Water Resources Management
https://doi.org/10.1007/s11269-024-03738-7

A Novel Strategic Water Loss Management Model and Its Optimization with Harris Hawk Algorithm

Cansu Bozkurt[1] · Abdullah Ates[2] · Mahmut Fırat[3] · Salih Yılmaz[4] · Özgür Özdemir[5]

Received: 19 January 2023 / Accepted: 3 January 2024
© The Author(s) 2024

Abstract
Water loss management (WLM) in water distribution systems (WDSs) is generally time consuming, costly and difficult. Therefore, the most appropriate methodology should be defined based on the current status, requirements and technical capacity of the utilities. In this study, a novel WLM model was developed to improve the WLM practices, to define the most appropriate and applicable targets and to determine the components that need to be improved primarily using the Harris Hawk Optimization (HHO) algorithm. Firstly, a total of 144 components in the CSAS proposed by Bozkurt et al. (Welcome Sigma J Eng Natural Sci 40:1–13, 2022b) are scored in three utilities. The averages of current scores in utilities are 3.4167, 2.9792 and 1.1597, respectively. Then the most appropriate targets were defined with the optimization algorithm by considering the current situation and the dynamic structure of utility. The target scores defined by optimization model in utilities are 4.4097, 4.3958 and 3.2292, respectively. This novel model will contribute to the prevention of time-consuming and costly processes by determining the most appropriate progress. It is thought that this study will provide significant benefits for creating a road map for practitioners in WLM.

Keywords Water distribution system · Water loss management · Optimization · Harris Hawk optimization

1 Introduction

Efficient use of existing water resources is imperative and critical due to industrialization, population growth, global warming and high-water consumptions. Utilities globally face the water losses and energy inefficiencies. Failures that occur in WDSs due to factors that are hydraulic, physical, environmental and operating, cause inefficient

✉ Abdullah Ates
abdullah.ates@inonu.edu.tr

[1] Technical Sciences Department, Ardahan University, Ardahan, Turkey

[2] Computer Engineering Department, Inonu University, Malatya, Turkey

[3] Civil Engineering Department, Inonu University, Malatya, Turkey

[4] Civil Engineering Department, Çankırı Karatekin University, Çankırı, Turkey

[5] Environmental Engineering Department, Adiyaman University, Adiyaman, Turkey

operation of the network, poor service quality and high operating and investment costs (Pearson 2019). The optimum design, renewal and rehabilitation of WDSs and water loss reduction are quite critical for sustainable infrastructure management (Babu and Vijayalakshmi 2013). The methods of the active leakage control (Azevedo and Saurin 2018; Gupta and Kulat 2018), minimum night flow (MNF) (Farah and Shahrour 2017; Marzola et al. 2021) and pressure management (AL-Washali et al. 2020), have been applied for ensuring the water and energy efficiency. Mutikanga et al. (2013) reviewed the methods and tools applied for WLM in detail. The recommendations that are the improvement of the data quality and implementation of leakage reduction and performance evaluation practices, were listed. Moreover, the works related to customer meter management (Fontanazza et al. 2015), and the performance assessment (Marzola et al. 2021) were proposed for sustainable WLM. These studies were generally implemented by considering the system's characteristics such as hydraulic, network, customer, environmental and operational characteristics. Therefore, it is important to develop a model that analyzes all the components in WLM and determines the appropriate method for WLM. The water losses couldn't be reduced to zero neither technically nor economically. Leakage reduction generally involve costly and time-consuming processes (Yilmaz et al. 2022). The water losses should be accurately assessed and prioritized (AL-Washali et al. 2020). Water utilities work hard to reduce water losses with high costs. The components should first be identified correctly in order to control water losses. Next, the costs that will be incurred to reduce water loss components should be analyzed (Moslehi et al. 2021). The most appropriate strategy for leakage management should be planned to ensure water and energy efficiency. (Bozkurt et al. 2022a) proposed an assessment system to evaluate the data quality and current implementation level of leakage management components.

Optimization algorithms have been used for urban water management. Molinos-Senante et al. (2014) developed an optimization model that considers the water resources and user profiles in order to maximize the revenue water. The optimization algorithm contributed to the solution of complex WDSs. Techniques (2021) used multivariate techniques and optimization algorithm to provide strategic planning of water resources. The results provide solutions closer to the optimum solution. Monsef et al. (2019) applied the non-dominated sorting genetic algorithm II, multi-objective differential evolution and multi-objective PSO algorithms for the optimum design of WDSs. Mathematical test functions were used to evaluate the performance of algorithms. Shende and Chau (2019) used the simple Benchmarking Algorithm which is a meta-heuristic algorithm, in order to provide reconfigurations in WDS networks with minimum cost. The results showed that SBA is better at achieving the lowest possible cost with rapid convergence.

Bilal and Pant (2020) proposed a combinatorial optimization study in order to optimize the total construction cost of the system. A hybrid metaheuristic algorithm was created by integrating the simulation with EPANET. Khalifeh et al. (2020) optimized the WDS with the HHO algorithm using water demand and storage data. The work provides the optimum network design with lower cost. Ezzeldin and Djebedjian (2020) presented the whale optimization technique for optimum design of WDSs with the lowest cost. The results of optimization were tested using a total of 14 different optimization techniques. Awe et al. (2019) reviewed different optimization methods in WDSs in order to minimize the cost or maximize revenue and system performance. Helmi et al. (2021) examined the problem of restructuring WDSs using the HHO algorithm in order to increase the efficiency of WDSs and to minimize energy losses. A new metaheuristic method using the HHO algorithm is proposed to obtain the optimal configurations.

In this study, a novel WLM model was developed to improve the WLM practices, to define the most appropriate and applicable targets and to determine the components that need to be improved primarily using the Harris Hawk Optimization (HHO) algorithm. The novelty of this study is that the current technical, equipment, financial and equipment capacity and the current implementation levels of WLM practices in the utility are considered. HHO, which is based on the hunting strategies of the Harris Hawk, is an effective method for parameter setting. Thus, the optimization model defines the most appropriate targets based on the current status. The developed model was tested with real data in three utilities. The current status of the utilities is scored between 0 and 5 with CSAS which consists of 8 main categories and 144 components, proposed by (Bozkurt et al. 2022b).

In the literature, water loss reduction and control studies are implemented in pilot areas using different methods within the scope of WLM. However, not every method can be applied in every system or region. Moreover, the same target cannot be set for every utility. Moreover, WLM components cannot be improved at the same level in every utility. The current status of the utility has a direct impact on the implementation and improvement of the components. The model proposed in this study, unlike the literature, considers the WLM components simultaneously and defines the most appropriate target according to the current situation.

