KSCE Journal of Civil Engineering (2012) 16(4):667-675
DOI 10.1007/s12205-012-1076-9

Water Engineering
www.springer.com/12205

Determining Rainfall-Intensity-Duration-Frequency Relationship Using Particle Swarm Optimization

Halil Karahan*

Received December 14, 2009/Revised 1st: June 14, 2011, 2nd: July 19, 2011/Accepted September 7, 2011

Abstract

This study proposes a Particle Swarm Optimization (PSO) algorithm to model the Rainfall-Intensity-Duration-Frequency (RIDF) relationship. The study is carried out under two scenarios. In scenario I, a data set with a length of 50 years is used. In Scenario II, the data set is extended to 68 years by adding the values of the recent 18 years. Scenario I is used for testing the robustness of the proposed PSO-RIDF model. The PSO-RIDF algorithm gives the same objective function value for different runs and this shows that the proposed algorithm is robust. Scenario II is used to investigate the influence of data length on model performance. It has been observed that the proposed PSO-RIDF model gives the same performance results as that of the Genetic Algorithm (GA) according to various error evaluation criteria. The PSO-RIDF model shows better performance than GA formulas when the number of parameters increases. It has also been observed that the length of the data set and the chosen formulation are influential on model performance. The weighting parameters of the RIDF model may be determined with PSO algorithm in one-stage instead of any statistical computations and/or trial-error procedure.

Keywords: *rainfall intensity-duration-frequency, particle swarm optimization, parameter estimation*

1. Introduction

Information regarding rainfall e.g., duration, intensity, distribution with respect to time and area is a prerequisite for planning purposes in agriculture, water resources, urbanization, drainage, flood control and transportation. Many engineering structures in these fields are designed, built and maintained according to available rainfall information. In general, the relationship between duration, intensity and return period of rainfall is called as the intensity-duration-frequency relationship. This relationship is expressed by various statistical and/or mathematical equations. Each form of RIDF takes T, as period; t, as duration; 1/T, as frequency and some weighting parameters, time and location.

The determination of weighting parameters generally requires mathematical transformations and/or statistical analysis. Finding out which distribution represents the observed data requires many trials and/or use of software developed for this purpose. The RIDF relationship is represented with different empirical and statistical forms of equations and they are given in Karahan *et al.* (2007). Their algorithm uses the genetic algorithm technique as an optimization method. The methodology is also applied for the South-eastern Anatolian region. The models took into account the geographical characteristics such as latitude, longitude and elevation. It has been shown that the proposed technique may be used for regional applications (Karahan *et al.*, 2008).

In recent years, Particle Swarm Optimization (PSO) technique, as an heuristic optimization, has been successfully applied in various fields of engineering problems. The PSO technique, first developed by Kennedy and Eberhart, was inspired by the behaviours of bird and fish swarms in situations such as finding food and escaping from danger (Kennedy and Eberhart, 1995; Eberhart and Kennedy, 1995). PSO is used by many researchers for the solution of problems in the fields of reactive power and voltage control (Yoshida *et al.*, 2000), optimum design of power systems (Abido, 2002), short-term hydro-thermal scheduling (Yu *et al.*, 2007), determination of chemical reaction parameters (Schwaab *et al.*, 2008), prediction of hard rock Tunnel Boring Machines (TBM) penetration (Yagiz and Karahan, 2011) and it has also been used as a learning algorithm in Artificial Neural Network (ANN) applications (Chau, 2006, 2007).

Recently, PSO algorithm was applied in the design of water supply systems (Montalvo *et al.*, 2008a), and wastewater collection networks (Izquierdo *et al.*, 2008), but as far as the writer knows, this study may be the first one in which the PSO is applied in the hydrological field.

Like other evolutionary techniques, PSO does not guarantee the global optimum and it does not need specific operators such as crossover and mutation (GA) or pheromone updating (ant colony optimization) because the particles in PSO update themselves with internal velocity (Montalvo *et al.*, 2008b). A critic about the PSO algorithm is the lack of population diversity. To avoid this problem, a method that regenerates the particles that

*Professor, Dept. of Civil Engineering, Pamukkale University, Denizli 20017, Turkey (E-mail: hkarahan@pamukkale.edu.tr)

occupy the leader's position has been proposed by (Montalvo *et al*., 2008b). It is reported that this regeneration avoids premature convergence since it prevents clone populations from dominating the search. In this study, Linear Fitness Scaling (LFS) approach is proposed for the conservation of population diversity. Thanks to LFS, the difference between the best and the worst objective function values in the swarm is controlled and if the difference is smaller than the allowed value, the particle with the worst objective function in the swarm is replaced with a randomly selected one. With this approach, the diversity of the swarm is conserved and the PSO algorithm is enabled to reach a global or near-global optimum.

Results show that the easy programmable and faster PSO technique (especially in obtaining the regional RIDF relationships requiring longer computation duration) is an effective solution tool for obtaining the best fitting intensity-duration relationship to the observed data.

2. Formulation of Intensity-Duration-Frequency Relations

The forms of the RIDF formula are explicitly derived from the emprical and statistical distributions. The emprical models are given as follows:

2.1 Emprical Formulation

We can formulate a generalised RIDF relationship in the form $I = a(T)/b(t)$ which has the advantage of a separable functional dependence of I on T and t. Where: I is the rainfall intensity (mm /min), T is the return period, t is the duration (Bernard, 1932; Chen, 1983; Kothyari and Garde, 1992; Koutsoyiannis *et al*., 1998; Karahan *et al*., 2007).

$$I = \frac{w_0 T^{w_1}}{t^{w_2}} \tag{1}$$

$$I = \frac{w_0 T^{w_1}}{(w_2 + t^{w_3})^{w_4}} \tag{2}$$

where: w_k (k = 0, 1, 2 for Eq. (1); k = 0,1,…,4 for Eq. (2)) represents the weighting coefficients.

