Firefly Algorithm for Demand Estimation of Water Resources

Hui Wang[1,2](✉), Zhihua Cui[3], Wenjun Wang[4], Xinyu Zhou[5], Jia Zhao[1,2], Li Lv[1,2], and Hui Sun[1,2]

[1] Jiangxi Province Key Laboratory of Water Information Cooperative Sensing and Intelligent Processing, Nanchang Institute of Technology, Nanchang 330099, China
huiwang@whu.edu.cn

[2] School of Information Engineering, Nanchang Institute of Technology, Nanchang 330099, China

[3] Complex System and Computational Intelligence Laboratory, Taiyuan University of Science and Technology, Taiyuan 030024, China

[4] School of Business Administration, Nanchang Institute of Technology, Nanchang 330099, China

[5] College of Computer and Information Engineering, Jiangxi Normal University, Nanchang 330022, China

Abstract. Firefly algorithm (FA) is an efficient swarm intelligence optimization technique, which has been used to solve many engineering optimization problems. In this paper, we present a new FA (called NFA) variant for demand estimation of water resources in Nanchang city of China. The performance of the standard FA highly depends on its control parameters. To tackle this issue, a dynamic step factor strategy is proposed. In NFA, the step factor is not fixed and it is dynamically updated during the search process. Three models in different forms (linear, exponential and hybrid) are developed based on the structure of social and economic conditions. Water demand in Nanchang city from 2003 to 2015 is considered as a case study. The data from 2003 to 2012 is used for finding the optimal weights, and the rest data (2013–2015) is for testing the models. Simulation results show that three FA variants can achieve promising performance. Our proposed NFA outperforms the standard FA and memetic FA (MFA), and the prediction accuracy is up to 97.91%.

Keywords: Firefly algorithm · Swarm intelligence · Water demand estimation · Water demand forecasting · Optimization

1 Introduction

Water is a valuable resource for human survival and social economic development. It is an irreplaceable basic natural resource and strategic economic resource. With the speeding up of urbanization process, the demand of water resources is increasing. However, the amount of water resources is limited in nature. Therefore, the optimal allocation of water resources is important to the sustainable utilization of water resources [1].

D. Liu et al. (Eds.): ICONIP 2017, Part IV, LNCS 10637, pp. 11–20, 2017.
https://doi.org/10.1007/978-3-319-70093-9_2

Demand estimation of water resources is a significant step in the allocation of water resources. The estimation of water demand aims to infer the water demand in the future according to the historical water consumption, current situation, and environment changes. As the basic means of water resources planning and management, water demand estimation is the prerequisite for the optimal allocation of water resources. Water demand is related to population, economy, social policy, ecological environment and other factors. Due to some uncertain factors, the estimation error exists. How to exactly estimate the water demand is worthy to be investigated.

Traditional estimation methods for water demand includes time series, regression analysis, gray predication, artificial neural network (ANN), quota method, and so on. For these methods, how to choose weighting parameters is a difficult task because of the random behaviors of water consumptions. In the past several years, some intelligent algorithms have been used to estimate the water demand [1–5]. In [1], a hybrid model based on soft computing techniques was used to improve the demand estimation of irrigation water. Fuzzy logic and genetic algorithm (GA) were combined with computational neural network (CNN). Experimental results show that the hybrid model outperforms the single CNN model. Do et al. [2] used GA to estimate the water demand in water distribution system. Simulation results show that GA can achieve good solutions on a 24-h period case study. In [3], harmony search (HS) was applied to short term water demand estimation, in which HS aims to search the parameter of a double seasonal ARIMA model. Romano and Kapelan [4] combined evolutionary algorithms (EAs) and ANN to construct a smart estimation model. Reported results show that the mean error is about 5%. Bai et al. [5] proposed a multi-scale method for urban water demand estimation. In the approach, an adaptive chaotic particle swarm optimization (PSO) was used to search the optimal parameters of the relevance vector regression model.