In this study, firstly, the current state of the system is analyzed in the proposed algorithmic structure. Then, optimization constraints were determined according to system constraints. Recently, various engineering problems have been solved with stochastic/deterministic optimization algorithms. The model structure in this study is novel and innovative in solving it with a numerical optimization algorithm. Numerical optimization methods can be applied for different engineering problems. The Enhanced equilibrium optimization algorithm is designed with a fractional and integer degree controller (Ates 2021). The optimum power flow problem was solved with the HHO algorithm (Akdağ et al. 2020). Similarly, a controller was designed for the 3 DOF hover system with the modified monarch butter fly optimization algorithm (Ates and Akpamukcu 2022). Finally, controllers were designed for the systems by increasing the performance of optimization algorithms with the optimization to optimization (OtoO) approach (Ates and Akpamukcu 2021). In this study, the parameters for simulation and real-time systems were optimized with the HHO algorithm and the results were presented comparatively.

2 Development of Optimization Model

Water use efficiency, leakage reduction and service quality improvement are critical issues for sustainable water management. Implementation of general approaches for WLM is difficult, costly and time consuming for all networks where the characteristic, environmental, operational and hydraulic behavior of each system differs. Therefore, the methods should be chosen by considering the economic, feasibility and current situation. Moreover, it is quite important to develop a model that considers and analyzes all components-actors and offers the most appropriate alternative. (Bozkurt et al. 2022b) proposed a model that evaluates the current implementation levels of WLM practices. This model includes the matrices of the CCAS, data quality, performance evaluation, and method. The CSAS consisting of 144 components comprising the WLM, evaluates the data quality of components (Bozkurt et al. 2022b). The main components are; Utility Management, Basic Data, Information Systems, Water Balance, Apparent Loss, Real Loss, Performance Monitoring and Economic Analysis (Bozkurt et al. 2022b).

2.1 Objective Function for Optimization Algorithm

There is a need for a model that defines appropriate targets for components by considering the current situation of the utility. In this study, an optimization model was developed to improve the components in CSAS proposed by (Bozkurt et al. 2022b). The most appropriate targets were identified using the HHO algorithm based on the scores. The objective function was created for optimization model. There are a total of 144 components that affect each other in CCAS. It is extremely difficult to develop a strategy with basic algorithmic queries in this space of alternatives. Each component affects system performance at different levels.

The steps of the optimization model are (i) scoring the components in CSAS, (ii) determining the model with the field data, (iii) optimizing the model parameters, and (iv) testing the model accuracy. WLM involves a large number of methods, data and information systems. These components affect each other directly/indirectly. Therefore, a coefficient expressed as the impact level or weight was defined for the components. The CSAS and data matrices in the model proposed by Bozkurt et al. (2022b) were considered. The components that each component in the CSAS affects in the CSAS matrix and data matrices were determined. For example, system flowrate is the main component in WLM. The measurement of this component directly affects the practices of analyzing water balance, performance and efficiency. The impact of this component in WLM model should be different from other components. Therefore, the effects of each component in the CSAS were determined (Table 1). The optimization model was established with this perspective. The data matrix (V_{Mat}) which are associated according to a linear structure, is expressed in Eq. (1).

Table 1 Weights used in the analyzes

AMat	**A1**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8**
1	33	5	107	31	27	19	49	7
2	30	21	19	7	45	124	4	3
3	26	75	79	45	17	113	3	9
4	28	9	24	9	4	93	49	2
5	24	32	63	5	2	34	5	2
6	20	36	39	25	12	1	21	1
7	14	37	52	15	5	29	10	6
8	13	17	4	31	1	9	6	5
9	13	4	6	8	15	12	15	3
10	16	10	43	2	13	2	9	3
11	4	2	6	10	2	12	18	1
12	9	3	1	14	1	9	11	1
13	8	3	66	13	1	20	6	2
14	7	5	14	3	1	3	48	1
15	5	13	4	2	9	3	2	1
16	3	3	20	3	5	3	3	1
17	2	5	5	1	2	2	1	1
18	1	13	34	1	1	9	1	1

$$V_{mat} = \begin{bmatrix} V_{mat(1,1)} & V_{mat(1,2)} & \cdots & V_{mat(1,i)} \\ V_{mat(2,1)} & V_{mat(2,2)} & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ V_{mat(18,1)} & V_{mat(18,1)} & \cdots & V_{mat(j,i)} \end{bmatrix} \text{i} = 1, 2, ..8 \;\; \text{j} = 1, 2, 3 \ldots 18 \tag{1}$$

There is a linear relationship between the components in the CCAS by considering this impact level and logical structure. A linear model is proposed based on the linear relationship. The efficiency of the linear model is increased with weights (Eq. (2)).

$$a_{mat} = \begin{bmatrix} a_{mat(1,1)} & a_{mat(1,2)} & \cdots & a_{mat(1,i)} \\ a_{mat(2,1)} & a_{mat(2,2)} & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ a_{mat(18,1)} & a_{mat(18,1)} & \cdots & a_{mat(j,i)} \end{bmatrix} \text{i} = 1, 2, ..8 \;\; \text{j} = 1, 2, 3 \ldots 18 \tag{2}$$

a_{mat} : the weight matrix. The system matrix (*SMat*) is formed by combining these matrices and performing mathematical operations. The calculation process of the objective function starts in the optimization processes. The coefficient matrix is optimized according to the calculation of the objective function. The values that the result matrix can take are defined according to the level of the variable.

$$\text{SMat} = \sum_{i=1}^{144} AMat(i)VMat(i) \tag{3}$$

The constraints for optimization model are shown in Table 2. The optimization algorithm searches the parameter space according to its search terminology. The results are checked at each iteration. The algorithm proceeds to the next iteration whether the result is accepted or not according to the suitability of these conditions.

The difficulties, impact levels, requirements, and costs of the components in CCAS were considered in defining the conditions. The scores in CCAS are defined between 0

Table 2 The constaints for optimization model

Level	Current Score	Result Matrix
The first group (components 1-6 in the CSAS	0,1 and 2	≤ 4
	3 and 4	≥ current score
The second group (components 7-12 in the CSAS	0,1 and 2	≤ 4
	3	≥3
	4	≥4
The third group (components 13-18 in the CSAS	0 and 1	≤ 3
	2 and 3	≥3
	4	≥4

and 5 in the design of the mathematical form. Bringing each component to the value of 5 may not be economically and technically possible under the current conditions.

A certain target for each utility is determined according to its current value. The configurations are calculated with optimization. Integer values between 0 and 5 are assigned recursively for each cell in the CCAS. Then, the objective function is calculated by considering the initial value of the component.

The current score and the value assigned by the algorithm is compared. If the current score is 0 or 1 and the value assigned by the algorithm is 2 or 3. Smat is multiplied by the transition1. If the current score is 0 or 1 is the value assigned by the algorithm, the Smat is multiplied by the transition1+ transition2. If the current score is 0 and 1, the value of the algorithm is 5, then the Smat is multiplied by the transition1+ transition2+ transition3.