Two empirical formulations proposed by (Karahan *et al*., 2007), which are also used in this study are:

$$I = \frac{\sum_{i=0}^{M} w_i[\ln(T)]^i}{(w_{M+1} + t^{M+2})^{M+3}} \tag{3}$$

$$I = \frac{\sum_{i=0}^{M} w_i[\ln(T)]^i}{\sum_{i=0}^{M} w_{(M+1)+i}[\ln(t)]^i} \tag{4}$$

where M is the rank of the equation and M = 4.

The statistical distribution models are given as follows:

2.2. Statistical Formulation

Commonly used statistical distributions to obtain RIDF relationship are: *Gumbel, Generalized Extreme Value* (GEV)*, Exponential, Pareto* and it has been shown that the RIDF problem fits better to Gumbel and GEV distributions (Koutsoyiannis *et al*., 1998). The formula of the Gumbel distribution is:

$$I = w_0 \frac{w_1 - \ln\left[-\ln\left(1 - \frac{1}{T}\right)\right]}{(w_2 + t^{w_3})^{w_4}} \tag{5}$$

and the GEV distribution is defined as:

$$I = \frac{w_0\left\{w_1 + \left[-\ln\left(1 - \frac{1}{T}\right)\right]^{-w_2}\right\}}{(w_3 + t^{w_4})^{w_5}} \tag{6}$$

The other two RIDF formulations which are less used than the distributions previously mentioned are Exponential and Pareto distributions (Koutsoyiannis *et al*., 1998).

The Exponential distribution is defined as:

$$I = \frac{w_0(w_1 + \ln(T))}{(w_2 + t^{w_3})^{w_4}} \tag{7}$$

and the Pareto distribution is defined as:

$$I = \frac{w_0(w_1 + T^{w_2})}{(w_3 + t^{w_4})^{w_5}} \tag{8}$$

3. Model Application

3.1 Particle Swarm Optimization Algorithm

The PSO technique was first developed by Kennedy and Eberhart (1995), inspired by the behaviors of bird and fish swarms in a multi-dimensional space in situations like finding food and escaping from danger (Kennedy and Eberhart, 1995; Eberhart and Kennedy, 1995). Each individual taking place in the PSO algorithm is called a particle, and the population generated by these particles is called swarm. In the PSO algorithm, each particle is evaluated as a candidate for the solution of the problem. The particle and swarm terms in the PSO algorithm account for the chromosome and population terms as in the GA (Goldberg, 1989).

PSO is an iterative solution method related to the search behaviours of swarm particles in a multi-dimensional search space. The dimension of the search space is equal to the number of the unknown variables of the investigated problem and the number of particles in the swarm shows the volume of the swarm.

Before solving the problem, the maximum and minimum possible values of the variables (x_d^{Max} and x_d^{Min}) are chosen according to the range of the possible values of the variable. Using these values, the maximum and minimum velocities are calculated as follows:

$$v_d^{Max} = (x_d^{Max} - x_d^{Min})/2 \quad (9a)$$

$$v_d^{Min} = -v_d^{Max} \quad (9b)$$

In this equation, the values of x_d^{Max} and x_d^{Min} are chosen according to the range of the variable and the initial locations and velocities of the particles that are randomly determined according to Eqs. (10) and (11).

$$x_{p,d}^k = x_d^{Min} + r(x_d^{Max} - x_d^{Min}) \quad (10)$$

$$v_{p,d}^k = v_d^{Max}(2r - 1) \quad (11)$$

Where, p is the particle number, d is the search direction, v is the particle velocity, x is the location of particle and r is a randomly generated number fitting to uniform distribution in the range [0,1]. According to these initial values, the suitability value of each particle is calculated according to the objective function defined for the problem and the location of the best particle is determined. Until the location and velocity values meet the stopping criteria, each particle updates its own location according to previous experience and the location of the best particle in the swarm.

$$v_{p,d}^{k+1} = \omega v_{p,d}^k + c_1 r_1 (x_{p,d}^{ind} - x_{p,d}^k) + c_2 r_2 (x_d^{glo} - x_{p,d}^k) \quad (12)$$

$$x_{p,d}^{k+1} = x_{p,d}^k + v_{p,d}^{k+1} \quad (13)$$

In Eqs. (12) and (13), k represents the iteration number. ω, c_1 and c_2 are search parameters, r_1 and r_2 are two random numbers with a uniform distribution in the range [0,1]. $x_{p,d}^{ind}$ is the best position found by the particle itself, while x_d^{glo} is the best position found by the whole swarm. The parameters c_1 and c_2 are the cognition and the social parameters (Kennedy and Eberhart, 1995; Eberhart and Kennedy, 1995). The parameter ω is called inertial weight and was not present in the original form of the algorithm. Originally, the ω coefficient is proposed to be 1 by Kennedy and Eberhart in the PSO algorithm. This parameter is effective on model performance and it is proposed by (Shi and Eberhart, 1998) in a linearly decreasing form as in Eq. (14). It has been used by many researchers in this form (Shi and Eberhart, 1998; Clerc, 1999; Yagiz and Karahan, 2011).

$$\omega = \omega_{\max} - (\omega_{\max} - \omega_{\min})\frac{k}{k_{\max}} \quad (14)$$

In Eq. (14), $k_{\max}$ is the maximum number of iterations and k is the current number of iterations.