Swarm intelligence is a new computational paradigm inspired by the social behaviors from the nature [6]. In recent years, some efficient swarm intelligence algorithms have been proposed, such as artificial bee colony (ABC) [7–10], firefly algorithm (FA) [11–16], cuckoo search (CS) [17, 18] and bat algorithm (BA) [19]. FA is inspired by the mating behaviors of flashing fireflies [11]. Some recent studies show that FA can achieve promising results on many benchmark functions and real-world problems [20]. In this paper, we present an application of FA on demand estimation of water resources in Nanchang city of China. To reduce the dependency of FA on its parameters, a dynamic step factor strategy is proposed. In our approach, the step factor is not fixed and it can be dynamically updated during the search process. Three models based on linear, exponential and hybrid forms are developed based on the structure of social and economic conditions of Nanchang city. In the experiments, the performance of proposed NFA is compared with the standard FA and memetic FA (MFA) [21].

The rest of the paper is organized as follows. In Sect. 2, the standard FA is briefly described. Estimation models are developed in Sect. 3. Our approach is proposed in Sect. 4. Results and discussions are presented in Sect. 5. Finally, this work is concluded in Sect. 6.

2 Firefly Algorithm

Like PSO, FA is also a population-based random search algorithm. Each individual (firefly) in the population represents a candidate solution. The search of FA is inspired by the mating behavior of flashing fireflies. When a firefly is attracted by other brighter ones, it can move toward other new positions and find potential solutions. To construct the search model, Yang [11] proposed three assumptions: (1) one firefly is attracted to other all brighter ones; (2) the attractiveness is determined by the brightness; and (3) the brightness is affected by the given objective function. The assumptions mean that a brighter firefly has a better fitness value.

Let $X_i = (x_{i1}, x_{i1}, \ldots, x_{iD})$ be the ith firefly in the population, where $i = 1, 2, \ldots, N$, N is the population size, and D is the dimensional size. For any two different fireflies X_i and X_j, their attractiveness can be calculated as follows [11].

$$\beta(r_{ij}) = \beta_0 e^{-\gamma r_{ij}^2} \tag{1}$$

where β_0 is the attractiveness at $r = 0$, γ is the light absorption coefficient, and r_{ij} is the distance between X_i and X_j. The distance r_{ij} is defined by [11]

$$r_{ij} = \left\| X_i - X_j \right\| = \sqrt{\sum_{d=1}^{D} (x_{id} - x_{jd})^2} \tag{2}$$

when X_j is brighter (better) than X_i, X_i will move toward X_j due to the attraction. In the standard FA, this movement is defined as follows [11].

$$x_{id}(t+1) = x_{id}(t) + \beta(r_{ij}) \cdot \left(x_{jd}(t) - x_{id}(t)\right) + \alpha\left(rand - \frac{1}{2}\right) \tag{3}$$

where x_{id} and x_{jd} are the dth dimensions of X_i and X_j, respectively, $\alpha \in [0, 1]$ is called step factor, and *rand* is a random value within [0, 1].

3 Estimation Models

In this paper, we focus on an application of FA to estimate the water demand in Nanchang city of China. The water demand is related to the social and economic conditions. Table 1 shows the historical water use in Nanchang city from 2003 to 2015 [22, 23]. It can be seen that the water use of Nanchang is distributed in three departments, agriculture, industry and residents. The average proportion of agricultural water use is up to 57%. It demonstrates that agriculture is the main department of water use. Industrial and residential water use also take large proportions. The ecological water use is only 3%. So, three factors related to agricultural, industrial, and residential water use are utilized to construct the estimation model, while the ecological factor is ignored.

Table 1. Historical water use in Nanchang city from 2013 to 2015 (10^8 m^3).

