), *Water resources and wetlands*, 4th International Conference Water resources and wetlands, 5-9 September 2018, Tulcea (Romania), p.312

Available online at http://www.limnology.ro/wrw2018/proceedings.html

4th International Conference Water resources and wetlands, 5-9 September 2018, Tulcea (Romania)

EVALUATION OF POPULATION BASED EVOLUTIONARY OPTIMIZATIONALGORITHMS IN THE CONCEPTUAL HYDROLOGICAL MODEL CALIBRATION

Umut OKKAN, Nuray GEDİK

Department of Civil Engineering, Hydraulic Division, Balikesir University, Cagis Campus, 10145, Balikesir, Turkey
E-mail: umutokkan@balikesir.edu.tr, ngedik@balikesir.edu.tr

Abstract

The conceptual hydrological models are generally deterministic and lumped, and these models are based on water budget equations and provide the identification of hydrological cycle elements by means of different parameters. The usage of these models in water resources engineering is rather crucial. These models are used in a variety of areas such as explaining rainfall-runoff relationships of the basin, simulating flows for basins without observation, and analyzing the possible effects of climate change on streamflows. The capability of conceptual models in representing rainfall-runoff relationship of a basin depends on accurate estimation of flow at the outlet of the basin. Hence, it is realized by calibrating of parameters related to hydrological model. This process transforms into an optimization problem based upon determining parameters that minimize the errors between model flows and observed flows. As the number of hydrological model parameters increases, it becomes more difficult to perform the optimization process with manual methods. For this reason, the usage of global optimization algorithms can increase the reliability of the predicted parameters. In the study, some population based evolutionary algorithms were chosen from different optimization techniques and they were assessed in the calibration of a five-parameter hydrological model termed dynamic water balance model. The utilized algorithms mimic some nature inspired principles such as swarm intelligence and the survival of the best within the numerical algorithms. In the study, convergence performances of genetic algorithm (GA), particle swarm optimization (PSO), differential evolution algorithm (DEA), invasive weed algorithm (IWA), and artificial bee colony (ABC) algorithm were compared and it was questioned which algorithms could be preferred in the hydrological model calibration.

Keywords: Hydrological modeling, model calibration, population based evolutionary algorithms

1 INTRODUCTION

The conceptual hydrological models are mostly preferred for the quantitative identification of transformation treatment of rainfall into runoff. They can be utilized in filling missing flow records, in assessing the runoff reaction to climate and/or land use changes and in analyzing hydrological drought events (Xu and Singh 1998). These models can conceptualize the hydrological process represented by soil moisture and groundwater storage functions (Tigkas et al. 2016).

In order to enhance the precision of runoff forecasts in the basin scale, a calibration process is needed to identify the conceptual parameters that minimize an error function computed between modeled and observed runoff series (Gupta et al. 1998). The outdated techniques were commonly based upon trial-and-error adjustments, which were termed as manual calibration. Recently, the manual types have been replaced by computer aided automatic calibration methods in which more robust parameter estimations can be provided (Wagener and Wheater 2006). The main necessaries for the application related to automatic calibration process comprise a qualified calibration data, an objective (fitness) function, and an effective optimization algorithm. A test (validation) data is also required to examine the capability of the calibrated hydrological model (Tigkas et al. 2016).

Because of the lack of direct measurement of assigned model parameters and the fact that the solution space can be quite large, the aforementioned calibration process is not a simple task. Additionally, due to large number of local minimums faced in the calibration process of hydrological models, it is stated that the effectiveness and reliability of conventional automatic methods including Newtonian types may not be sufficient (Madsen et al. 2002; Jha et al. 2006). Thus, in recent years, population based evolutionary algorithms have been used to find global minimum in parameter optimization problem of hydrological models that have probably many local minimums. Since these algorithms have a stochastic character, they can scan a wide range of solutions and get rid of traps by switching between the local minimums due to their adaptive capabilities (Madsen et al. 2002; Tigkas et al. 2016).

