Annals of Nuclear Energy 163 (2021) 108545

Contents lists available at ScienceDirect

Annals of Nuclear Energy

journal homepage: www.elsevier.com/locate/anucene

# Introducing optimum parameters of separation cascades for $^{123}$Te using GWO based on ANN

Morteza Imani, M. Aghaie *, Mohammademad Adelikhah

*Engineering Department, Shahid Beheshti University, G.C., P.O. Box: 1983963113, Tehran, Iran*

ARTICLE INFO

*Article history:*
Received 14 March 2021
Received in revised form 28 June 2021
Accepted 3 July 2021
Available online 15 July 2021

*Keywords:*
Square cascade
Separation
$^{123}$Te
Gray wolf algorithm
Artificial Neural network

ABSTRACT

In recent decades, separation of stable isotopes due to their substantial role in human health has been widely increased. The present research deals with square cascades optimization in order to separate the $^{123}$Te by the Gray Wolf Optimization algorithm (GWO). The separation of $^{123}$Te has significant application in medical science, and production of radioisotopes. In this study, attempts have been made to find the desired concentration of product (99.9%) for a given amount of natural Tellurium feed within four connected cascades. In this analysis, instead of solving nonlinear equations of concentration distribution in cascades, two different artificial neural networks (ANN) are trained to predict the objective functions. Two test cases for $^{123}$Te separation with different objective functions have been considered. The aim is to gain the maximum product from a specified amount of feed in different configurations. In the first case, the neural network has 20 inputs and considers four connected cascades. To train the network, 5000 randomly generated data from the results is used. In the second case, the network has 22 inputs and 10,000 random data is used. In both cases, the Levenberg-Marquardt algorithm with 40 hidden layers is selected to train the networks. Prediction of the objective functions using a neural network leads to a 98% reduction in execution time and significantly improves the speed of the optimization process. Using this method, the optimal cascades for separation of $^{123}$Te with 99.9% concentration from 15 kg of natural Tellurium during a year are introduced.



## 1. Introduction

Today, the efficient methods for extracting the stable isotopes from natural resources is under development. The stable isotopes are widely used in medical sciences. One of the efficient methods for separation of stable isotopes is the use of Gas Centrifuge (GC) cascades (Zeng and Ying, 2000a). Tellurium consists of eight stable isotopes with different applications. For example, $^{123}$Te is used in $^{123}$I radioisotope production, which can be applied in life science for healthcare, medical applications, imaging and pharmaceuticals industries (Smirnov and Sulaberidze, 2014). Taper and square cascades can be used to separate multicomponent isotopes by GCs. In taper cascades, using R, Q and Quasi ideal models, a cascade can be designed that is used to separate a special isotope (Zeng et al., 2011). This cascade cannot be used to separate other isotopes, unless redesign the cascade structure. Therefore, it is not economically feasible to design a separation cascade capable of separating the isotopes, individually. Accordingly, the design and construction of cascades in commercial applications is going to the direction that separation of a wide range of stable isotopes would be possible by a single cascade. Nowadays, square cascades are used for this purpose (Zeng et al., 2014). Due to the recycle flows in the first and last stages, square cascades can be used in low-feed flow rates and different cuts, which have a high flexibility in operation and can be adjusted by the control valves between the stages of the cascade. Separation of isotopes using a square cascade is associated with great complexity. Determination of optimum parameters to reach the desired concentration is not easy. For example, for a square cascade, the effective parameters are the cascade feed flow rate, feed location, cascade cut, feed flow of GCs and cut of the first stage. Changing the parameters, different concentrations of the desired isotope can be obtained. It is not possible to reach a definite theoretical solution method to determine the optimal process for separating of each isotopes regarding the effective parameters, but optimization algorithms can be used to attain the desired goal (Ying and Zeng, 2015).