In this way, in the current status, the transition from 0 and 1 to 2 or 3, from 2 or 3 to 4, and from 4 to 5 is planned to be unequal. Due to the nature of the problem, it is more difficult to bring a component from 4 to 5 than to bring the component from 0 or 1 to 2 or 3. The transitions in the model, also contribute to defining the difficulty level in the nature of this problem. The coefficients of transitions (transition_1=1 transition_2=3 and transition_3=5 was determined as a result of continuous tests. The objective function (AF) used during the optimization was calculated by Eq. (4).

$$\text{AF} = (\text{HD} - \sum_{i=1}^{144} \left((tran_1 + tran_2 + tran_3)AMat(i)VMat(i)\right) / \sum_{i=1}^{144} AMat(i)) \tag{4}$$

HD: Target value, tran_1= transition_1, tran_2= transition_2 and tran_3= transition_3.

In each iteration of the optimization, a candidate solution is proposed. AF is calculated according to these structures. Then, the compatibility of the proposed solution is tested based on the conditions given in Table 2. If these conditions are met, the algorithm accepts this solution, if it does not, it passes to another parameter vector space.

The following algorithmic structure was used to determine the transitions during the optimization. YVMat(i) is the new value suggested by the optimization algorithm. If the current score is 0, 1, 2, 3, the function given below is used to determine the transition.

$$G(tran_1, tran_2, tran_3) = \left\{ \begin{array}{l} tran_1 = 0; tran_2 = 0; tran_3 = 0; (YVMat(i) = 0 \,||\, YVMat(i) == 1) \\ tran_1 = 1; tran_2 = 0; tran_3 = 0; (YVMat(i) = 2 \,||\, YVMat(i) == 3) \\ tran_1 = 1; tran_2 = 3; tran_3 = 0; (YVMat(i) = 4) \\ tran_1 = 1; tran_2 = 3; tran_3 = 5; (YVMat(i) = 5) \end{array} \right\}$$

$$G(tran_1, tran_2, tran_3) = \left\{ \begin{array}{l} tran_1 = 0; tran_2 = 0; tran_3 = 0; (YVMat(i) = 0 \,||\, YVMat(i) == 1) \\ tran_1 = 0; tran_2 = 0; tran_3 = 0; (YVMat(i) = 2 \,||\, YVMat(i) == 3) \\ tran_1 = 0; tran_2 = 3; tran_3 = 0; (YVMat(i) = 4) \\ tran_1 = 0; tran_2 = 3; tran_3 = 5; (YVMat(i) = 5) \end{array} \right\}$$

If the current score is 4, the function given below is used to determine the transition coefficient.

$$\text{G}(\text{tran}_1,\text{tran}_2,\text{tran}_3) = \left\{ \begin{array}{l} \text{tran}_1 = 0;\ \text{tran}_2 = 0;\text{tran}_3 = 0;\ (\text{YVMat(i)} = 0 \,\|\, \text{YVMat(i)} == 1) \\ \text{tran}_1 = 0;\ \text{tran}_2 = 0;\text{tran}_3 = 0;\ (\text{YVMat(i)} = 2 \,\|\, \text{YVMat(i)} == 3) \\ \text{tran}_1 = 0;\ \text{tran}_2 = 0;\text{tran}_3 = 0;\ (\text{YVMat(i)} = 4) \\ \text{tran}_1 = 0;\ \text{tran}_2 = 0;\text{tran}_3 = 5;\ (\text{YVMat(i)} = 5) \end{array} \right\}$$

A current score of 5 indicates that there is no need for optimization. The transitions are transition1=0; pass2=0; pass3=0; taken in the form. It is not possible for any value to go back during the optimization process for the real-time model of the system. Transitions are assigned as (pass1=0; pass2=0; pass=0) so that the back-off states do not get stuck in local points of the algorithm in order to provide this terminology.

2.2 Optimization of WLM Parameters with HHO Algorithm

The HHO algorithm is an optimization algorithm proposed by Heidari, inspired by the behavior of Harris hawks, which are good hunters in nature (Heidari et al. 2019; Akdağ et al. 2020; Akdag 2021). Harris hawks called an intelligent swarm-based algorithm, act as a flock while hunting rabbits. There are various phases that are the discovery, the transition from exploration to attack, and the attack. The optimization of the Vmat was made with objective function and constraints.

Discovery phase shows the initial solution of the algorithm. Harris hawks wait in the desert and are constantly observing. These observers are part of the flock. Others await the next mission if there is hunt. It is tried to find the most suitable points by making some kind of random suggestions (Heidari et al. 2019; Akdağ et al. 2020).

$$V_{mat}(t+1) = \begin{cases} V_{mat_{rand}}(t) - r_1 \left| V_{mat_{rand}}(t) - 2r_2 V_{mat}(t) \right|, \\ (V_{mat_{rabbit}}(t) - xV_{mat_m}(t)) - r_3(AL - r_4(UL - AL)), \; q < 0.5 \end{cases} \tag{5}$$

$V_{mat}(t+1)$: the position vector, $V_{mat_{rabbit}}(t)$: the hunt position vector, $V_{mat_{rand}}(t)$: the hawk randomly selected hawk at the current population, $V_{mat_m}(t)$: the average position of current hawk position, $V_{mat}(t+1)$: the current position of hawk, (r_1,r_2,r_3,r_4 ve q) the random numbers (0,1). AL, and ÜL: the lower and upper values, respectively (Heidari et al. 2019; Akdağ et al. 2020).

$$Vmat_m(t) = \frac{1}{N} \sum_{i=1}^{N} Vmat_i(t) \tag{6}$$

The transition from exploration to attack: Harris hawks can switch to different attack styles according to the energy of the hunt. After the hunt is detected, the task of a group of hawks in the flock is to reduce the energy of the hunt.

$$E = 2E_0\left(1 - \frac{t}{T}\right) \tag{7}$$

E: energy of escaped hunt, E_0 : initial energy of hunt, N; the number of hawks, T: maximum number of iterations.

The Attack is determined according to the strategies of soft, hard, soft with progressive rapid dives, and hard with progressive rapid dives (Heidari et al. 2019; Akdağ et al. 2020).

Soft surround: The Harris hawk aims to reduce the energy of the hunt with sudden attacks. ($r \geq 0.5, E \geq 0.5$).

$$\mathrm{Vmat(t+1)} = \Delta\mathrm{Vmat(t)} - \mathrm{E}\left|\mathrm{JVmat_{rabbit}(t)\text{-}Vmat(t)}\right| \tag{8}$$

$$\Delta\mathrm{Vmat(t)} = \mathrm{Vmat_{rabbit}\text{-}Vmat(t)} \tag{9}$$

Hard surround: It is a situation where the energy of the hunt is considerably reduced. ($r \geq 0.5, |E| \leq 0.5$).

$$\mathrm{DVmat(t+1)} = \mathrm{Vmat_{rabbit}(t)} - \mathrm{E|DVmat(t)|} \tag{10}$$

Soft surround with progressive rapid dives: Hunt has enough energy. The Harris hawk aims to lower the energy of the hunt with sudden attacks.

$$\mathrm{Y=Vmat_{rabbit}}(t)\text{-}\mathrm{E|\Delta Vmat}(t)|\text{-}\mathrm{Vmat}(t) \tag{11}$$