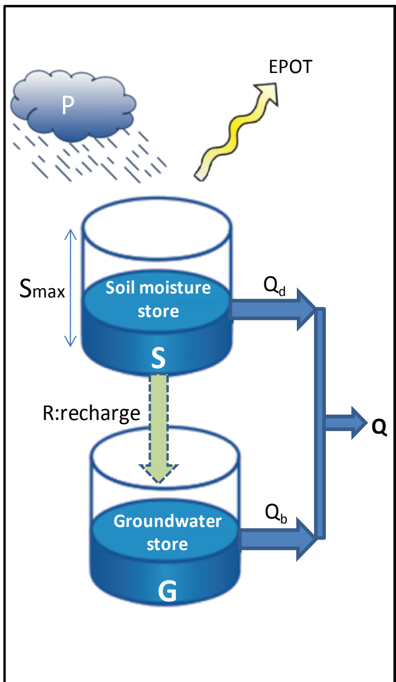
Genetic algorithm (GA) was among the first that has been evaluated in the hydrological model calibration phase (e.g. Wang 1997; Franchini and Galeati 1997; Ndiritu and Daniell 2001; Zhang et al. 2009; Tigkas et al. 2016). After the genetic algorithm, the approaches that were introduced in the literature were also inspired by social behaviors of animal or plant species. While the parameter solution in GA is represented by “chromosomes”, it is represented by “particles or food resources” in particle swarm optimization (PSO), “individuals or chromosomes” in differential evolution algorithm (DEA), “weeds” in invasive weed algorithm (IWA), and “food resources” in artificial bee colony (ABC) algorithm. Of course, it should be mentioned that evolutionary algorithms are not limited to the algorithm examples given above. A broad literature review regarding this issue, as well as several algorithm implementations in the computationally intensive hydrologic models, were presented by Zhang et al. (2009) and Tigkas et al. (2016).

The present work compares the performances of five population based evolutionary algorithms GA, PSO, DEA, IWA and ABC in a challenging automatic calibration task of a five-parameter monthly hydrological model dynamic water budget model (DYNWBM) which is conceptual, deterministic and lumped. The results obtained from the comparative study are thought to contribute to the hydrology literature.

2 MATERIALS AND METHODS

2.1 Study area description

The effectiveness of optimization techniques is linked with the characteristics of the fitness function response surface of used model, which is related to the basin properties (Duan et al. 1992; Zhang et al. 2009). To assess the performance of GA, PSO, DEA, IWA and ABC algorithms, a watershed at eastern region of Turkey was selected. It covers the Gordes watershed which is named by one of the prominent tributaries of Gediz River and it is located at Aegean Region of Turkey. The study region has typical Mediterranean climate characteristics. Gordes watershed has a drainage area of 1070 km^2 and annual mean runoff values obtained from Hacıhıdır flow gauging station is 10.78 mm/month. Potential evapotranspirations (EPOT) were estimated with Hamon’s empirical equation. To achieve detailed information about the study area, readers can look through the work presented by Okkan and Kirdemir (2016).



$$X(t) = P(t)\left[1 + \frac{S_{max} - S(t-1) + EPOT(t)}{P(t)} - \left[1 + \left(\frac{S_{max} - S(t-1) + EPOT(t)}{P(t)}\right)^{\frac{1}{1-\alpha_1}}\right]^{1-\alpha_1}\right]$$

$$Q_d(t) = P(t) - X(t)$$

$$W(t) = X(t) + S(t-1)$$

$$Y(t) = W(t)\left[1 + \frac{EPOT(t) + S_{max}}{W(t)} - \left[1 + \left(\frac{EPOT(t) + S_{max}}{W(t)}\right)^{\frac{1}{1-\alpha_2}}\right]^{1-\alpha_2}\right]$$

$$R(t) = W(t) - Y(t)$$

$$E(t) = W(t)\left[1 + \frac{EPOT(t)}{W(t)} - \left[1 + \left(\frac{EPOT(t)}{W(t)}\right)^{\frac{1}{1-\alpha_2}}\right]^{1-\alpha_2}\right]$$

$$S(t) = \max(0, Y(t) - E(t))$$

$$Q_b = dG(t-1)$$

$$G(t) = \max\left[0, (G(t-1) + R(t))\xi - Q_b(t)\right]$$

$$Q_m(t) = Q_d(t) + Q_b(t)$$

DYNWBM model descriptions:
- •P is shared by Q_d direct flow and X the amount of water retained by the basin for other water budget elements.
- •W: amount of water available in the basin
- •Y:sum of actual evapotranspiration (E) and soil moisture (S) values
- •R: water drained into the groundwater store system (recharge)
- •Q_b is the baseflow