In this paper, the enrichment of $^{123}$Te (middle isotope of the Tellurium) has been investigated using square cascades of GCs. To obtain the best-connected square cascades, the process parameters, which include cascade feed flow rate, feed location, cascade cut, feed flow of GCs and cut of the first stage have been studied

* Corresponding author.

https://doi.org/10.1016/j.anucene.2021.108545
0306-4549/

**Nomenclature**

| | |
|---|---|
| $C$ | Concentration |
| $C^{'}$ | Product concentration |
| $C^{''}$ | Waste concentration |
| *Feed* | Feed flow rate of cascade |
| $F$ | External feed flow of each stage |
| $L$ | Feed flow |
| $L^{'}$ | Product flow |
| $L^{''}$ | Waste flow |
| $M$ | Molecular weight |
| $N$ | Number of stages |
| *Product* | Amount of desired isotope produced in 1 year |
| $P$ | External product flow of each stage |
| $f_{GC}$ | Feed flow rate of gas centrifuge |
| $W$ | External waste flow of each stage |

*Greek characters*

| | |
|---|---|
| $\theta$ | Interstage Cut |
| $\theta 1$ | The cut of the first stage |
| $\alpha$ | Unit separation factor |
| $\varepsilon$ | Recycle flow |

*Subscripts and superscripts*

| | |
|---|---|
| $i$ | Isotope number |
| $n$ | Stage number |
| cas | Cascade |

(see Fig. 1). The optimization of the parameters is carried out by a *meta*-heuristic algorithm. The optimization process in order to enrich the product stream from a certain feed material has been done by the GWO algorithm based on ANN. ANNs are computational models which copy the human brain's learning and decision-making abilities. They are created to undergo a learning process before they are used, (Sahin and Koyuncu, 2012). Several fundamental aspects effect on their decision-making ability and correctness of the decisions such as the network structure, type of activation functions in the neurons, learning period, and number of neurons used during the learning process. The ANN is a well-established tool for different forecasting problems in different areas, with over 40 years of application (Naserbegi et al., 2020b). Because of their universal approximate functional form, ANN seems to be a good choice for modeling non-linear dependency (Mahmoudi and Aghaie, 2019).

In optimization problems, the correct definition of objective function is significant (Aghaie and Mahmoudi, 2016). The objective functions have so far been studied, are separation capacity, total feed flows (Sulaberidze et al., 2018), total number of GCs (Palkin, 2014), D parameter along with total interstage flow rates (Mansourzadeh et al., 2018), amount of product (Imani et al., 2020), and the desired isotope concentration along with total interstage flows (Palkin, 2013). In this study, using GWO-ANN, two test cases are evaluated. In first test case, attempts have been made to maximize the amount of product for a given amount of natural Tellurium in a fixed configuration. In second case, the total number of the GCs are also minimized. The ANNs with lower time consumption and proper accuracy are used for objective functions evaluations. Finally, the optimized parameters are reported and results are discussed.

## 2. Simulation of square cascades

In a square cascade, GCs are connected in parallel and in series. The number of GCs is equal in all stages; i.e., the internal feed flow rates in all stages are same. Another feature of square cascades is the use of recycle flows in the first and last stages of the cascade (see Fig. 2). Square cascade analysis can, generally, be divided into two parts; (1) calculation of interstage flow rates, and (2) calculation of concentration distribution. In the first step, using the equations obtained from the total mass balance at each stage and mixing points, all flow rates between the steps will be calculated. In the second step, by solving the mass conservation equations for each component in stages and mixing points, the concentration distribution along the cascade will be calculated.

### 2.1. Calculation of interstage flows

The external hydraulic parameters of cascade are introduced as cascade feed flow $(F_n)$, waste flow $(W_n)$ and product flow $(P_n)$. The internal hydraulic parameters of cascade are product flow of stages $(L_n^{'})$, waste flow of stages $(L_n^{''})$ and interstage cut $(\theta_n)$ (Zeng and Ying, 2002).

Another internal hydraulic parameter is feed to the stages $(L_n)$. By placing the following equations in a linear equation system, the