One of the most serious critiques made to the PSO algorithm is the lack of diversity with the progress of iteration. This lack of diversity can be understood as the superposition of particles in the solution domain, getting closer each time to a leader. In depending on the initial position of particles and the lack of diversity that the algorithm may experiment, particles can converge to some local optimum, losing the possibilities to explore other regions of the solution domain. After a number of trials, it has been decided to check the superposition only on the best particle in the swarm and to regenerate a new particle completely randomly if superposition occurs. The random regeneration of the particles which occupy the leader's position avoids premature convergence since it prevents clone populations from dominating the search (Montalvo *et al*., 2008b). For this purpose regeneration is used by using Linear Fitness Scaling (LFS) technique. The LFS is given by Eq. (16).

$$(f_{best} - f_{worst}) < \varepsilon_{div} \quad (15)$$

Where; f_{best} and f_{worst} are the values of the best and the worst objective functions in the swarm and ε_{div} is the required condition to conserve diversity. In this study ε_{div} is chosen to be 0.001.

In this study, Mean Square Error (MSE) is chosen to be the objective function and the aim is to decrease this function to a minimum value. The objective function is given in Eq. (16).

$$MinF(x) = \frac{1}{N}\sum_{i=1}^{N}(I_i^{Obs} - I_i^{Est})^2 \quad (16)$$

In Eq. (16), I_i^{Obs} is the observed, I_i^{Est} is the predicted rainfall intensity and N is the number of observations.

Optimization with PSO continues until a desired stopping condition occurs. In this study, the PSO algorithm is run until the objective function is minimized. The calculation steps of the optimization process with PSO are:

Step 1. Initialize the search parameters:
- Niter: number of iterations;
- Npt: number of particles;
- Nd: number of searched dimensions;
- $\mathbf{x}^{MIN}$ and $\mathbf{x}^{MAX}$: vectors of length Nd with searching limits;
- c_1, c_2, $\omega_{\min}$, $\omega_{\max}$ PSO searching parameters;
- ε_{div} (diversity tolerance value).
- set k=0 (iteration counter).

Step 2. Calculate the maximum and minimum particle velocities along each direction d according to Eqs. (9a) and (9b)

Step 3. Calculate initial particle positions and velocities according to Eqs. (10) and (11).

Step 4. Calculation of the objective function using Eq. (16).

Step 5. Update $\mathbf{x}^{glo}$, a vector with dimension Nd that contains the best position found by the whole particle swarm.

Step 6. Calculate the inertial weight value by using Eq. (14).

Step 7. Update the particle velocities for p=1…Npt; d=1…Nd by using Eq. (12).

Step 8. If the absolute particle velocity is higher than the maximum permitted value then:
$v_{p,d}^{k+1} = v_d^{Max} sign(v_{p,d}^{k+1})$

Step 9. Update the particle positions by using Eq. (13).

Step 10. If the particle position is not inside the searching limits, the particle is placed at the violated searching limit.

Step 11. Controlling diversity according to Eq. (15).

Step 12. Check the termination criteria. If termination criteria is not satisfied, then go to Step 4.

In this study two termination criteria are used. The search is terminated if the maximum number of iterations is achieved (k=Niter) or Eq. (15) is true.

The following definitions are used in the process of optimization with PSO-RIDF:

The number of particles in the swarm (NP)	The number of weight parameters ×10
The number of weighting parameters	Eq. (1): 3, Eq. (2): 5 Eq. (3): 8, Eq. (4): 10 Eq. (5): 5, Eq. (6): 6 Eq. (7): 5, Eq. (8): 6
c_1 and c_2 coefficient	2 for Eq. (12)
ω inertia coefficient	Calculated according to Eq. (14)
Diversity control	Calculated according to Eq. (15) ε_{div}=0.001

3.2 Numerical Applications

Data for the period of 1938-1987 have been collected from the Turkish State Meteorological Service meteorology station in Izmir, which is located in the western part of Turkey. For the application of the developed model, the maximum annual standard rainfall observation values between 1938 and 2005 have been used under two scenarios. In scenario I, maximum annual standard duration rainfall observation data used in (Karahan *et al*., 2007) covers 50 years between 1938 and 1987. In Scenario I, this data have been used for the confirmation of the developed model and the proposed PSO-RIDF algorithm has been compared to GA in respect of sensitivity, computation time and programming. In scenario II, the data set in Scenario I is increased to 68 years by adding the observation values between 1988 and 2005 for investigating the influence of data set length and recent data on results and model performance. The statistical summary of the data sets for the two scenarios are given in Tables 1 and 2. A general increase in the standard deviations of Scenario II can be observed from these tables.

As the number of data used in the model is limited, the return periods, T_{jl}, are obtained for each rainfall intensity value by using the Gringorten equation (Cunnane, 1978). The Gringorten equation is as follows:

$$T_{ji} = \frac{m_j + 0.12}{L - 0.44} \quad j = 1, 2, 3, ..., 14; l = 1, 2, 3, ..., L \tag{17}$$

where, m_j is the order number for the investigated rainfall duration and L; the number of observation. With the developed solution algorithm, the weighting coefficients of Eqs. (1)-(8) are determined by minimizing the objective function given in Eq. (16) and summarized in Table 3a for Scenario I.