Parameters:
S_{max}:Maximum soil moisture capacity
α_1: The parameter controlling the retention for the water balance components except for direct runoff
α_2:Evapotranspiration efficiency parameter
d: Baseflow parameter
ξ: Groundwater store parameter

Figure 1. Brief information about DYNWBM hydrological model (Okkan and Kirdemir 2018)

2.2 DYNWBM description

In the study presented, DYNWBM developed by Zhang et al. (2008) was exposed to optimization algorithms. It is based upon the Budyko's approach, which relies on interrelationships between water supply and atmospheric demand in the current hydrological cycle, and takes into account hydrological processes such as soil moisture storage, evapotranspiration, and groundwater storage. Monthly total precipitation and EPOT derived from Hamon's formula are used as inputs in the model. The original model with four parameters was modified by Okkan and Kirdemir (2018). In DYNWBM, the sum of the direct and base flow outputs gives the Q_m modeled total flow. Both required definitions and calculation steps for the DYNWBM model are summarized in Fig. 1.

2.3 Population based algorithms used in the study

Due to the page limit, the algorithms used in this study are briefly described below.

GA: It is one of the stochastic search and global optimization methods based upon biological selection principles by simulating a problem in the computer environment. This method has been developed by Goldberg (1989) in the light of comprehensive trials. The working mechanism of algorithm is basically dependent on the principle of community change which represents each individual as chromosomes. These operations are performed by some operators such as crossover, mutation, selection.

PSO: It is another population-based heuristic algorithm proposed by Kennedy and Eberhart (2001), inspired by the social behavior of birds. For each particle that is initially randomized, the local best (pbest) is found in each generation (or iteration). The number of pbest in the swarm is equal to the number of particles. After a sufficient iteration, the global best (gbest) solution is determined from the local solutions by means of velocity and position update operators.

DEA: It is also a population-based heuristic optimization tool proposed by Storn and Price (1997), which is similar to GA in terms of its operator names. Through the use of operators, it is tried to improve the solution of the problem during iterative operation. Unlike GA, the process of acquiring a new individual by treating each chromosome in DEA is carried out through three randomly selected chromosomes. This procedure of DEA has shortened the calculation steps of GA and made it possible to code the problem more easily.

IWA: Mehrabian and Lucas (2006) developed IWA algorithm, inspired by weed colonies. Some operators were created in the process of colonization with the goal of mimicking the colonization behavior of weeds and adapting them to the algorithmic structure. These are operators such as the creation of weed populations, the generation of seeds according to their fitness values when the seeds become flowering plants, the random scattering of the produced seeds in the solution space, and then the competitive selecting process.

ABC: This algorithm, improved by Karaboga and Basturk (2008), is a new generation optimization algorithm that simulates the food search behavior of honey bees. The operators in the algorithm are represented by three kinds of bees: worker bees, observer bees and explorer bees. After the worker bees store the nectar from the food sources and turn the bucket, the observer bees look for new sources of food with a certain probability by watching the dancing of the worker bees. Worker bees who consume food resources within a certain number of trials turn into the explorer bees so as to search for new sources. All of these steps continue until the maximum number of cycles is reached.

3 RESULTS

In the study, the first 15 years (1981–1995) of 30-year observations of Hacihidir flow gauging station were used in calibration, and the other part (1996–2010) was used in validation stage. Due to stochastic properties of the examined algorithms, 30 independent runs were carried out for 1981-1995 calibration period. The maximum number of iterations (or generations/cycles) was set to 200. In the study, sum of squared errors (SSE) was selected as fitness function. For all runs conducted, fitness function was initially calculated between random feasible parameter values. These are typically sampled based upon the probable solution space using a uniform distribution. However, to analyze the ability to find the global result from a large solution space, an extreme parameter range was chosen for $S_{\max}$, α_1, α_2, d and ξ parameters defined in the DYNWBM model.