$\Delta \mathrm{Vmat}(t)$: the difference between the current position in t^{th} iteration and the current position of the hunt. It is compared to previous dives to check if the move is a good move. If not suitable, sudden dives are continued. Levy flight distribution structure is used in the decision-making process.

$$\mathrm{Z=Y+S*LF}(D) \tag{12}$$

D: the problem size, S is a random vector of size 1xD, LF: the levy functions.

$$\mathrm{LF(x)} = 0.001\mathrm{x}\left(\frac{\mu \mathrm{x}\sigma}{|\mu|^{\frac{1}{\beta}}}\right), \sigma = \left[\frac{\Gamma(1+\beta)\mathrm{x}\sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\frac{(1+\beta)}{2}\mathrm{x}\beta\mathrm{x}2^{\left(\frac{\beta-1}{2}\right)}}\right] \tag{13}$$

u, v : the random number (0,1), β : 1.5. Equation (15) is used to update the positions of the hawks in the soft surround.

$$\mathrm{V_{mat}(t+1)} = \left\{\begin{array}{l} \mathrm{Y\ if\ F(Y) < f(V_{mat}(t))} \\ \mathrm{Z\ if\ F(Z) < F(V_{mat}(t))} \end{array}\right\} \tag{14}$$

Hard surround with progressive rapid dives: The hunt does not have enough energy. Because of this, the Harris hawk makes a tough surround.

$$\mathrm{V_{mat}(t+1)} = \left\{\begin{array}{l} \mathrm{Y\ if\ F(Y) < f(V_{mat}(t))} \\ \mathrm{Z\ if\ F(Z) < F(V_{mat}(t))} \end{array}\right\} \tag{15}$$

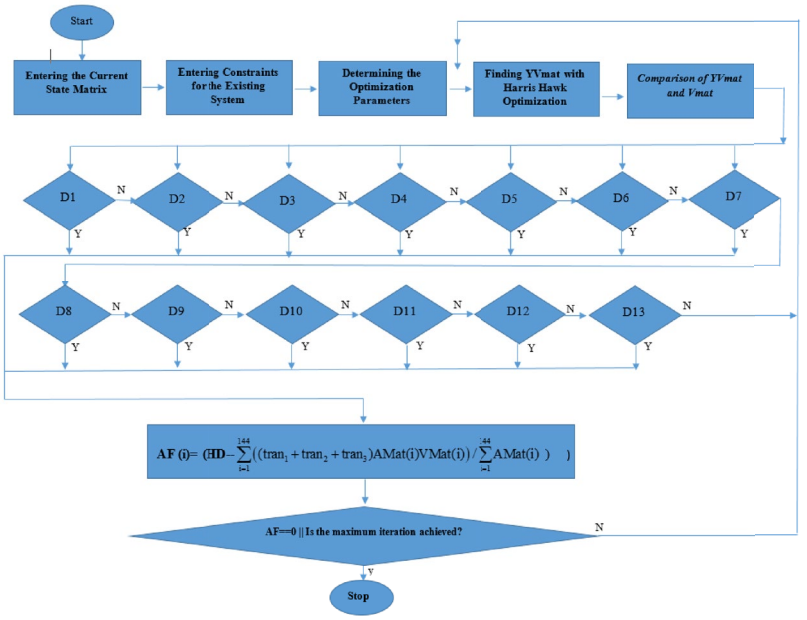
$$\mathrm{Y = Vmat_{rabbit}(t) - E|\Delta JVmat(t)| - Vmat(t)} \tag{16}$$

$$\mathrm{Z = Y + SXLF(D)} \tag{17}$$

2.3 Finding the $\mathbf{V_{mat}}$ Matrix with the HHO Algorithm

The following steps were applied to optimize the Vmat matrix with the HHO algorithm (Fig. 1).

Step 1: VMat are entered into the algorithm. The target, constraints and transitions (transition1=1; transition2=3; transition3=5) are determined.
Step 2: The parameters of N, T, AL and ÜL are defined.



State 1 (D1)	(Vmat(i)==0 || Vmat(i)==1) && (YVMat(i)=0 || YVMat(i)==1) transition1=0; transition2=0; transition0;
State 2 (D2)	(Vmat(i)==0 || Vmat(i)==1) && (YVMat(i)=2 || YVMat(i)==3) transition1=1; transition2=0; transition0;
State 3 (D3)	(Vmat(i)==0 || Vmat(i)==1) && (YVMat(i)=4) transition1=1; transition2=3; transition0;
State 4 (D4)	(Vmat(i)==0 || Vmat(i)==1) && (YVMat(i)=5) transition1=1; transition2=3; transition5;
State 5 (D5)	(Vmat(i)==2 || Vmat(i)==3) && (YVMat(i)=0 || YVMat(i)==1) transition1=0; transition2=0; transition0;
State 6 (D6)	(Vmat(i)==2 || Vmat(i)==3) && (YVMat(i)=2 || YVMat(i)==3) transition1=0; transition2=0; transition0;
State 7 (D7)	(Vmat(i)==2 || Vmat(i)==3) && (YVMat(i)=4) transition1=0; transition2=3; transition0;
State 8 (D8)	(Vmat(i)==2 || Vmat(i)==3) && (YVMat(i)=5) transition1=0; transition2=3; transition5;
State 9 (D9)	(Vmat(i)==4) && (YVMat(i)=0 || YVMat(i)==1) transition1=0; transition2=0; transition0;
State 10 (D10)	(Vmat(i)==4) && (YVMat(i)=2 || YVMat(i)==3) transition1=0; transition2=0; transition0;
State 11 (D11)	(Vmat(i)==4) && (YVMat(i)=4) transition1=0; transition2=0; transition0;
State 12 (D12)	(Vmat(i)==4) && (YVMat(i)=5) transition1=0; transition2=0; transition5;
State 13 (D13)	(Vmat(i)>YVMat(i)) transition1=0; transition2=0; transition0;

Fig. 1 Flow chart for proposed model and state definition

Step 3: The location of each hawk is determined in each iteration according to N. Vmat matrix is created according to the constraints.

$$\sum_{i=1}^{144}\left((\text{tran}_1+\text{tran}_2+\text{tran}_3)\text{AMat(i)VMat(i)}\right)/\sum_{i=1}^{144}\text{AMat(i)})$$

Step 4: The new value of the Vmat matrix is determined by considering the current situation.

Step 5: The position of the rabbit corresponding to minimizing the best Objective Function intended is calculated.

Step 6: The initial energy of the hunt is compared with the fitness. If the fitness is smaller than the initial energy; the initial energy of the hunt is equalized to the fitness.
Step 7: The energy of the hunt is modeled.
Step 8: When E is greater than or equal to 1, the position vector is updated using Eq. 1.
Step 9: If E is less than 1, whichever of the following situations is provided, go to that step.
Step 9.1: When (|E| ≥0.5andr ≥ 0.5) the position is updated using Eq. 4.
Step 9.2: ((|E| ≥ 0.5 and r < 0.5) the position is updated using Eq. 6.
Step 9.3: (|E| < 0.5 and r ≥ 0.5) the position is updated using Eq. 10.
Step 9.4: (|E| < 0.5 and r < 0.5) the position is updated using Eq. 11.
Step 10: Step 5 is repeated until the stopping criteria are reached.
Step 11: Find the location of the prey and its fitness.