As the new locations of the particles in PSO algorithm are updated according to the location of the best individual of the swarm, the diversity of the swarm decreases or it is completely lost because of the superposition of the particles during the solution process. In this situation, the algorithm may be trapped in a local optimum value. In order to eliminate this situation, the effectiveness of the LFS technique proposed in this study is tested. For this, the Eqs. (1)-(8) are run for 20 times and the results are given in Table 3b. As can be seen from this table, obtaining the same MSE values despite different weighting coefficients indicate that the function is multimodal and the

Table 1. The Statistical Summary of DMI Izmir Station Rainfall Data used in Scenario I

1938-1987	Minutes				Hours									
	5	10	15	30	1	2	3	4	5	6	8	12	18	24
Min (mm)	3.30	3.60	4.40	6.00	9.50	11.20	14.80	16.70	17.50	18.50	18.50	18.50	18.50	31.10
Max (mm)	18.40	29.70	37.40	52.30	60.80	76.50	76.50	76.50	76.50	76.50	76.50	84.90	93.90	143.10
Average (mm)	7.56	10.45	12.54	16.51	21.47	26.40	30.00	32.86	35.38	37.42	40.49	45.70	50.92	61.65
Std. Dev. (mm)	2.88	4.43	5.46	7.38	10.10	11.98	11.94	11.78	11.86	11.78	12.52	15.29	18.64	20.31

Table 2. The Statistical Summary of DMI Izmir Station Rainfall Data Used in Scenario II

1938-2005	Minutes				Hours									
	5	10	15	30	1	2	3	4	5	6	8	12	18	24
Min (mm)	1.50	2.50	2.80	4.00	6.00	6.80	7.00	7.10	9.90	10.60	11.70	11.80	11.80	31.10
Max (mm)	19.60	29.70	37.40	52.30	60.80	76.50	83.00	84.90	99.90	104.50	104.60	104.70	108.00	143.10
Average (mm)	7.46	10.54	12.75	17.03	22.41	27.96	32.14	34.97	37.65	39.68	42.56	47.07	51.54	63.51
Std. Dev. (mm)	4.31	5.36	7.27	9.96	12.59	14.16	14.59	14.90	15.23	15.51	16.85	18.84	20.63	19.85

Table 3(a). The Weighting Coefficients of Equations for Scenario I

Equation number	w_0	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9
1	2.808	0.304	0.589							
2	4.976	0.305	2.139	0.812	0.879					
3	3.914	6.038	-2.662	0.732	-0.057	1.758	0.706	1.027		
4	17.558	19.997	-4.879	0.692	0.036	5.388	8.155	2.485	-1.451	0.392
5	4.489	2.591	1.359	0.438	1.843					
6	9.143	-0.334	0.201	1.830	0.724	1.000				
7	3.061	1.426	1.817	0.733	0.983					
8	17.104	-0.813	0.103	9.979	1.745	0.376				

Table 3(b). Testing Robustness of the PSO Algorithm for Scenario I

Equation number		w_0	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	Objective function (MSE
1	Min	2.808	0.304	0.589								0.0045
	Max	2.808	0.304	0.589								0.0045
	Std. Dev.	0.000	0.000	0.000								0.0000
2	Min	4.826	0.305	1.889	0.726	0.837						0.0038
	Max	5.399	0.305	2.243	0.847	1.000						0.0038
	Std. Dev.	0.242	0.000	0.151	0.051	0.070						0.0000
3	Min	3.476	5.223	-6.700	0.569	-0.132	1.423	0.400	0.842			0.0014
	Max	10.459	15.803	-2.217	1.784	-0.017	2.108	0.838	2.131			0.0015
	Std. Dev.	2.143	3.281	1.511	0.440	0.040	0.215	0.146	0.417			0.0000
4	Min	-10.983	-17.148	-8.423	-2.110	-0.166	-10.676	-11.432	-8.093	-2.398	-0.301	0.0015
	Max	15.361	19.982	7.641	2.243	0.167	8.509	13.821	6.825	2.207	0.421	0.0015
	Std. Dev.	10.461	15.403	6.379	1.684	0.123	5.967	8.575	4.600	1.467	0.280	0.0000
5	Min	2.519	2.591	1.364	0.441	1.111						0.0031
	Max	4.435	2.591	1.606	0.659	1.826						0.0031
	Std. Dev.	0.550	0.000	0.070	0.060	0.202						0.0000
6	Min	8.339	-0.338	0.200	1.582	0.601	0.877					0.0015
	Max	10.975	-0.334	0.201	2.006	0.808	1.248					0.0016
	Std. Dev.	0.929	0.001	0.000	0.162	0.075	0.133					0.0000
7	Min	2.900	1.426	1.635	0.643	0.899						0.0025
	Max	3.443	1.426	1.981	0.792	1.148						0.0025
	Std. Dev.	0.173	0.000	0.107	0.047	0.078						0.0000
8	Min	31.795	-0.888	0.067	1.916	0.762	0.691					0.0024
	Max	37.778	-0.873	0.075	2.779	1.001	0.942					0.0024
	Std. Dev.	1.710	0.005	0.003	0.285	0.077	0.079					0.0000

PSO-RIDF algorithm converges at near global optima.

In order to test the performance of the developed PSO-RIDF algorithm, the results are compared with the results obtained with GA by using the same data set and formulation in (Karahan *et al*., 2007). The performance evaluation of Eqs. (1)-(8) according to various error evaluation criteria for both scenarios is given in Tables 4-6.

In these tables, MSE is the mean square error; MAE is the mean absolute error; r is the correlation coefficient and E is the modified coefficient of efficiency.

