The impact of choice of the optimizer on the performance of catchment runoff models

J.J. Napiorkowski, A.P. Piotrowski, E. Karamuz, T.B. Senbeta

Institute of Geophysics Polish Academy of Sciences, Warsaw

Introduction

The performance of conceptual catchment runoff models may highly depend on specific methodological choices made by the user. In this paper we show how the performance of two different models, HBV [1] and GR4J [2], is affected by the choice of the optimizer, the method of Average Areal Rainfall estimation and computational period.

Conceptual catchment runoff models

This paper considers two lumped conceptual models, that are built of interconnected reservoirs with mathematical transfer functions used to describe the transfer of water between reservoirs and into the river. The input variables to both models are daily precipitation totals, mean air temperature and estimated potential evapotranspiration. Precipitation may occur in the form of rainfall, snowfall or a mixture of snowfall and rainfall.

HBV model

The HBV model, introduced by Bergström [2], is a standard tool for runoff simulations and flood forecasting. Details of the version adopted in this paper can be found in [3]. The model has five state variables representing storage of snow pack, snowmelt water, soil moisture, fast runoff and base flow. The model requires calibration of 13 parameters.

GR4J model

The detailed mathematical description of the GR4J model may be found in [4]. Since our study is concerned with Polish climatic conditions, the original model is extended by adding a snow module similar to that used in HBV model. This extended version of GR4J has seven parameters, namely three parameters in the snow routine and four original parameters representing maximum capacity of production store, groundwater exchange coefficient, one-day-ahead maximum capacity of routing store and time base of unit hydrograph. However, we believe that the introduction of the name GR7J may be confusing.

Average Areal Precipitation estimation

Areal precipitation estimation is to be determined in three ways: Thiessen Polygon approach, joint optimization of conceptual rainfall-runoff model parameters, and E-OBS gridded dataset.

The choice of the optimizer

This paper focus on direct comparison between two families of optimization algorithms: Particle Swarm Optimization and Differential Evolution, for conceptual rainfall-runoff model calibration. After quarter century of research, one could find hundreds of DE and PSO variants in the literature ([5], [6]) . Various variants of DE and PSO may highly differ from, and are often much more complicated than the basic versions of these algorithms. In this study we point at mod-ern DE and PSO variants, instead of their historical, simple versions. Five relatively recently proposed DE variants, and five PSO variants were selected for calibration of HBV and GR4J models. All tested variants were proposed in the main computer science journals or proceedings from the leading conferences in Evolutionary Computation.

Kamienna catchment.

Tests are based on the Kamienna catchment located in the central Poland that have relatively complicated orography, and is composed of lowland and low-mountainous part (A = 2007,9 km², L= 138 km, Q_{avr} = 10 m³/s).

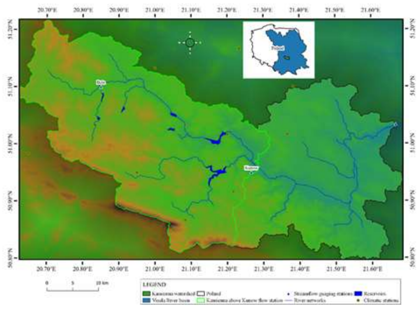

Table 1. Compared variants of Differential Evolution and Particle Swarm Optimization algorithms. In comments, D means problem dimensionality, [MINi, MAXi] are bounds set for i-th dimension. Algorithms 1-5 are Differential Evolution-based, 6-10 are Particle Swarm Optimization-based