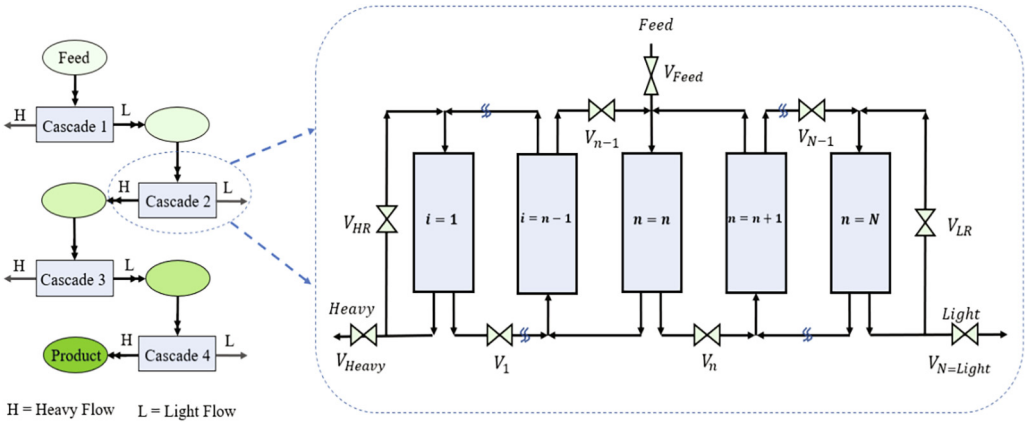


**Fig. 1.** Schematic diagram of a separation path and a square cascade.

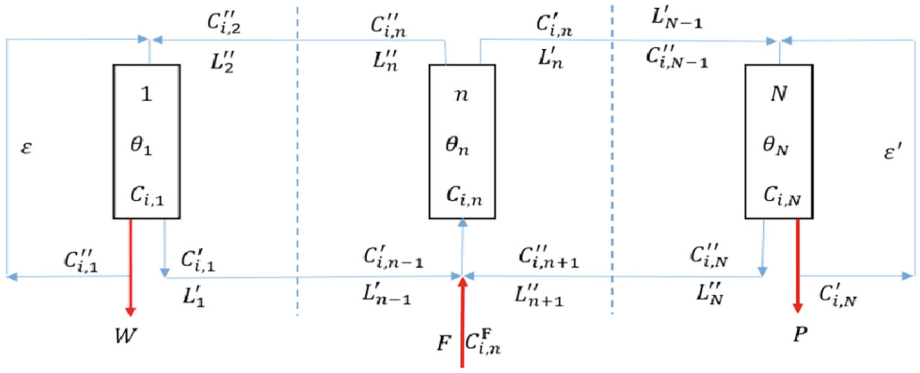


**Fig. 2.** Schematic of a square cascade.

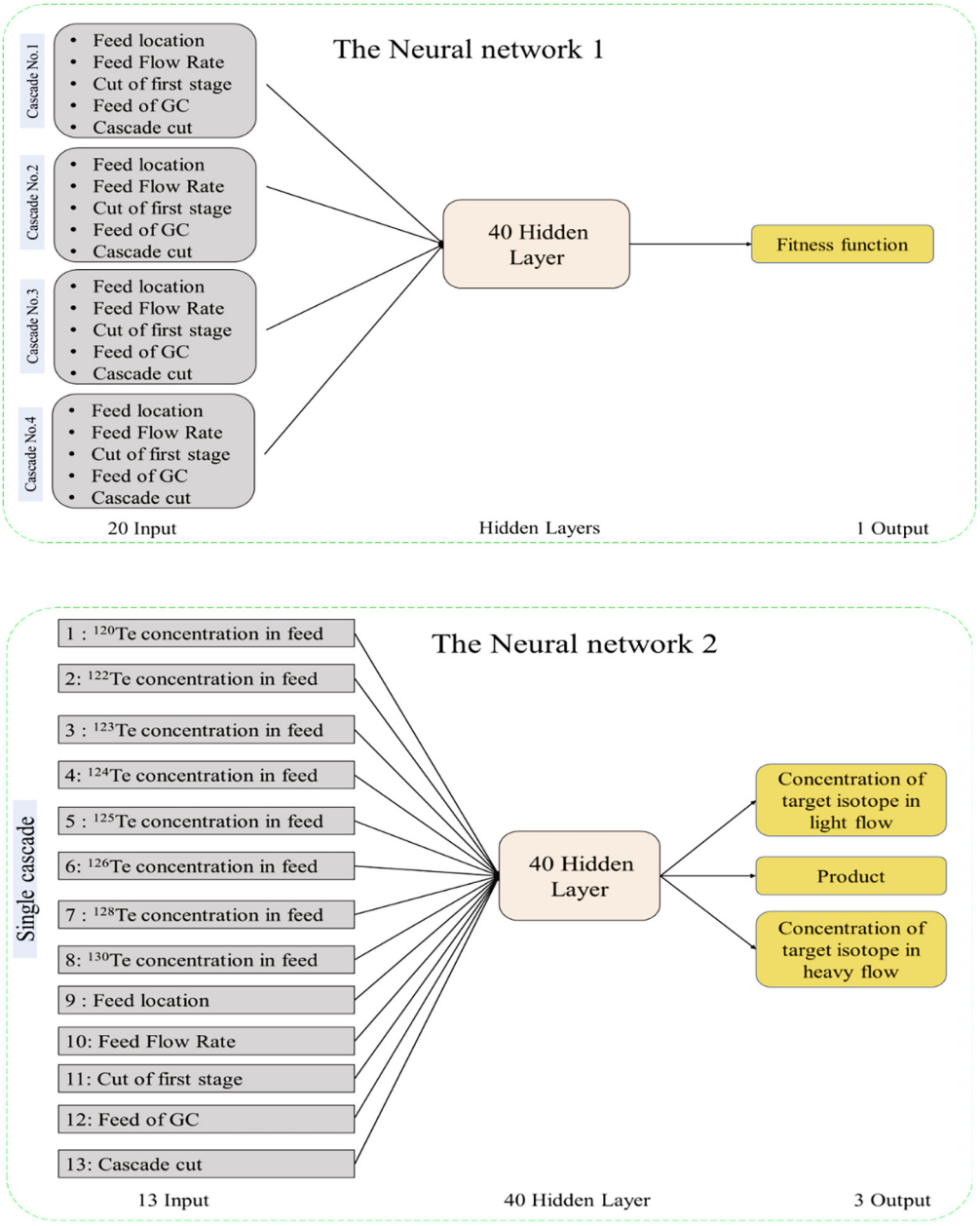


**Fig. 3.** Schematic diagram of inputs and outputs for two ANNs used in GWO algorithm.

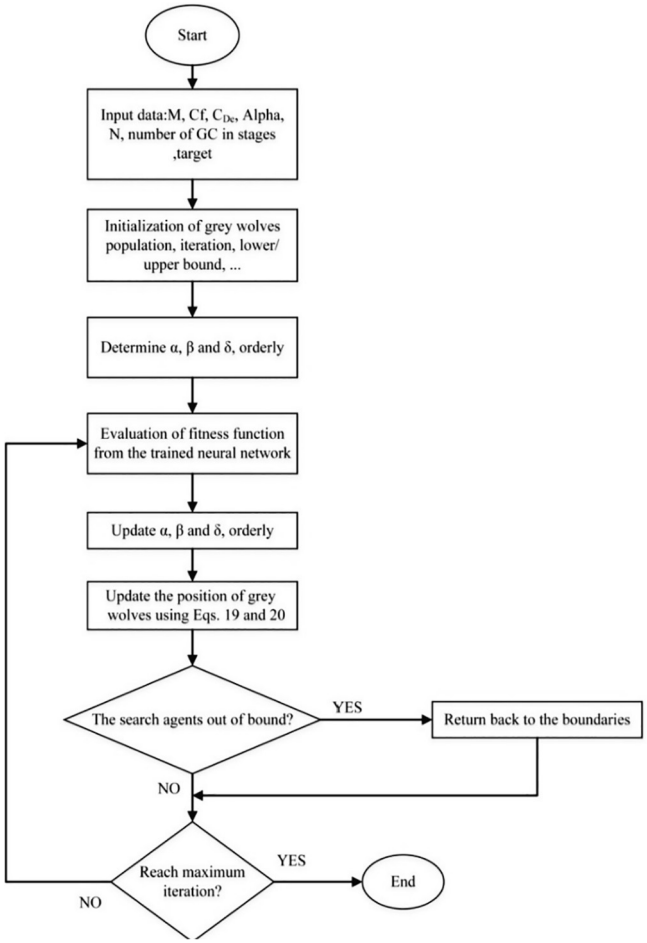


**Fig. 4.** GWO-ANN optimization algorithm.

**Table 1**
Natural concentration of Tellurium isotopes.

| $^{130}$Te | $^{128}$Te | $^{126}$Te | $^{125}$Te | $^{124}$Te | $^{123}$Te | $^{122}$Te | $^{120}$Te |
|---|---|---|---|---|---|---|---|
| 0.3416 | 0.3170 | 0.1880 | 0.0707 | 0.0474 | 0.0089 | 0.0255 | 0.0009 |

hydraulic parameters are calculated $(L_n^{''}, L_n^{'}, \theta_n)$ (Zeng et al., 2003; Zeng and Ying, 2000b).