For the comparison of PSO-RIDF and GA-RIDF models according to computation time, both models are run 20 times and the average values of the runs are presented in Table 4. As can be seen from Table 4, the PSO-RIDF model requires less computation time than GA-RIDF model. Furthermore, there is no significant sensitivity difference between the performance evaluation criteria (MSE, MAE, r and E) of the models except for Eq. (3). Results of PSO-RIDF are better than GA for Eq. (3). Though the parameter number of the exponential distribution is the same as the empirical formulation (2) and Gumbel distribution, the exponential distribution provides better results in terms of performance evaluation criteria. Similarly, the GEV distribution gives better results than the Pareto distribution though they have the same number of parameters. In addition, the GEV distribution has an advantage over the other polynomial type of distribution although it has six parameters.

Table 4. Performance Comparison of PSO-RIDF and GA-RIDF Models for Scenario I

Eqn #	PSO-RIDF					GA-RIDF					GA (Karahan *et al*., 2007)
	MSE	MAE	r	E	CPU (s)	MSE	MAE	r	E	CPU (s)	MSE
(1)	0.0045	0.0376	0.9910	0.8950	1.18	0.0045	0.0374	0.9910	0.8956	4.56	0.0045
(2)	0.0038	0.0298	0.9921	0.9167	7.86	0.0038	0.0300	0.9921	0.9161	16.02	0.0038
(3)	0.0014	0.0184	0.9971	0.9487	51.26	0.0026	0.0254	0.9946	0.9290	81.32	0.0026
(4)	0.0015	0.0188	0.9969	0.9474	67.63	0.0015	0.0188	0.9969	0.9476	100.26	0.0013
(5)	0.0031	0.0235	0.9936	0.9343	30.17	0.0031	0.0238	0.9935	0.9336	41.30	0.0031
(6)	**0.0015**	**0.0177**	**0.9968**	**0.9506**	39.43	**0.0015**	**0.0192**	**0.9965**	**0.9483**	55.62	**0.0015**
(7)	0.0025	0.0227	0.9949	0.9365	7.13	0.0025	0.0230	0.9948	0.9356	37.62	NA
(8)	0.0024	0.0231	0.9949	0.9355	10.81	0.0027	0.0250	0.9944	0.9303	51.40	NA

Table 5. The Weight Coefficients of Equations for Scenario II

Equation number	w_0	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9
1	2.700	0.311	0.574							
2	5.247	0.311	1.917	0.714	0.996					
3	2.941	7.784	-4.610	1.509	-0.154	1.508	0.669	1.041		
4	3.306	9.317	-5.725	1.891	-0.194	1.446	1.653	1.001	-0.474	0.104
5	3.487	2.293	1.577	0.545	1.379					
6	18.069	-0.674	0.123	1.879	0.710	1.000				
7	3.375	1.227	1.909	0.711	1.000					
8	20.000	-0.844	0.098	4.507	1.303	0.498				

Table 6. Performance Variation of the Equations for Scenario II

Scenario II	MSE	MAE	r	E
Eq. (1)	0.0073	0.0469	0.9858	0.8702
Eq. (2)	0.0066	0.0384	0.9869	0.8936
Eq. (3)	0.0013	0.0184	0.9974	0.9490
Eq. (4)	0.0013	0.0187	0.9974	0.9482
Eq. (5)	0.0021	0.0217	0.9958	0.9399
Eq. (6)	**0.0013**	**0.0173**	**0.9973**	**0.9522**
Eq. (7)	0.0027	0.0245	0.9946	0.9321
Eq. (8)	0.0032	0.0268	0.9936	0.9258

The data used have been changed for Scenario II and the same process is repeated as in Scenario I. The weighting coefficients of Eqs. (1)-(8) are determined and summarized in Table 5.

The performance evaluation of Eqs. (1)-(8) according to various error evaluation criteria for Scenario II is given in Tables 6.

As can be seen from Tables 4 and 6, though the performance of Scenario II is better than Scenario I for the formulas (3-6) according to all evaluation criteria, it is worse in formulas (1, 2, 7 and 8).

Tables 1 and 2 show the statistical summary for the observations used in Scenario I and II. It can be seen that the standard deviation value of the observations in Scenario II is higher than in Scenario I. For this reason, the empirical and statistical formulations with a fewer number of parameters remain insufficient for representing all of the observations. Besides, the formulas with a higher number of parameters show better performance results with an increased number of observations.

As can be seen from Tables 4 and 7, Eq. (6) shows the best performance results for both scenarios according to all evaluation criteria. This situation is clearly depicted in the scatter diagram in Fig. 1, where the observation values and model results are presented together.

By using Eq. (6) which gives the smallest MSE value for both scenarios, the change of rainfall intensity corresponding to rainfall duration and return periods is shown in Fig. 2. Return periods of T=100,200 and 500 years in the related figure show the estimated values for longer periods of time, longer than the observation period (50 and 68 years for Scenario I and II, respectively). Though there has been a similarity between the rainfall intensities for greater return periods in both scenarios, it has been observed that the estimated values for Scenario II are smaller than the ones for Scenario I with the increase of return periods.