Year	Total water use	Industrial water use	Agricultural water use	Residential water use	Ecological water use
2003	24.21	9.81	11.55	2.53	0.32
2004	26.22	8.72	14.47	2.75	0.28
2005	28.14	8.30	16.92	2.60	0.32
2006	27.71	8.11	16.73	2.52	0.35
2007	32.55	7.51	21.27	2.92	0.85
2008	30.42	6.90	19.73	2.94	0.85
2009	33.42	6.57	20.15	3.21	3.49
2010	30.87	7.51	17.37	3.49	2.50
2011	31.26	8.97	17.70	4.03	0.56
2012	28.82	9.20	14.68	4.36	0.58
2013	32.62	9.35	18.23	4.45	0.59
2014	31.42	8.92	17.35	4.54	0.61
2015	30.64	9.17	16.21	4.64	0.62
Average	29.87	**8.39 (28%)**	**17.10 (57%)**	**3.46 (12%)**	**0.92 (3%)**

From the above analysis, we use gross agricultural production, gross industrial production, and population to associate with agricultural, industrial, and residential water use, respectively. Table 2 lists the total water use, population, gross industrial production, and gross agricultural production in Nanchang city from 2003 to 2015 [22, 23]. By the suggestions of [24], both linear and exponential forms of models for water demand estimation are defined as follows.

Linear estimation model:

$$Y_l = x_1 \cdot W_1 + x_2 \cdot W_2 + x_3 \cdot W_3 + x_4 \tag{4}$$

Exponential estimation model:

$$Y_\mathrm{e} = x_1 \cdot W_1^{x_2} + x_3 \cdot W_2^{x_4} + x_5 \cdot W_3^{x_6} + x_7 \tag{5}$$

where W_1, W_2, and W_3 are the population, gross industrial production, and gross agricultural production, respectively, and $x_i \in [0, 1]$ are the corresponding weights.

In this paper, we propose a hybrid model, which is a middle phase between linear and exponential models. The new model is defined by

$$Y_h = x_1 \cdot Y_l + (1 - x_1) \cdot Y_e \tag{6}$$

where Y_l and Y_e are linear and exponential models, respectively, and $x_1 \in [0, 1]$ is the weighting factor. The hybrid model can be written as follows.

$$\begin{aligned} Y_h = {} & x_1 \cdot (x_2 \cdot W_1 + x_3 \cdot W_2 + x_4 \cdot W_3 + x_5) \\ & + (1 - x_1)\left(x_6 \cdot W_1^{x_7} + x_8 \cdot W_2^{x_9} + x_{10} \cdot W_3^{x_{11}} + x_{12}\right) \end{aligned} \tag{7}$$

Table 2. The total water use, population, gross industrial production, and gross agricultural production in Nanchang city from 2003 to 2015.

Year	Total water use (10^8 m^3)	Population	Gross industrial production (10^8 yuan)	Gross agricultural production (10^8 yuan)
2003	24.21	4437476	250.95	51.29
2004	26.22	4469671	306.08	99.11
2005	28.14	4500672	374.93	115.76
2006	27.71	4530776	448.15	124.58
2007	32.55	4563025	532.75	142.84
2008	30.42	4597936	676.61	171.14
2009	33.42	4632067	753.20	187.20
2010	30.87	5042567	952.75	204.66
2011	31.26	5088996	1223.72	229.70
2012	28.82	5131564	1290.93	249.35
2013	32.62	5184231	1398.63	266.12
2014	31.42	5240179	1500.70	283.63
2015	30.64	5302914	1619.50	296.92

4 Proposed Approach

4.1 Dynamic Parameter Strategy

The performance of FA is seriously affected by its control parameters α and β. In our previous study [25], we analyzed the relations between the step factor α and convergence. If FA is convergent, α should satisfy the following condition.

$$\lim_{t \to \infty} \alpha = 0 \tag{8}$$

where t is the index of iteration.

According to Eq. 8, a dynamic step factor strategy is designed to automatically adjust the parameter α as follows.

$$\alpha(t+1) = \alpha(t) \cdot \exp\left(-k \cdot \frac{t}{T_{\max}}\right) \tag{9}$$

where T_{max} is the maximum number of iterations. $k \in (0, 1]$ is called decreasing rate, which can adjust the decreasing speed of α. In this paper, $k = 0.2$ is used based on empirical studies. In some recent literature, the parameter α was limited in the range [0, 1]. So, the initial $\alpha(0)$ is set to 0.5, which is the midpoint of the range.