3 Simulation Results of HHO Algorithm

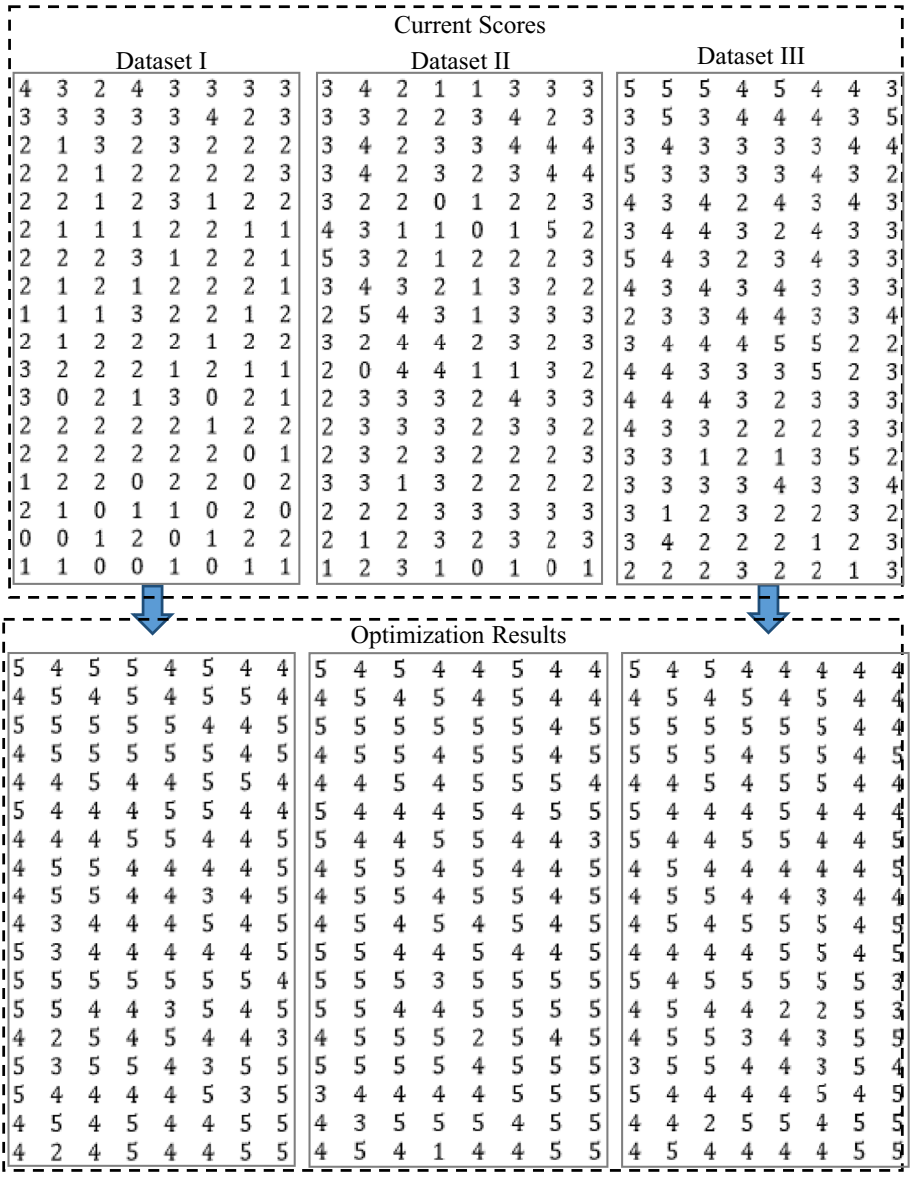
3.1 Testing the Algorithm with Sample Data

The most appropriate target for the utility is suggested in mathematical form according to the current situation. Sample data sets were produced in accordance with the score logic of the basic, moderate and advanced components in the CCAS (Fig. 2). Then, the algorithm was tested with these data sets. The components in the CSAS are divided into three subgroups that are the basic, moderate and advanced level, based on the technical and economic requirements for the implementation of the components. Each component in the CSAS was scored to create a sample data set based on these criteria. These data do not reflect the situation of any utility. It was created solely for the purpose of developing and testing the model.

The components at basic level should be performed in utilities. The implementation of moderate level components is generally more difficult than basic level. Moreover, the advanced level components are more costly and time-consuming than basic and moderate levels. Accordingly, it was predicted that (i) basic level scores were higher than moderate and advanced level, (ii) moderate level scores could be higher than advanced level. The targets were defined for each data matrix by considering the arithmetic average of scores.

The target provides the current average of the utility to a higher score class. Sample datasets were created as poor, moderate and good performance. The current scores and target for Data Set-I are 1.7222 and 2-3, respectively. The current scores and target for Data Set-II are 2.4931 and 4-5, respectively. The current scores and target for Data Set-III are 3.1667 and 4-5, respectively. The parameters were optimized in order to increase the score of the datasets. Random experiments are implemented to determine the target. If the score reaches the target, this is the sought value. If the target cannot be reached, the experiment is continued.

The parameters of N, T and target are defined based on the CCAS. After determining the targets, the target is kept constant and appropriate values of N and T are determined. Since there are 144 components in the matrix, the N must be at least 144. After the target and N are determined, these values are taken as constant. The tests are continued by changing the T parameter. The current situation is analyzed in low and high iterations. Appropriate iteration intervals are determined for each data set (Table 3). The purpose here is; (i) to determine the optimal range of target, N and T in order to increase the current score to the targe, (ii) to develop a mathematical model that reveals the relationship between the current and the target scores.



a)

Fig. 2 **a** The result matrices for sample data sets defined by HHO algorithm **b** Current scores and optimization results in pilot utilities