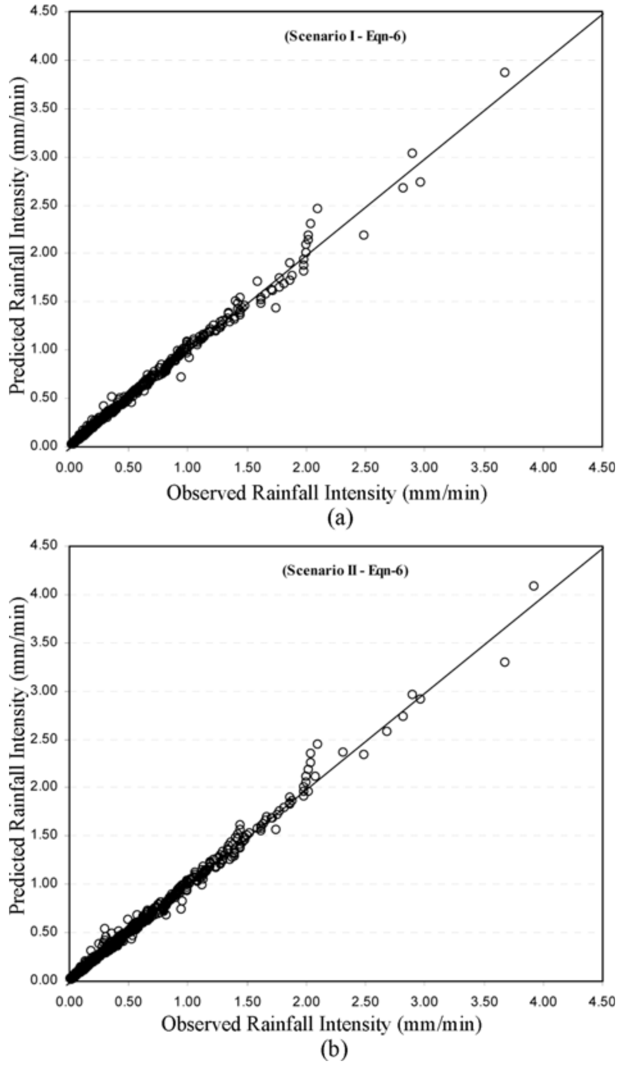

Fig. 1. Predicted and Observed Rainfall Intensities with the Selected PSO-RIDF Model: (a) Scenario I, (b) Scenario II

4. Discussion

No significant difference was found between the Scenario I covering the years 1938-1987 and Scenario II covering 1938-2005 after the evaluation of the PSO-RIDF model results obtained for the related scenarios in Section 3 together with the Table 1 and Table 2 showing the statistical analysis of the data sets used in

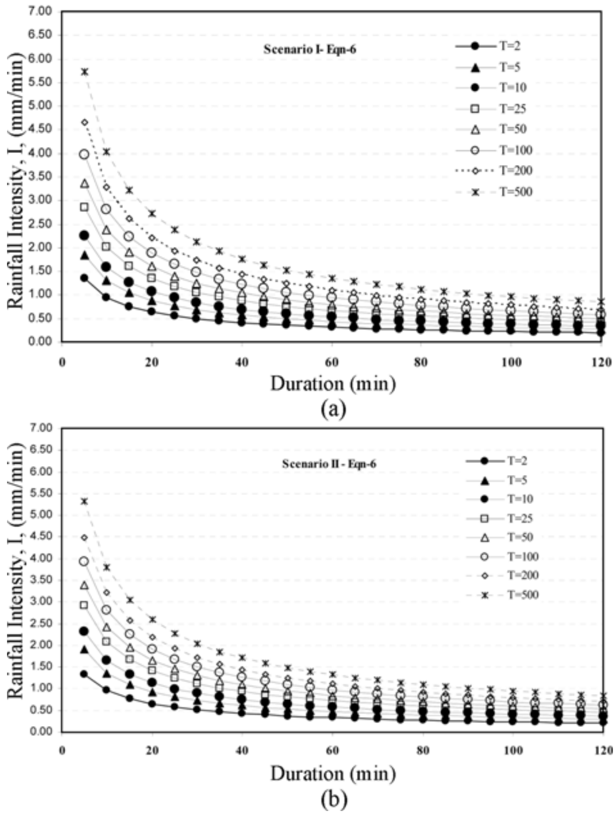

Fig. 2. RIDF Curves Generated (Fitted) by the PSO-RIDF Model: (a) Scenario I, (b) Scenario II

the scenarios. It was thought that this situation was because the majority of the data set used in Scenario II consists of the data set of Scenario I and for investigating if there is a time dependent variation in precipitation values, the data set covering 68 years was divided into two so that it covers the years 1938-1971 and 1972-2005 and the statistical evaluation is summarized in Tables 7 and 8.

Tables 7 and 8 show that two equal data sets following each other show significant difference, and it was seen that the precipitation values are bigger and that the standard deviations increase in the second half covering the years 1972-2005. It was evaluated that the subjects like to which extent this situation can be influential on the PSO-RIDF Model results and how much difference can the situation cause on the estimations aimed at future, is the trend periodic are important subjects that should be investigated hydrologically and as a start for the solution of these problems, PSO-RIDF model was run for two equal periods and the influence of this change in dataset on the model performance were evaluated for all equations and the findings were summarized in Table 9.

It is seen from Table 9 that, for all the equations and error evaluation criteria, the performance of the model covering the years 1938-1972 is significantly better than the performance of the model covering 1972-2005. It is also seen that, when the complete data set is used, the model performance is generally lower than the first half and higher than the second half. For investigating the influence of this situation on the future estimations. the precipitation values for T=68 years for short term precipitations were calculated by using Eq. (6) related to GEV distribution and they were compared with observed values in Table 10.