4.2 Normalization

In this paper, the historical data from 2003 to 2015 listed in Table 2 is used for training and testing the estimation models for water demand. To eliminate the influences of different units of data, the normalization method is used. In Table 2, the total water use, population, gross industrial production, and gross agricultural production are normalized as follows.

$$W^* = \frac{W - W_{\min}}{W_{\max} - W_{\min}} \tag{10}$$

where W^* is the normalized value, W is the value to be normalized, W_{min} and W_{max} are the minimum and maximal values for the corresponding variable, respectively.

4.3 Fitness Evaluation Function

The data from 2003 to 2012 is used to optimize the weighting factors of the estimation models, and the rest data (2013–2015) is applied to test the models. To evaluate the quality of obtained weighting factors, sum of squared errors (SSE) is employed to construct the fitness evaluation function.

$$f(X) = \sum_{i=1}^{m} \left(Y_{pre} - Y_{act}\right)^2 \tag{11}$$

where Y_{act} and Y_{pre} are the actual and predicted water demand, respectively, and m is the number of training samples.

5 Simulation Experiments

5.1 Experimental Setup

In the experiments, the proposed NFA is applied to estimate the water demand in Nanchang city. The performance of NFA is compared with standard FA and memetic FA (MFA) [21]. To have a fair comparison, the same parameter settings are used. The population size N and *MaxFEs* are set to 30 and 1.0E+05, respectively. In the standard FA, α and β_0 are set to 0.5 and 1.0, respectively. For MFA, the initial α, γ, β_0, and $\beta_{\min}$ are set to 0.5, 1.0, 1.0, and 0.2, respectively. In NFA, $\alpha(0)$ and $\beta_0(0)$ are equal to 0.5 and 1.0, respectively. The γ is set to $1/\Gamma^2$, where Γ is the length of search range.

Data from 2003 to 2012 listed in Table 2 is used to optimize the weighting factors of the estimation models, and the rest data (2013–2015) is applied to test the models. For each model, each algorithm is run 20 times and mean results are recorded. In the experiments, we use relative error (RE) and mean relative error (MRE) to measure the performance of FA.

$$RE = \left|\frac{Y_{pre} - Y_{act}}{Y_{act}}\right| \tag{12}$$

$$MRE = \frac{1}{n} \cdot \sum_{i=1}^{n} \left|\frac{Y_{pre}(i) - Y_{act}(i)}{Y_{act}(i)}\right| \tag{13}$$

where $Y_{pre}(i)$ and $Y_{pre}(i)$ are the predicted and actual water demand on the ith test sample, respectively, and n is the number of test samples.

5.2 Results

Tables 3, 4 and 5 present the results for the linear, exponential, and hybrid estimation models, respectively. As seen, FA, MFA, and NFA can achieve promising results on three estimation models. The best MRE is 2.09% and the worst one is only 5.76%. It means that the prediction accuracy is between 94.24% and 97.91%. For each model, NFA achieves better results than FA and MFA, and FA obtains the worst performance among three algorithms. The exponential model is better than the linear one. Results on the hybrid model are in line with our idea, which aims to provide a middle phase between the linear and exponential models. For all FA variants, the mean MRE on the hybrid model is better than the linear one, but worse than the exponential one. It is surprised that the best MRE on the hybrid model is better than other two models.

Table 3. Results for the linear estimation model.

Algorithm	Best MRE	Mean MRE	Std	Worst MRE
FA	4.94%	4.97%	3.60E−04	5.05%
MFA	4.92%	4.96%	**2.05E−04**	4.98%
NFA	**4.89%**	**4.95%**	2.33E−04	**4.96%**

Table 4. Results for the exponential estimation model.

Algorithm	Best MRE	Mean MRE	Std	Worst MRE
FA	2.42%	2.53%	1.95E−03	3.00%
MFA	2.36%	2.40%	3.70E−04	2.47%
NFA	**2.27%**	**2.35%**	**3.46E−04**	**2.38%**

Table 5. Results for the hybrid estimation model.