$$L_n - L^{'}{}_n - L^{''}{}_n - P_n - W_n = 0 \tag{1}$$

$$L_n = L'_{n-1} - L^{''}{}_{n+1} + F_n \tag{2}$$

$$\theta_n = \frac{L^{'}{}_n}{L_n} \tag{3}$$

$$\theta_{cas} = \frac{P}{F} \tag{4}$$

$$W + \varepsilon = L^{''}{}_1 \tag{5}$$

$$P + \varepsilon^{'} = L^{'}{}_N \tag{6}$$

### 2.2. Simulation of concentration distribution in the cascade

The mass conservation equation for isotope $i$ in step $n$ is written as follows (Zeng and Ying, 2001):

$$L_n C_{i,n} - \left(L_n^{''} + W_n\right) C_{i,n}^{''} - \left(L_n^{'} + P_n\right) C_{i,n}^{'} = 0 \tag{7}$$

Similarly, for mixing points, the mass conservation equation for isotope $i$ is obtained as Eq. (8).

$$\left(L_{n+1}^{''}\right) C_{i,n+1}^{''} + \left(L_{n-1}^{'}\right) C_{i,n-1}^{'} + F_n C_{i,n}^{F} = L_n C_{i,n} \tag{8}$$

In Eqs. (7) and (8), $C_{i,n}^{'}$is the $i$ isotope concentration in the product stream, and $C_{i,n}^{''}$represents the $i$ isotope concentration in the waste stream. Another relation used is the step separation factor is as Eq. (9) (Zeng and Ying, 2001).

$$\frac{C_{i,n}^{'}/C_{i,n}^{''}}{C_{j,n}^{'}/C_{j,n}^{''}} = \alpha_0^{M_j - M_i} \tag{9}$$

In addition to the relationships mentioned above, the following conditions should be calculated in the feed and withdrawal flows of each stage.

$$\sum_i C_{i,n} = 1 \sum_i C_{i,n}^{'} = 1 \sum_i C_{i,n}^{''} = 1 \tag{10}$$

It is clear, the number of unknown flows in the cascade is fewer than the equations. It is necessary to specify two parameters, so the number of the unknowns and the equations will be equal. In this paper, these two parameters are the cascade cut and the cut of the first stage. Using Equations (7) and (8), the following equation is obtained:

$$-L^{'}n - 1C^{'}i, n - 1 + \left(L_n^{'} + P_n\right) C_{i,n}^{'} + \left(L_n^{''} + W_n\right) C_{i,n}^{''} - L^{''}n + 1C^{''}i, n + 1 = F_c C_{i,n}^{F}. \tag{11}$$

In Eq. (11), the known values are of the right side of the equation, and the values of the unknown concentrations are of the left side. The coefficients can be named and the relation can be written in the following simplified form.

$$-\varphi_{i,n-1} C_{i,n-1}^{'} + \omega_{i,n} C_{i,n}^{'} + \delta_{i,n} C_{i,n}^{''} - \gamma_{i,n+1} C_{i,n+1}^{''} = r_{i,n} \tag{12}$$

To solve the above equation in each time interval, the $q$ iteration method is used (Zeng and Ying, 2000c). In this method, a parameter called $q$ is defined, which is equal to the ratio of one isotope concentration in the enriched stream to the concentration of the same isotope in the prepared stream.

$$q = C_i^{'}/C_i^{''} \tag{13}$$

By applying this definition to the separation factor relation, the following relation will be obtained.

$$\frac{C_{i,n}^{'}/C_{i,n}^{''}}{C_{j,n}^{'}/C_{j,n}^{''}} = \frac{q_{i,n}}{q_{j,n}} = \alpha_0^{M_j - M_i} \tag{14}$$

$$q_{i,n} = q_{j,n} \alpha_0^{M_j - M_i} \tag{15}$$

Using Eq. (15), the value of concentration $C_{i,n}^{'}$can be determined as follows:

$$C_{i,n}^{'} = C_{i,n}^{''} q_{i,n} = C_{i,n}^{''} q_{j,n} \alpha_0^{M_j - M_i} \tag{16}$$

By placing the above relation to the left side of Eq. (12), the following relation is obtained (Zeng and Ying, 2001):

$$-\varphi_{i,n-1} q_{i,n-1} C_{i,n-1}^{''} + \omega_{i,n} q_{i,n} C_{i,n}^{''} + \delta_{i,n} C_{i,n}^{''} - \gamma_{i,n+1} C_{i,n+1}^{''} = r_{i,n} \tag{17}$$