As can be seen in Table 10, similar results were obtained for Scenarios I and II. However. it can be seen from the last two columns of the table that the estimation values made by using the data sets of the periods 1938-1971 and 1972-2005 show significant difference. It was also seen from Table 10 that there is a very good agreement between the estimations made by using the observed rainfall intensity values (1938-2005) and the data covering the up-to-date values between the years 1972-2005. This situation clearly shows the importance of the existing data set being up-to-date and having enough length. However, to have a better understanding of this subject, a more detailed study is needed and it may be investigated whether there is a similar trend in the neighbouring stations. The next study is thought to be on this subject.

5. Conclusions

With the proposed PSO algorithm, the weighting parameters

Table 7. The Statistical Summary of DMI Izmir Station Rainfall Data for 1938-1971 Period

1938-1971	Minutes				Hours									
	5	10	15	30	1	2	3	4	5	6	8	12	18	24
Min (mm)	3.80	5.90	6.20	8.70	12.10	12.10	14.80	16.70	18.30	18.50	18.50	18.50	18.50	33.20
Max (mm)	14.10	17.10	19.50	27.30	35.40	37.40	37.40	37.40	39.90	43.30	48.10	48.40	69.60	84.20
Average (mm)	7.55	10.25	11.76	15.19	18.57	21.91	24.10	25.95	27.71	29.54	31.57	33.51	37.63	56.33
Std. Dev. (mm)	2.73	3.52	4.26	5.37	6.79	7.50	7.11	7.08	6.60	6.80	7.96	8.59	14.10	15.12

Table 8. The Statistical Summary of DMI Izmir Station Rainfall Data for 1972-2005 Period

1972-2005	Minutes				Hours									
	5	10	15	30	1	2	3	4	5	6	8	12	18	24
Min (mm)	3.30	3.60	4.40	6.00	9.50	11.20	16.70	17.50	17.50	18.80	18.80	18.80	18.80	31.10
Max (mm)	18.40	29.70	37.40	52.30	60.80	76.50	76.50	76.50	76.50	76.50	76.50	84.90	93.90	134.10
Average (mm)	7.51	10.50	12.64	16.81	22.57	28.15	32.35	35.69	38.50	40.58	44.04	50.22	55.90	63.64
Std. Dev. (mm)	3.03	4.90	5.91	8.18	11.20	13.24	12.91	12.38	12.29	11.91	12.16	14.28	17.33	21.91

Table 9. Performance Comparison of PSO-RIDF for Different Observation Periods

	Observation period											
	1938-2005 (68 years)				1938-1971 (34 years)				1972-2005 (34 years)			
	MSE	MAE	r	E	MSE	MAE	r	E	MSE	MAE	r	E
Eq. (1)	0.0073	0.0469	0.9858	0.8702	0.0038	0.0340	0.9905	0.8971	0.0126	0.0620	0.9794	0.8416
Eq. (2)	0.0066	0.0384	0.9869	0.8936	0.0036	0.0287	0.9910	0.9134	0.0112	0.0528	0.9813	0.8652
Eq. (3)	0.0013	0.0184	0.9974	0.9490	0.0011	0.0174	0.9971	0.9474	0.0019	0.0248	0.9969	0.9367
Eq. (4)	0.0013	0.0187	0.9974	0.9482	0.0011	0.0175	0.9971	0.9470	0.0018	0.0236	0.9971	0.9396
Eq. (5)	0.0021	0.0217	0.9958	0.9399	0.0014	0.0181	0.9964	0.9452	0.0032	0.0292	0.9947	0.9254
Eq. (6)	**0.0013**	0.0173	0.9973	0.9522	**0.0013**	0.0166	0.9967	0.9497	**0.0027**	0.0288	0.9956	0.9264
Eq. (7)	0.0027	0.0245	0.9946	0.9321	0.0018	0.0202	0.9953	0.9390	0.0047	0.0377	0.9922	0.9038
Eq. (8)	0.0032	0.0268	0.9936	0.9258	0.0021	0.0224	0.9946	0.9324	0.0059	0.0419	0.9903	0.8931

Table 10. Comparison of Estimated and Observed RI for Different Scenarios and Observation Periods

Duration (min)	Observed RI (mm/min)	Estimated RI (mm/min)			
	1938-2005 Period (All data)	1938-1987 Period (Scenario I)	1938-2005 Period (Scenario II)	1938-1971 Period	1972-2005 Period
5	3.920	3.627	3.621	2.932	4.167
10	2.970	2.563	2.591	2.017	3.060
15	2.493	2.045	2.082	1.596	2.480
30	1.743	1.347	1.389	1.048	1.666
60	1.013	0.861	0.900	0.674	1.083
120	0.638	0.540	0.571	0.427	0.690

can be determined with an optimization technique in one step instead of statistical calculations and/or many trial and error calculations. Because of its ease of programmability and high performance, the optimization technique is proposed as an effective method for RIDF analysis and the PSO algorithm that can obtain results close to the global optimum is proposed.

For testing the proposed PSO-RIDF algorithm, the data set and the formulation given in (Karahan *et al*., 2007) is used and two additional relationships based on exponential and Pareto distributions are proposed. It has been observed that the performance of the PSO-RIDF shows better performance results than the GA approach does for higher number of parameters.

To investigate the influence of data set length on model performance, the weighting coefficients of the proposed formulas are determined again by applying the procedure after adding the recent data of 18 years. No significant difference can be observed between the results obtained for the two data sets.