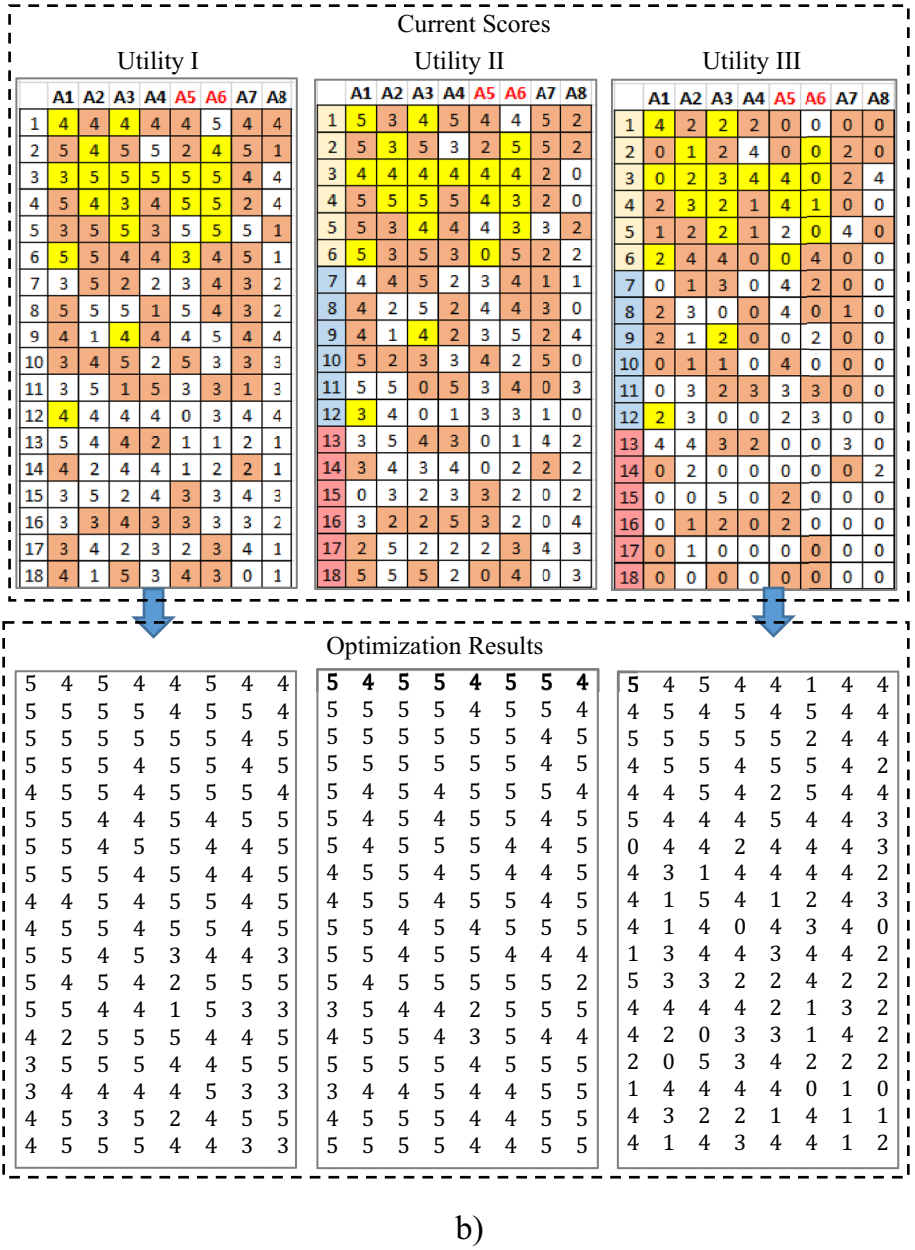


Fig. 2 (continued)

Targeto; the target for dataset, Current_mean; the current mean score of datasets; A_mean_new; the new average score defined by optimization. Thus, the data sets (Fig. 2) were analyzed with the HHO algorithm (Fig. 3). The algorithm updates the parameters and

Table 3 The optimization parameters and targets for sample data sets and field data sets

Sample Data Sets								
Data Set	N	T	Targeto	Current_mean	Basic_level_mean	Moderate_level_mean	Advanced_level_mean	A_mean_new
I (poor)	150	5000	0.25	1.7222	4.3542	4.2500	4.3819	3.0208
II (moderate)	500	5000	0.25	2.4931	4.5417	4.5208	4.4583	4.5069
III (good)	500	5000	0.25	3.1667	4.4583	4.4167	4.1667	4.3472
Field Data Sets								
Utility	N	T	Targeto	Current_mean	Basic_level_mean	Moderate_level_mean	Advanced_level_mean	A_mean_new
I	150	5000	0.5	3.4167	4.6667	4.4792	4.0833	4.4097
II	4000	100	0.5	2.9792	4.7083	4.1667	4.3125	4.3958
III	500	100	0.07	1.1597	4.1667	2.9792	2.5417	3.2292

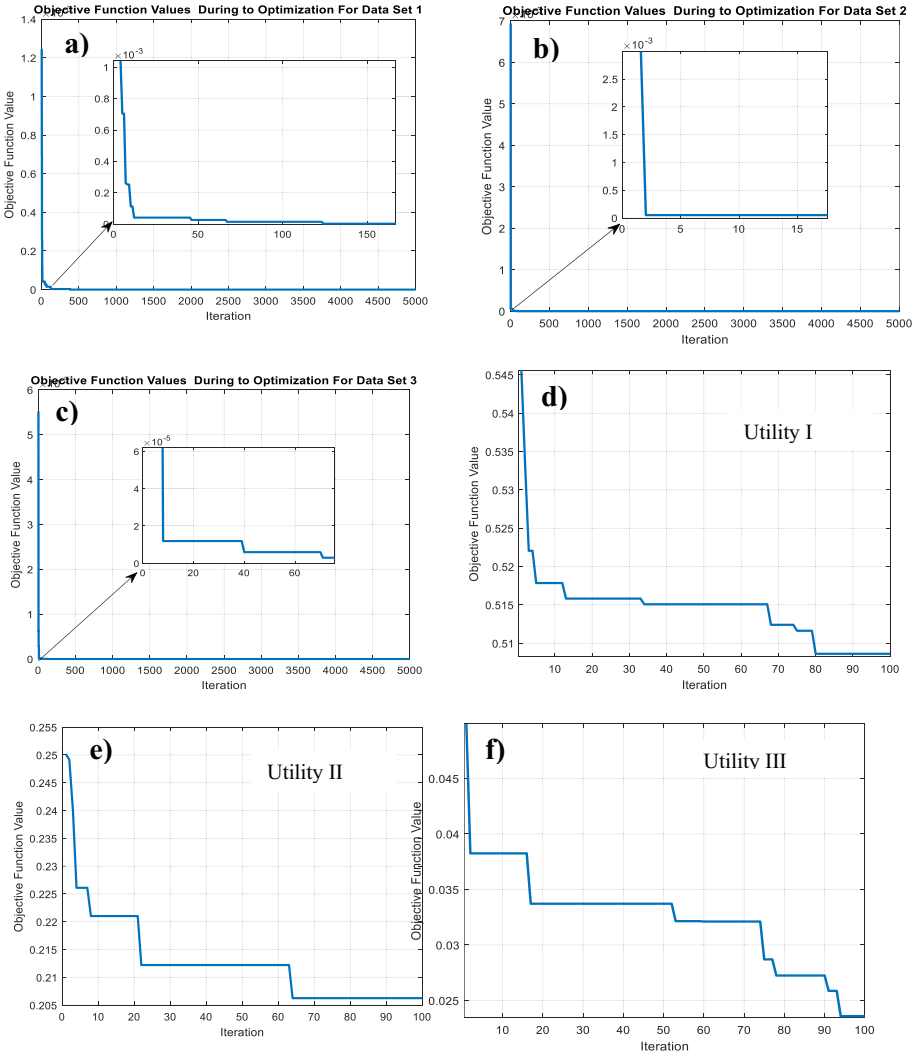


Fig. 3 **a** Objective function values during to optimization for data sets1 **b** objective function values during to optimization for data sets2 **c** objective function values during to optimization for data sets3 **d** objective function values during to optimization for utilities 1 **e** objective function values during to optimization for utilities 2 **f** objective function values during to optimization for utilities 3

reaches the optimal value even if the sample data does not work. The error function was minimally solved to achieve the goal, considering the system dynamics.