Since the algorithms used in the study are all population based, the contribution of the different population sizes (N_p) to convergence sensitivities of algorithms has been also investigated. For this purpose, algorithms were separately run using 20, 50 and 100 population sizes. (In IWA algorithms, these population sizes represent maximum weed values. In all variants of IWA, five weeds were assigned as starting population). All utilized algorithms were encoded in the MATLAB environment by the first author. After several trials, the values assigned to the variables controlling the operators defined in the algorithm and some modifications made to the original equations were not shared because of page limit. Variations in fitness values obtained for each run under different population sizes have been extracted throughout the generations. Only GA and DEA responses during generations are given as an example because of page constraint (Fig. 2).

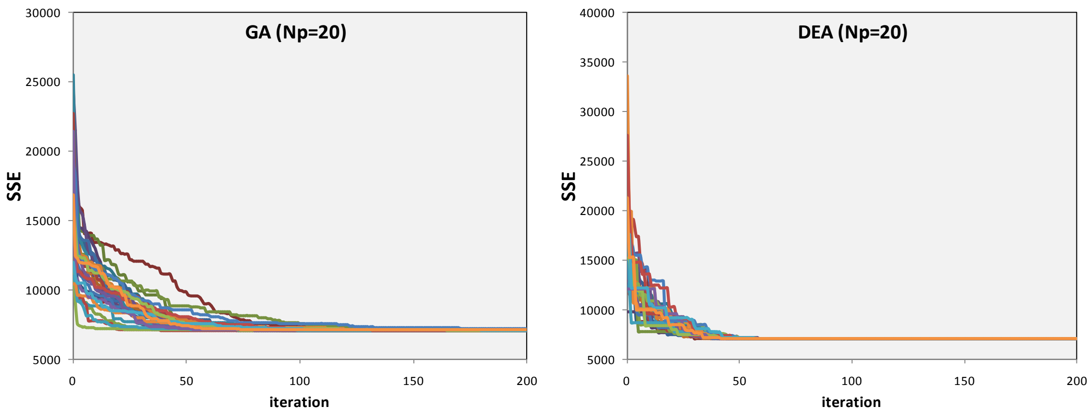


Figure 2. Behavior of GA and DEA algorithms along with iterations

It can be seen from Fig. 2 that the employed algorithms tend to provide a stable global solution after a sufficient cycle even if they start with different random solutions. Since operators of each algorithm have different structures in terms of local and global search capabilities, we have focused on the variability of the solutions in the last iteration in order to reach a general deduction. At this stage, it is useful to display the results obtained at the end of each run with box-plots (Fig. 3).