The empirical and statistical formulations with a lower number of parameters remain insufficient for representing all of the observations. Besides, the formulas with a higher number of parameters show better performance results with the increased number of observations. The performances of equations with a high number of parameters are close to each other and no significant difference can be observed between the two data sets.

For different return periods, the variation of rainfall intensity according to time is obtained by using the equation related to GEV distribution. The obtained results show that the solution algorithm proposed as an alternative technique in the determination of RIDF relationship, gives good results for both scenarios.

The proposed PSO-RIDF algorithm gives the same objective function value for different runs and this shows that the algorithm is robust for this set of data and the LFS technique is effective in the conservation of diversity.

The method proposed by the author is an advanced technique for the first time in literature for the RIDF analysis and evaluation of hydrological variables. The engineers or researchers may easily calculate the rainfall intensity by using the given weight coefficients in the related equations for rainfalls with different return periods and durations.

Acknowledgements

A part of this study is supported by “The Scientific and Technological Research Council of Turkey (TUBITAK)”. The author would like to thank TUBITAK for giving this support. The author is also thankful to the anonymous reviewers and the Editor for their insightful comments.

References

Abido, M. A. (2002). “Optimal design of power system stabilizers using particle swarm optimization.” *IEEE Trans Energy Conv*, Vol. 17, Issue 3, pp. 723-729.

Bernard, M. M. (1932). “Formulas for rainfall intensities of long durations.” *Trans. ASCE* Vol. 96, pp. 592-624.

Chau, K. W. (2006). “Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River.” *Journal of Hydrology*, Vol. 329, Issues 3-4, pp. 363-367.

Chau, K. W. (2007). “A split-step particle swarm optimization algorithm in river stage forecasting.” *Journal of Hydrology*, Vol. 346, Issues 3-4, pp. 131-135.

Chen, C. I. (1983). Rainfall intensity-duration-frequency formulas.” *J. Hydraul. Eng.,* Vol. 109, Issue 12, pp. 1603-1621.

Clerc, M. (1998). “The swarm and the queen: Towards a deterministic and adaptive particle swarm optimization.” *Proceedings of the ICEC*, Washington, D.C., 1951-7.

Cunnane, C. (1978). “Unbiased plotting positions: A review.” *Journal of Hydrology*, Vol. 37, pp. 205-222.

Eberhart, R. C. and Kennedy, J. (1995). “A new optimizer using particle swarm theory.” *Proceedings, Sixth International Symposium on Micro Machine and Human Science*, Nagoya, Japan: IEEE Service

Center; pp. 39-43.

Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley.

Izquierdo, J., Montalvo, I., Perez, R., and Fuertes, V. S. (2008). "Design optimization of wastewater collection networks by PSO." *Computers and Mathematics with Applications*, Vol. 56, Issue 3, pp. 777-784.

Karahan, H., Ayvaz, M. T., and Gurarslan, G. (2008). "Determination of intensity-duration-frequency relationship by Genetic Algorithm: Case study of GAP." *Teknik Dergi*, Vol. 19, Issue 2, pp. 4393-4407.

Karahan, H., Ceylan, H., and Ayvaz, M. T. (2007). "Predicting rainfall-intensity using genetic algorithm approach." *Hydrological Processes*, Vol. 21, Issue 4, pp. 470-475.

Kennedy, J. and Eberhart, R. C. (1995). "Particle swarm optimization." *Proceedings of IEEE international Conference on Neural Networks*, Perth, Australia, pp. 1942-1948.

Kothyari, U. C. and Garde, R. J. (1992). "Rainfall intensity-duration-frequency formula for India." *J. Hydraul. Eng.*, Vol. 118, Issue 2, pp. 323-336.

Koutsoyiannis, D., Kozonis, D., and Manetas, A. (1998). "A mathematical framework for studying rainfall intensity-duration-frequency relationships." *Journal of Hydrology*, Vol. 206, Issues 1-2, pp. 118-135.

Montalvo, J., Izquirerdo, R. Perez, and Iglesias, P. L. (2008b). "A diversity-enriched variant of discrete PSO applied to the design of water distribution networks." *Engineering Optimization*, Vol. 40, Issue 7, pp. 655-668.

Montalvo, J., Izquirerdo, R. Perez, and Tung, M. M. (2008a). "Particle swarm optimization applied to the design of water supply systems." *Computers and Mathematics with Applications*, Vol. 56, Issue 3, pp. 769-776.

Schwaab, M., Biscaia, E. C., Monteiro, J. L., and Pinto, J. C. (2008). "Nonlinear parameter estimation through particle swarm optimization." *Chemical Engineering Science*, Vol. 63, Issue 6, pp. 1542-1552.

Shi, Y. and Eberhart, R., (1998). "A modified particle swarm optimizer." *Proceedings of the IEEE International Conference on Evolutionary Computation*, Anchorage, Alaska, pp. 69-73.

Yagiz, S. and Karahan, H. (2011). "Prediction of hard rock TBM penetration rate using particle swarm optimization." *Int. J. Rock Mech. Min. Sci.*, Vol. 48, Issue 3, pp. 427-433.

Yoshida, H., Kawata, K., Fukuyama, Y., Takayama, S., and Nakanishi, Y. (2000). "A particle swarm optimization for reactive power and voltage control considering voltage security assessment." *IEEE Trans Power Syst*, Vol. 15, Issue 4, pp. 1232-1239.

Yu, B., Yuan, X., and Wang, J. (2007). "Short-term hydro-thermal scheduling using particle swarm optimization method." *Energy Conversion and Management*, Vol. 48, Issue 7, pp. 1902-1908.