3.2 Study Area

Three utilities in Turkey (Utility I: Bursa, Utility II: Kayseri, Utility III: Malatya) are the study area (Fig. 4) to test the optimization model. Utility I located in the west of Turkey is the fourth most populous city with a population of 3,139,744. The number of customers is 1,250,000. The network length is approximately 7,100 km. Utility II located in the Central

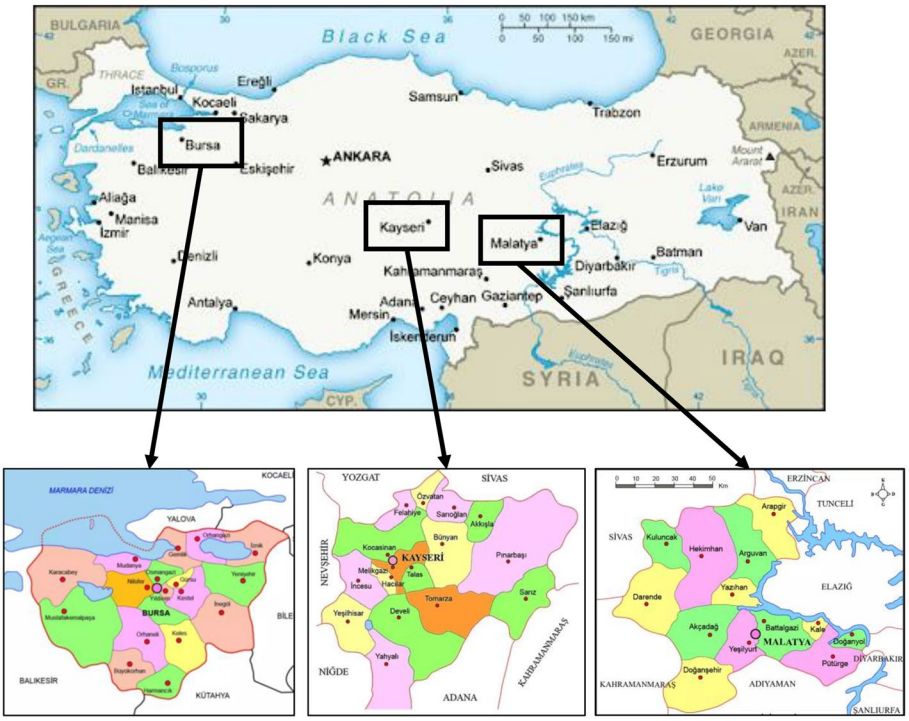


Fig. 4 Study area

Anatolia region has a population of 1,434,357, and customers of 550,000. The network length is approximately 12,321 km. Utility III located in the east of Turkey, has a population of 808,692, with a total customer of 350,000. The network length is approximately 2,500 km.

3.3 Optimization Results for Field Data Matrices

The current status of the utilities was defined with scoring system and CSAS (Fig. 2). To make an objective evaluation, the components in this matrix are scored between 0 and 5 by experts outside the institution. The departments of water management, information technology, customer management, Geographic Information System and SCAD were visited in in the utilities. Data, reports, activity documents and databases in the utility are taken as basis for scoring each component.

The utilities have different score classes. It is necessary to determine the appropriate target for each utility. The model was run using the component's scores in utilities. Targets were found according to the current average of utility. After the basic components reach to the target, the other levels are improved. Thus, the gradual structure was dynamically defined. The result matrices for the utilities are given in Fig. 3.

3.3.1 Evaluation for Utility I

The current status is at good level. The appropriate target and new average are 0.8 and 4.4097, respectively. The HHO algorithm has optimized 144 parameters by reducing the error function.

The basic components currently have score of 2 or higher. Three components at moderate level are 0 and 1. Seven components at advanced level are 0 or 1. The components with a high weight and the components with 0 and 1 were improved. Then, the improvement levels of moderate components were determined by the model according to the current score. The new scores recommended for the moderate components in this utility are at higher level than in other utilities. The main reason is that the current scores in the utility are at better than other utilities. Thus, the current scores of the utility are considered by the model. Similarly, the new scores recommended by the optimization are higher than other utilities.

3.3.2 Evaluation for Utility II

The appropriate target and new average are 0.5 and 4.3958, respectively. A consistent optimization process was obtained to reach the target by the HHO. Three components at basic level, seven components at moderate level and eight components at advanced level are 0 and 1 in current status. The components with a high weight and the components with 0 and 1 were improved by the model. The components in the basic level are generally suggested as good or quite good in result matrix. The current scores of basic components should meet the necessary conditions in order to apply the components in moderate and advanced level. It is thought that the target scores suggested by the model are in accordance with the natural structure of the problem.

3.3.3 Evaluation for Utility III

The appropriate target and new average are 0.07 and 3.2292, respectively. A total of 22 components at basic level, 30 components at moderate level and 37 components at advance level are 0 or 1. These components were firstly improved in result matrix. Since the current scores in the utility are generally at poor level. The improvements suggested by the model were more limited than other utilities. The new targets for moderate and advanced components include gradual improvement based on current scores.

3.4 Definition of the Current Situation and Target Relationship

In the previous sections, the relationship between the current average and target values obtained was presented. The optimization model was run a large number of times to define the relationship between the current situation and the target score (Table 3).

The table presents the current score, target mean value, targeto value, and A_mean_new values representing the new mean in the result matrix.

A control matrix (comparison matrix) was established to check the optimization conditions. Moreover, a control counter was added to the algorithm. Analysis is carried out by defining a value for the control counter, depending on how many of the conditions are desired to be met. For example, utility's current score is 3.416 and the target average is between 4-5 according to the score classes.

The target, which will increase the average of this utility to the band of 4-5, was found as 0.8. In this way, the targets that will increase the current average of each utility to the target average were found by testing and the diagram given above was obtained.

The target averages differ since the current situation of each utility is at different levels. A curve was fitted with linear regression to the target values obtained by the tests. This equation shows the relationship between the current mean and target value. The accuracy of the obtained equation is high. In this way, the target can be calculated for the utility based on the current score and the optimization algorithm works by targeting this value.

As a result, the current situation was analyzed by applying the scoring system and an optimization model was proposed according to the dynamic structure of utility. The priority components in improvement were determined. This novel model will contribute to the prevention of time-consuming and costly processes and reduction of construction costs by determining the most appropriate method.