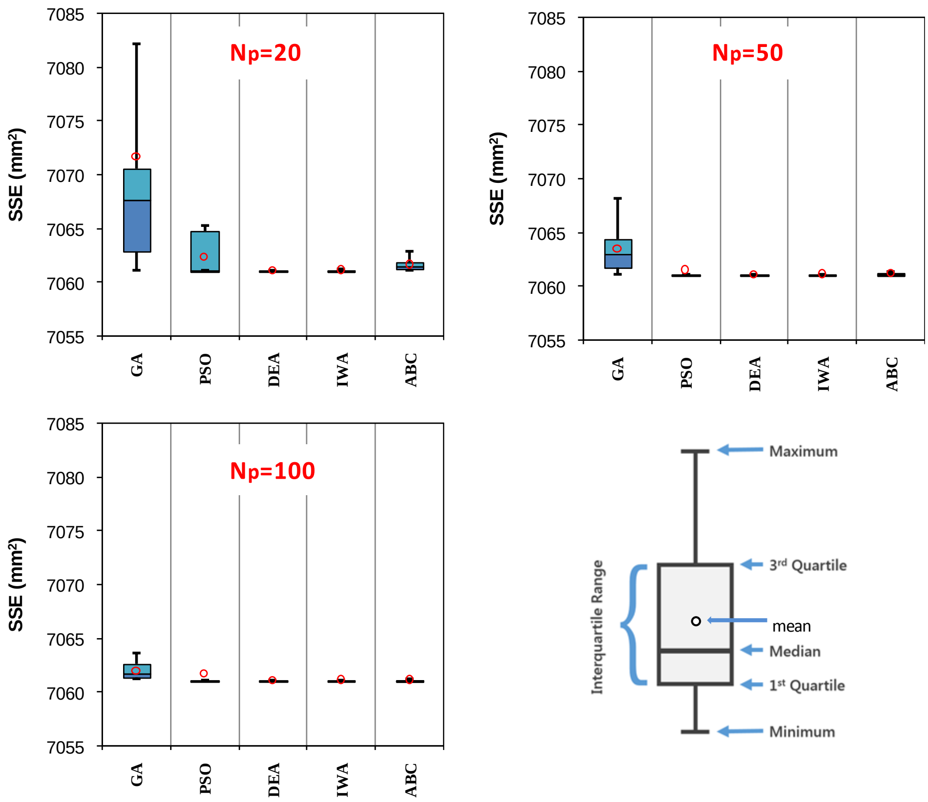


Figure 3. Box-plots offitness function values (30 optimization runs per population size)

As can be seen from Fig. 3, GA does not work well when it run with fewer populations. In the case of processing with 100 populations, 30 runs gave closer results for GA. However, in all variations, GA showed poorer performance than the other algorithms utilized. PSO appears to be more unstable than DEA, IWA and ABC when it is operated with 20 populations. However, when it reached 50 populations, it was quite reasonable. On the other hand, DEA, IWA and ABC algorithms seem to be fairly successful even when they run in low populations. In case of using proficient population size (for example Np≥50), it can be seen that PSO, DEA, IWA and ABC present a more robust performance than GA, as the optimal fitness function values are almost fully clustered around the mean statistics.

In the population based evolutionary algorithms, it is also substantial to make out how rapid they converge to the desired optimal point per iteration or their average convergence ratios (CRs). In this study, a geometric convergence rate formula proposed by He and Lin (2016) was considered. The sufficient iteration number is also derived from the implementation of this formula (see Figure 4).

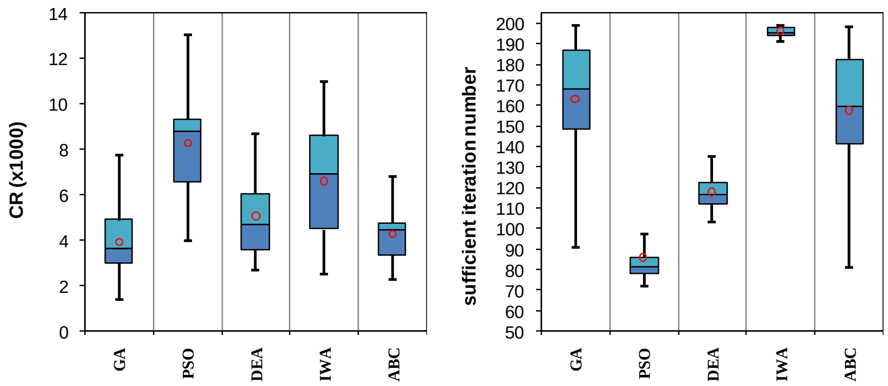


Figure 4. For the case N_p=50, distributions about the CRs provided by the different methods over 200 iterations and the required iteration numbers (the distributions were obtained over 30 runs)

When both the convergence speed and the required iteration numbers are also evaluated, PSO and DEA are more successful than the others in case of assigning a proper population size. If a general assessment is made, the PSO performs faster convergence performance with a sufficient population coming up to 50, while the DEA algorithm yields reasonable results in low populations like N_p=20.