Algorithm	Best MRE	Mean MRE	Std	Worst MRE
FA	2.25%	3.55%	1.39E−02	5.76%
MFA	2.22%	2.82%	7.04E−03	4.25%
NFA	**2.09%**	**2.79%**	**6.98E−03**	**4.06%**

Tables 6, 7 and 8 show the best relative errors for the linear, exponential, and hybrid estimation models, respectively. It can be seen that the linear model achieves good fitting for year 2013, and the exponential model is suitable for year 2014. Results on the hybrid model show that the combination of the linear and exponential models can provide more chances of finding better solutions. Due to the space limitation, some figures and forecasting results from 2018 to 2020 are not presented.

Table 6. The best relative errors (RE) for the linear model.

Year	FA RE	MFA RE	NFA RE
2013	**1.04%**	1.07%	1.14%
2014	4.75%	4.71%	**4.64%**
2015	9.02%	8.99%	**8.90%**
Average	4.94%	4.92%	**4.89%**

Table 7. The best relative errors (RE) for the exponential model.

Year	FA RE	MFA RE	NFA RE
2013	**4.23%**	4.32%	4.62%
2014	0.14%	**0.01%**	0.69%
2015	2.90%	2.75%	**1.50%**
Average	2.42%	2.36%	**2.27%**

Table 8. The best relative errors (RE) for the hybrid model.

Year	FA RE	MFA RE	NFA RE
2013	4.41%	4.31%	**3.90%**
2014	0.51%	0.13%	**0.000037%**
2015	**1.82%**	2.23%	2.36%
Average	2.25%	2.22%	**2.09%**

6 Conclusions

In this paper, we present a new FA (NFA) to estimate the water demand in Nanchang city of China. To improve the performance of the original FA, a dynamic strategy is proposed to adjust the step factor during the search process. By analyzing the historical water use of Nanchang, three estimation models (linear, exponential and hybrid) are developed. Moreover, the normalization method is employed to eliminate the effects of different units of test data.

Data from 2003 to 2012 is used to optimize the weighting factors of the estimation models, and the rest data (2013–2015) is applied to test the models. Simulation results show that FA, MFA and NFA can achieve promising performance. NFA outperforms FA and MFA and the prediction accuracy is up to 97.91%.

This paper only uses three factors (population, gross industrial production, and gross agricultural production) to construct the estimation model. However, there are some uncertain factors, such as social policy and climate change, which may affect the water demand. This will be further investigated in the future work.

Acknowledgement. This work was supported by the National Natural Science Foundation of China (No. 61663028), the Distinguished Young Talents Plan of Jiangxi Province (No. 20171BCB23075), the Natural Science Foundation of Jiangxi Province (No. 20171BAB202035), and the Open Research Fund of Jiangxi Province Key Laboratory of Water Information Cooperative Sensing and Intelligent Processing (No. 2016WICSIP015).