4 Conclusion

A novel optimization model was proposed using the HHO algorithm to define the most appropriate strategy based on the CSAS. A unique scoring system was applied in three utilities. The current status of utilities I, II and III were at good, moderate and poor level, respectively. The developed model was first tested with virtual datasets and then tests were conducted with real datasets. Since each utility's current score was at different level at the beginning, a dynamic improvement system was used for each utility's target. The current scores in utility I were 4.0625, 3.3958, and 2.7917, respectively. In utility I, the optimization results for basic, moderate and advanced levels were 4.6667, 4.4792 and 4.0833, respectively. The current scores in utility II were 3.5417, 2.9717, 2.6042, respectively. In utility II, the optimization results for basic, moderate and advanced levels were 44.7083, 4.1667 and 4.3125, respectively. Finally, the current scores of components in utility III were 1.6042, 1.1875, 0.6875, respectively. In utility III, the optimization results for basic, moderate and advanced levels were 4.167, 2.9792 and 2.5417, respectively. Optimization results were supported by error function graphs according to the target value being reached. The model and algorithmic structure presented in this way are valid for the proposed model and that results consistent with the field can be produced.

It is very important to define the most appropriate management strategy, since WLM involves quite complex and costly processes. This model proposed in this study provides important gains for technical personnel in utilities, in terms of planning WLM activities and identifying the components that need to be improved primarily. Additionally, it will contribute to practitioners in terms of revealing the most appropriate improvement targets on a system-wide or component basis.

The optimization model suggested the graded scores according to the current score and dynamic structure of the utility. As a result, the model developed within the scope of the study suggested new levels in accordance with the nature of the problem, considering the current scores of the components in utilities.

Author Contributions Cansu Bozkurt, Abdullah Ates, Mahmut Fırat wrote the main manuscript. Cansu Bozkurt and Abdullah Ates coded optimization algorithm. Salih Yılmaz and Ozgur Ozdemir tested proposed optimization via experimental system.

Funding Open access funding provided by the Scientific and Technological Research Council of Türkiye (TÜBİTAK). This research was supported by TUBITAK under the Project Number 220M091.

Availability of Data and Material All data and models that support the findings of this study are available from the corresponding author upon reasonable request.

Code Availability All codes that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethical Approval and Informed Consent Not applicable.

Consent to Participate Not applicable.

Consent to Publish The authors are indeed informed and agree to publish.

Conflict of Interest Authors declare that we have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Akdag O (2021) Modification of Harris hawks optimization algorithm with random distribution functions for optimum power flow problem. Neural Comput App 33:1959–1985

Akdağ O, Ateş A, Yeroğlu C (2020) Harris Şahini Optimizasyon Algoritması ile Aktif Güç Kayıplarının Minimizasyonu Minimization of Active Power Losses Using Harris Hawks Optimization Algorithm. Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen Ve Mühendislik Dergisi 22:481–490

AL-Washali T, Sharma S, Lupoja R et al (2020) Assessment of water losses in distribution networks: Methods, applications, uncertainties, and implications in intermittent supply. Resources, Conserv Recycling 152

Ates A (2021) Enhanced equilibrium optimization method with fractional order chaotic and application engineering. Neural Comput App 33:9849–9876

Ates A, Akpamukcu M (2021) Optimization to optimization (OtoO): optimize monarchy butterfly method with stochastics multi - parameter divergence method for benchmark functions and load frequency control. Eng Comp 38:1735–1754

Ates A, Akpamukcu M (2022) Modified monarch butterfly optimization with distribution functions and its application for 3 DOF Hover flight system. Neural Comput Appl 1–26

Awe OM, Okolie STA, Fayomi OSI (2019) Optimization of Water Distribution Systems: A Review. J Phys Conf Ser 1378

Azevedo BB, Saurin TA (2018) Losses in Water Distribution Systems: A Complexity Theory Perspective. Water Resources Manag 32:2919–2936

Babu KSJ, Vijayalakshmi DP (2013) Self-adaptive PSO-GA hybrid model for combinatorialwater distribution network design. J Pipeline Sys Eng Prac 4:57–67

Bilal Pant M (2020) Parameter optimization of water distribution network–a hybrid metaheuristic approach. Mater Manuf Proc 35:737–749

Bozkurt C, Firat M, Ates A (2022a) Development of a new comprehensive framework for the evaluation of leak management components and practices. AQUA - Water Infrastructure, Ecosys Soc 71(5):642–663

Bozkurt C, Firat M, Ateş A (2022b) Strategic water loss management : Current status and new model for future perspectives. Welcome Sigma J Eng Natural Sci 40:1–13

Ezzeldin RM, Djebedjian B (2020) Optimal design of water distribution networks using whale optimization algorithm. Urban Water J 17:14–22

Farah E, Shahrour I (2017) Leakage detection using smart water system: Combination of water balance and automated minimum night flow. Water Resources Manag 31:4821–4833

Fontanazza CM, Notaro V, Puleo V, Freni G (2015) The apparent losses due to metering errors: a proactive approach to predict losses and schedule maintenance. Urban Water J 12:229–239

Gupta A, Kulat KD (2018) A selective literature review on leak management techniques for water distribution system. Water Resources Manag 32:3247–3269

Heidari AA, Mirjalili S, Faris H et al (2019) Harris hawks optimization: Algorithm and applications. Future Generation Comp Sys 97:849–872

Helmi AM, Carli R, Dotoli M, Ramadan HS (2021) Harris hawks optimization for the efficient reconfiguration of distribution networks. 2021 29th Mediterr Conf Control Autom MED 214–219

Khalifeh S, Akbarifard S, Khalifeh V, Zallaghi E (2020) MethodsX Optimization of water distribution of network systems using the Harris Hawks optimization algorithm (Case study : Homashahr city). MethodsX 7

Marzola I, Alvisi S, Franchini M (2021) Analysis of MNF and FAVAD models for leakage characterization by exploiting smart-metered data: The case of the gorino ferrarese (fe-Italy) district. Water (Switzerland) 13:643

Molinos-Senante M, Hernández-Sancho F, Mocholí-Arce M, Sala-Garrido R (2014) A management and optimisation model for water supply planning in water deficit areas. J Hydrol 515:139–146

Monsef H, Naghashzadegan M, Jamali A, Farmani R (2019) Comparison of evolutionary multi objective optimization algorithms in optimum design of water distribution network. Ain Shams Eng J 10:103–111

Moslehi I, Jalili-Ghazizadeh M, Yousefi-Khoshqalb E (2021) Developing a framework for leakage target setting in water distribution networks from an economic perspective. Struct Infrastructure Eng 17:821–837

Mutikanga HE, Sharma SK, Vairavamoorthy K (2013) Methods and tools for managing losses in water distribution systems. J Water Resources Plan Manag 139:166–174

Pearson D (2019) Standard Definitions for Water Losses. IWA Publishing, London, UK

Shende S, Chau KW (2019) Design of water distribution systems using an intelligent simple benchmarking algorithm with respect to cost optimization and computational efficiency. Water Sci Technol Water Supply 19:1892–1898

Yilmaz S, Firat M, Ates A, Özdemir Ö (2022) Analyzing the economic water loss level with a discrete stochastic optimization algorithm by considering budget constraints. AQUA - Water Infrastructure, Ecosys Soc 71(7):835–848

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.