Table 1. Summary of the results obtained from DEA operated by 20 populations

(a) Calibrated DYNWBM parameters

S_{max}	α_1	α_2	d	ξ
294.0106	0.6347	0.7371	0.9155	0.6223

(b) Calibration statistics

RMSE (mm)	R^2	NS	RSR	PBIAS
6.2633	0.8905	0.8894	0.3326	3.1049

(c) Validation statistics

RMSE (mm)	R^2	NS	RSR	PBIAS
7.9635	0.8719	0.8516	0.3853	14.3079

In the context of being an example, the parameters obtained by the DEA algorithm operated with 20 populations and the derived statistics for calibration-validation periods are presented in Table 1. Calibration and validation period predictions related the parameters found with DEA are also given in Fig. 5 as time series plot. As determined here, the runoff predictions produced by DYNWBM calibrated by DEA matched successfully with the observed values (Q_O). According to Table 1, the evaluation indices including NS, RSR, PBIAS revealed a satisfactory response of the model for both calibration and validation periods.

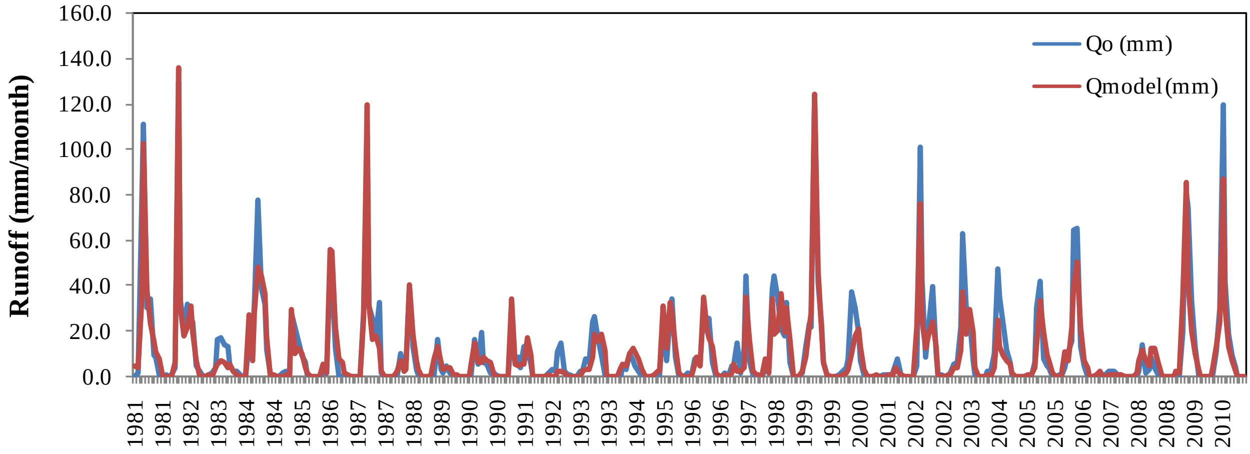


Figure 5. DYNWBM runoff predictions for Gordes watershed

4 CONCLUSIONS

In the study, five population-based heuristic algorithms used for benchmark functions in the literature were selected and applied on DYNWBM, an outstanding hydrological model. Some of the deductions derived from the study are:

- If a familiar global optimization algorithm such as GA is used, it is recommended not to use the low populations. For example, as the number of parameters of the hydrological model increases, the success of GA will be proportional to the population size.
- PSO is a very practical and fast technique in terms of variables controlling the algorithm. Unfortunately, it does not work sensitively with a small population.
- DEA, IWA and ABC are able to provide more effective results in low populations compared to GA and PSO. In particular, DEA and PSO have been found to be very successful in terms of fast convergence and stability in parameters compared to other algorithms in case of using enough population sizes.
- It should be clarified that the responses of employed algorithms have been collected from a small watershed in Turkey. Further implementations of the explained procedure subjected to the study, for other rivers with different climatic characteristics, will provide better generalization. Regarding the other optimization algorithms (for example hybrid techniques) and the other fitness functions will shed light on the problem as well.

ACKNOWLEDGMENTS

This study was funded by “Scientific Research Projects Unit of Balikesir University” under Grant No.2017/134.

REFERENCES

Duan, Q., Sorooshian, S., and Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resources Research*, **28**(4), 1015-1031. Doi:10.1029/91WR02985

Franchini, M., and Galeati, G. (1997). Comparing several genetic algorithm schemes for the calibration of conceptual rainfall-runoff models, *Hydrological Sciences Journal*, **42**(3), 357-379. Doi:10.1080/02626669709492034

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley: Reading, MA.