Alg.	Short name	Descriptive name	Reference	Year	Comments
1	HARD-DE	Hierarchical archive-based DE	Meng and Pan 2019	2019	Population size is parabolically (quicker at the end of the search) decreased during run from $25\ln(D)D^{0.5}$ to 4. Algorithm performed best on calibration of other hydrological problem – air2water model calibration (Zhu et al. 2021).
2	MDE_pBX	Memetic adaptive DE with new mutation and crossover	Islam et al. 2012	2012	Population size = 100. One among older classical adaptive DE variants.
3	L-SHADE	SHADE with linear population size reduction	Tanabe and Fukunaga 2014	2014	A version of Successful History Adaptive DE with population size linearly reduced during run from $18D$ at the beginning to 4 at the end. State-of-the-art algorithm that was a kick-off point for many other Differential Evolution variants.
4	OLSHADE-CS	DE orthogonal array-based initialization and new selection strategy	Kumar et al. 2022	2022	Algorithm proposes a new DE selection technique and a new initialization method. The population size is decreasing linearly from $6D^2$ to 4 during the search.
5	EnsDE	Ensemble of DE algorithms	Wu et al. 2018	2018	Population size = 100. State-of-the-art ensemble of DE variants.
6	DEPSO	Dual environmental PSO	Zhang et al 2019	2019	Population size = 50. Velocities are initialized randomly within $[MIN_i,MAX_i]$ interval.
7	EPSO	Ensemble of PSO variants	Lynn and Sugan-than 2017	2017	Population size = 40, divided into two uneven swarms. Population size = 50. Velocities are initialized randomly within $[MIN_i,MAX_i]$ interval. Classical algorithm that manages an ensemble of different PSO variants.
8	PPSO	Pyramid PSO	Li et al 2022	2022	Population size = 64. An algorithm builds a 4-layer pyramid with number of particles from the top to the bottom layer set to 4, 8, 20, 32. Some particles in each layer may communicate solely with particles from the same layer, others may communicate with particles from the same layer and from the one upper layer. Velocities are initialized with $\pm 0.1 \cdot [MIN_i-MAX_i]$.
9	PSO-sono	PSO for single-objective problems	Meng et al. 2022	2022	Population size = 100. A very recent PSO variant that uses a ring topology. Velocities are initialized randomly within [-30,30] interval.
10	TAPSO	Triple-archive PSO	Xia et al. 2020	2020	Population size = 60. Velocities are initialized randomly within $[MIN_i,MAX_i]$ interval, but during run are effectively restricted to $\pm 0.2 \cdot [MIN_i-MAX_i]$.

Table 2. MSE results obtained for HBV model

Measure	HARD-DE	MDE-pBX	L-SHADE	OL-SHADE-CS	EnsDE	DEPSO	EPSO	PPSO	PSO-sono	TAPSO
Calibration set										
Mean	14.154	14.504	14.666	20.242	14.624	15.200	14.647	14.088	15.727	**15.041**
Rank mean	2	3	6	10	4	8	5	1	9	**7**
Median	14.505	14.505	14.505	20.365	14.603	14.926	14.629	14.039	16.066	**15.065**
Rank median	2	2	2	10	5	7	6	1	9	**8**
Best	12.158	14.487	14.487	18.389	14.505	14.506	14.303	13.042	14.572	**13.270**
Rank best	1	5	5	10	7	8	4	2	9	**3**
Worst	14.505	14.585	16.294	22.327	14.892	16.295	14.983	16.161	16.349	**16.008**
Rank worst	1	2	7	10	3	8	4	6	9	**5**
Standard deviation	0.811	0.017	0.481	0.939	0.086	0.608	0.163	0.988	0.710	**0.607**
Validation set										
Mean	39.314	40.365	40.450	46.087	42.072	38.832	40.600	37.196	40.339	**39.410**
Rank mean	3	6	7	10	9	2	8	1	5	**4**
Median	40.026	40.027	40.025	44.484	41.886	39.367	39.937	37.118	39.974	**39.349**
Rank median	7	8	6	10	9	3	4	1	5	**2**
Best	32.350	39.964	39.354	36.918	37.665	37.076	38.811	35.099	38.859	**37.291**
Rank best	1	10	9	3	6	4	7	2	8	**5**
Worst	41.718	41.738	43.285	60.342	45.759	41.264	45.108	40.014	43.525	**43.710**
Rank worst	3	4	5	10	9	2	8	1	6	**7**
Standard deviation	**2.644**	**0.677**	**0.886**	**5.908**	**1.545**	**1.197**	**1.578**	**1.329**	**1.256**	**1.315**