References

1. Pulido-Calvo, I., Gutiérrez-Estrada, J.C.: Improved irrigation water demand forecasting using a soft-computing hybrid model. Biosyst. Eng. **102**(2), 202–218 (2009)
2. Do, N., Simpson, A., Deuerlein, J., Piller, O.: Demand estimation in water distribution systems: solving underdetermined problems using genetic algorithms. Procedia Eng. **186**, 193–201 (2017)
3. Oliveira, P.J., Steffen, J.L., Cheung, P.: Parameter estimation of seasonal Arima models for water demand forecasting using the harmony search algorithm. Procedia Eng. **186**, 177–185 (2017)
4. Romano, M., Kapelan, Z.: Adaptive water demand forecasting for near real-time management of smart water distribution systems. Environ. Model Softw. **60**, 265–276 (2014)
5. Bai, Y., Wang, P., Li, C., Xie, J.J., Wang, Y.: A multi-scale relevance vector regression approach for daily urban water demand forecasting. J. Hydrol. **517**, 236–245 (2014)
6. Torres-Treviño, L.M.: Let the swarm be: an implicit elitism in swarm intelligence. Int. J. Bio-Inspired Comput. **9**(2), 65–76 (2017)
7. Sun, H., Wang, K., Zhao, J., Yu, X.: Artificial bee colony algorithm with improved special centre. Int. J. Comput. Sci. Math. **7**(6), 548–553 (2016)
8. Yu, G.: A new multi-population-based artificial bee colony for numerical optimization. Int. J. Comput. Sci. Math. **7**(6), 509–515 (2016)
9. Lv, L., Wu, L.Y., Zhao, J., Wang, H., Wu, R.X., Fan, T.H., Hu, M., Xie, Z.F.: Improved multi-strategy artificial bee colony algorithm. Int. J. Comput. Sci. Math. **7**(5), 467–475 (2016)
10. Lu, Y., Li, R.X., Li, S.M.: Artificial bee colony with bidirectional search. Int. J. Comput. Sci. Math. **7**(6), 586–593 (2016)
11. Yang, X.S.: Nature-Inspired Metaheuristic Algorithms. Luniver Press, Beckington (2008)
12. Marichelvam, M.K., Geetha, M.: A hybrid discrete firefly algorithm to solve flow shop scheduling problems to minimise total flow time. Int. J. Bio-Inspired Comput. **8**(5), 318–325 (2016)
13. Wang, H., Wang, W., Sun, H., Rahnamayan, S.: Firefly algorithm with random attraction. Int. J. Bio-Inspired Comput. **8**(1), 33–41 (2016)
14. Kaur, M., Sharma, P.K.: On solving partition driven standard cell placement problem using firefly-based metaheuristic approach. Int. J. Bio-Inspired Comput. **9**(2), 121–127 (2017)
15. Yu, G.: An improved firefly algorithm based on probabilistic attraction. Int. J. Comput. Sci. Math. **7**(6), 530–536 (2016)
16. Wang, H., Wang, W.J., Zhou, X.Y., Sun, H., Zhao, J., Yu, X., Cui, Z.: Firefly algorithm with neighborhood attraction. Inf. Sci. **382–383**, 374–387 (2017)
17. Cui, Z., Sun, B., Wang, G., Xue, Y., Chen, J.: A novel oriented cuckoo search algorithm to improve DV-Hop performance for cyber-physical systems. J. Parallel Distrib. Comput. **103**, 42–52 (2017)

18. Zhang, M., Wang, H., Cui, Z., Chen, J.: Hybrid multi-objective cuckoo search with dynamical local search. Memet. Comput. (2017, in press). doi:10.1007/s12293-017-0237-2
19. Cai, X., Gao, X.Z., Xue, Y.: Improved bat algorithm with optimal forage strategy and random disturbance strategy. Int. J. Bio-Inspired Comput. **8**(4), 205–214 (2016)
20. Fister, I., Fister Jr., I., Yang, X.S., Brest, J.: A comprehensive review of firefly algorithms. Swarm Evol. Comput. **13**, 34–46 (2013)
21. Fister Jr., I., Yang, X.S., Fister, I., Brest, J., Memetic firefly algorithm for combinatorial optimization, arXiv preprint arXiv:1204.5165 (2012)
22. Statistic Bureau of Jiangxi: Jiangxi Statistical Yearbook, Chinese Statistical Press, Beijing (2004–2016)
23. Statistics Bureau of Nanchang: Statistical bulletin of national economic and social development of Nanchang, Nanchang (2003–2015)
24. Assareh, E., Behrang, M.A., Assari, M.R., Ghanbarzadeh, A.: Application of PSO (particle swarm optimization) and GA (genetic algorithm) techniques on demand estimation of oil in Iran. Energy **35**, 5223–5229 (2010)
25. Wang, H., Zhou, X.Y., Sun, H., Yu, X., Zhao, J., Zhang, H., Cui, L.Z.: Firefly algorithm with adaptive control parameters. Soft. Comput. (2016, in press). doi:10.1007/s00500-016-2104-3

Springer

http://www.springer.com/978-3-319-70092-2

Neural Information Processing
24th International Conference, ICONIP 2017,
Guangzhou, China, November 14–18, 2017,
Proceedings, Part IV
Liu, D.; Xie, S.; Li, Y.; Zhao, D.; El-Alfy, E.-S.M. (Eds.)
2017, XVIII, 898 p. 326 illus., Softcover
ISBN: 978-3-319-70092-2