Gupta, H. V., Sorooshian, S., and Yapo, P. O. (1998). Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resources Research*, **34**(4), 751-763. Doi:10.1029/97WR03495

He, J., and Lin, G. (2016). Average convergence rate of evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, **20**(2), 316-321. Doi: 10.1109/TEVC.2015.2444793

Jha, M. K., Kumar, A., Nanda, G., and Bhatt, G. (2006). Evaluation of traditional and nontraditional optimization techniques for determining well parameters from step-drawdown test data, *Journal of HydrologicEngineering,* **11**(6), 617-630. Doi:10.1061/(ASCE)1084-0699(2006)11:6(617)

Karaboga, D., and Basturk, B. (2008). On the performance of artificial bee colony (ABC) algorithm, *Applied Soft Computing*, **8**, 687-697, 2008. Doi:10.1016/j.asoc.2007.05.007

Kennedy, J., and Eberhart, R.C. (1995). Particle Swarm Optimization, In Proceedings of IEEE International Conference on Neural Networks, 4, 1942-1948, Perth, Australia, IEEE Service Center, Piscataway, NJ.

Madsen, H., Wilson, G., and Ammentorp, H. C. (2002). Comparison of different automated strategies for calibration of rainfall-runoff models, *Journal of Hydrology*, **261**(1-4), 48-59. Doi:10.1016/S0022-1694(01)00619-9

Mehrabian, A., and Lucas, C. (2006). A novel numerical optimization algorithm inspired from weed colonization, *Ecological Informatics*, **1**(4), 355-366. Doi: 10.1016/j.ecoinf.2006.07.003

Ndiritu, J., and Daniell, T. (2001). An improved genetic algorithm for rainfall-runoff model calibration and function optimization, *Mathematical and ComputerModelling*, **33**(6-7), 695-706. Doi: 10.1016/S0895-7177(00)00273-9

Okkan, U., and Kirdemir, U. (2016). Investigation of a dam reservoir behavior under climate change scenarios of IPCC-AR5, pp. 769-779. In *Proceeding of 3rd International Scientific Meeting: State and Trends of Civil and Enviromental Engineering – EGTZ 2016, Editors: Adnan Ibrahimović and Damir Zenunović, 2– 4 June 2016, Bosnia, Tuzla.*

Okkan, U., and Kirdemir, U. (2018). Investigation of the behavior of an agricultural-operated dam reservoir under RCP scenarios of AR5-IPCC, *Water Resources Management*, **32**(8), 2847–2866. Doi: 10.1007/s11269-018-1962-0

Storn, R., and Price, K. (1997). Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces, *Journal of Global Optimization*, **11**, 341–359. Doi: 10.1023/A:1008202821328

Tigkas, D., Christelis, V., and Tsakiris, G. (2016). Comparative study of evolutionary algorithms for the automatic calibration of the Medbasin-D conceptual hydrological model, *Environmental Processes*, **3**(3), 629-644. Doi: 10.1007/s40710-016-0147-1

Wagener, T., and Wheater, H. S. (2006). Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty, *Journal of Hydrology*, **320**(1-2), 132-154. Doi: 10.1016/j.jhydrol.2005.07.015

Wang, Q. (1997). Using genetic algorithms to optimise model parameters, *Environmental Modelling & Software*, **12**(1), 27-34. Doi: 10.1016/S1364-8152(96)00030-8

Xu, C.Y., and Singh, V.P. (1998). A review on monthly water balance models for water resources investigation and climatic impact assessment, *Water Resources Management*, **12**, 31–50. Doi: 10.1023/A:100791681646

Zhang, L., Potter, N., Hickel, K., Zhang, Y., and Shao, Q. (2008). Water balance modeling over variable time scales based on the Budyko framework – Model development and testing, *Journal of Hydrology*, **360**(1-4), 117-131. Doi: 10.1016/j.jhydrol.2008.07.021

Zhang, X., Srinivasan, R., Zhao, K., and Liew, M. V. (2009). Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model, *Hydrological Processes*, **23**(3), 430-441. Doi: 10.1002/hyp.7152