Table 3. MSE results obtained for GR4J model

measure	HARD-DE	MDE-pBX	L-SHADE	OL-SHADE-CS	EnsDE	DEPSO	EPSO	PPSO	PSO-sono	TAPSO
Calibration set										
Mean	15.819	16.127	16.091	19.909	15.824	16.946	15.979	16.567	16.661	**16.567**
Rank mean	1	5	4	10	2	9	3	6	8	**6**
Median	15.719	15.719	15.719	19.871	15.719	15.828	15.737	15.828	15.815	**15.828**
Rank median	1	1	1	10	1	7	5	7	6	**7**
Best	15.719	15.719	15.719	17.030	15.719	15.719	15.719	15.719	15.719	**15.719**
Rank best	1	1	1	10	1	1	1	1	1	**1**
Worst	18.819	18.821	18.821	21.529	18.819	18.868	18.823	18.874	18.863	**18.874**
Rank worst	1	3	3	10	1	7	5	8	6	**8**
Standard deviation	0.557	1.054	0.932	1.344	0.556	1.521	0.764	1.356	1.423	**1.356**
Validation set										
Mean	30.279	30.279	31.296	38.513	30.240	35.873	30.837	33.913	34.104	**33.913**
Rank mean	2	2	5	10	1	9	4	6	8	**6**
Median	29.759	29.759	29.759	35.996	29.757	30.275	29.759	30.221	29.759	**30.221**
Rank median	2	2	2	10	1	9	2	7	2	**7**
Best	29.759	29.759	29.749	28.874	29.049	28.840	29.079	28.840	28.816	**28.840**
Rank best	8	8	8	5	6	2	7	2	1	**2**
Worst	45.870	45.870	45.816	46.794	45.816	45.518	45.535	45.518	45.871	**45.518**
Rank worst	7	7	5	10	5	1	4	2	9	**2**
Standard deviation	**2.894**	**2.894**	**4.765**	**5.992**	**2.894**	**7.753**	**3.939**	**6.937**	**7.372**	**6.937**

Conclusions

We have compared five DE and five PSO algorithms on calibration of HBV and GR4J conceptual models for rainfall-runoff modelling of the Kamienna catchment. We aimed at finding whether DE or PSO algorithms would be better suited for that task. Each algorithm was run 30 times to obtained a sample of results.

It turned out that the results obtained by different optimizers are roughly similar for GR4J model that has a few parameters. For GR4J model one may point at an inferior algorithm – OLSHADE-CS, rather than a winners, as many optimizers performed very similarly.

In the case of HBV model the results were much different. OLSHADE-CS also performed by far the poorest, but results obtained by other algorithms were diversified. Which method could be termed a winner depends whether one focus on calibration, or validation set, and whether one is interested in mean/ median performance, or in finding the best possible solution in one among 30 runs.

Overall two algorithms, PPSO and HARD-DE, performed best on HBV model calibration, and DE algorithms slightly outperformed PSO ones.

References

[1] Bergstrom, S. (1975). Development of a snow routine for the HBV-2 model, Nordic Hydrology, 6(2), 73-92

[2] Perrin, C., Oudin, L., Andreassian, V., Rojas-Serna1, C., Michel, C., & Mathevet, T. (2007). Impact of limited streamflow data on the efficiency and the parameters of rainfall–runoff models, Hydrological Sciences Journal, 52(1), 131-151.

[3] Piotrowski, A.P., Napiorkowski, M.J., Napiorkowski, J.J., Osuch, M., Kundzewicz, Z.W. (2017). Are modern metaheuristics successful in calibrating simple conceptual rainfall–runoff models? Hydrological Sciences Journal , 62, 4, 606–625.

[4] Perrin, C., Michel, C., Andreassian, V., (2003). Improvement of a parsimonious model for streamflow simulation. Journal of Hydrology, 279, 1-4, 275-289.

[5] Das, S., Mullick, S.S., Suganthan, P.N. (2016) Recent advances in Differential Evolution – an updated survey, Swarm and Evolutionary Computation 27, 1-30

[6] Bonyadi, M.R., Michalewicz, Z. (2017) Particle Swarm Optimization for single objective continuous space problems: A re-view, Evolutionary Computation 25(1), 1-54.