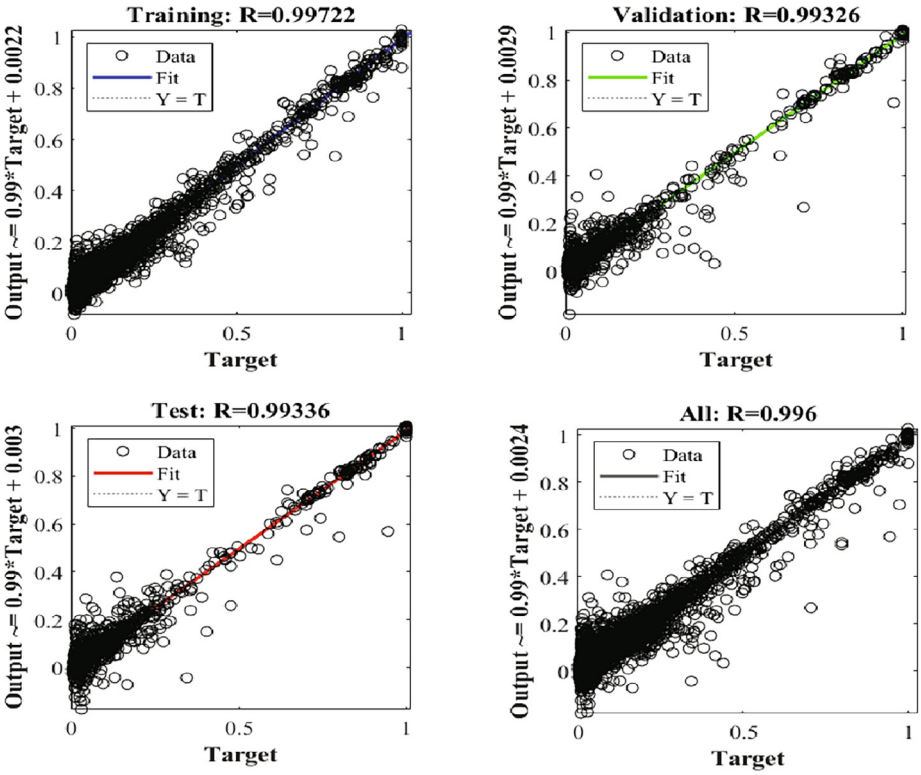


**Fig. 5.** The neural network training regression for the neural network 1.

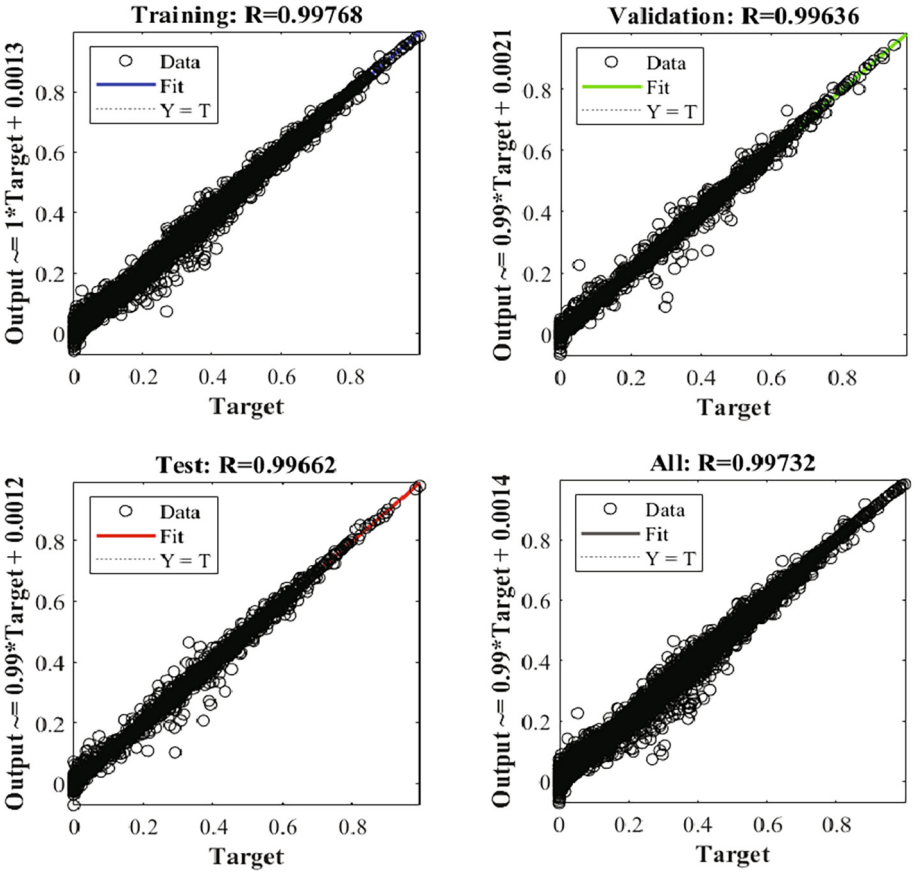


**Fig. 6.** The neural network training regression for the neural network 2.

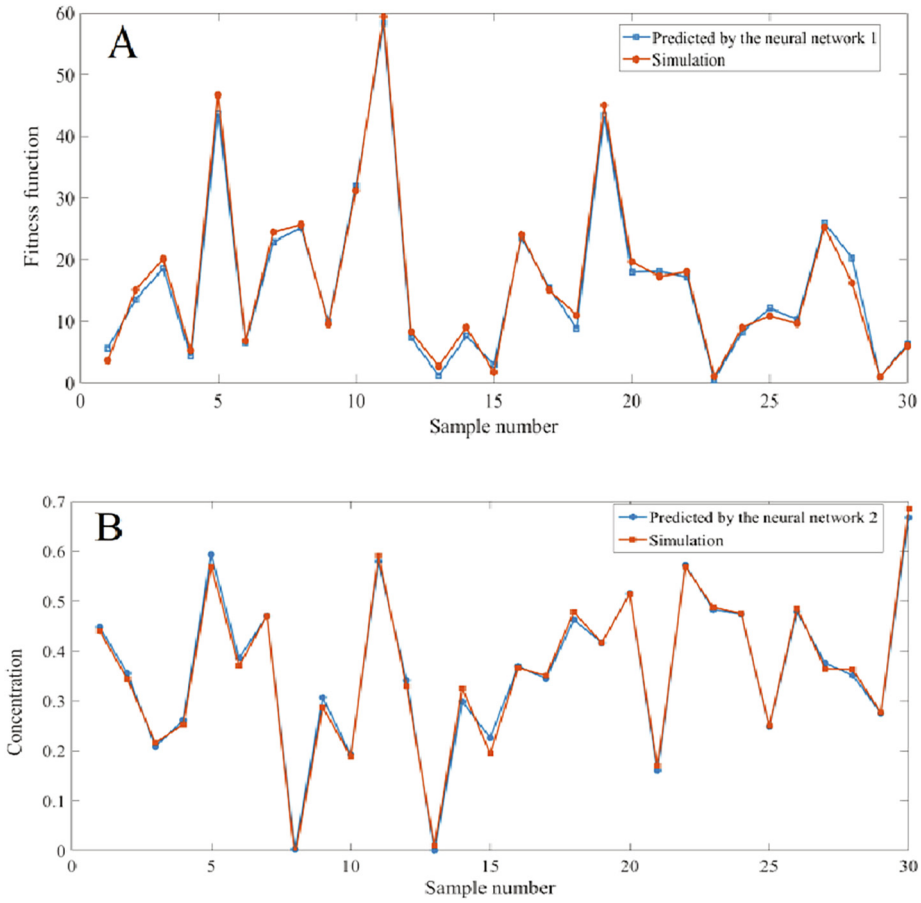


**Fig. 7.** The real and predicted values for 30 random samples A) ANN1 B) ANN2.

**Table 2**
The error values for both networks.

| | *MAE* | *RMSE* | *MAPE* |
|---|---|---|---|
| ANN 1 | 0.57 | 1.03 | 2.7% |
| ANN 2 | 0.26 | 0.33 | 1.7% |

**Table 3**
The average calculation time for a single run, 10 repeated tests (four cascades).

| Used network | NN 1 | NN 2 |
|---|---|---|
| Time (s) | 0.00124 | 0.00928 |

The above equation must be solved at each step for each isotope to calculate the concentration of all components.

### 2.3. Determining the connection flow between cascades

In separation of middle isotopes, it is not possible to reach high concentrations in one separation step. In addition, in the presence of a long separation cascade, the concentration of isotopes in light or heavy sides of the cascade is limited to specific values. These values are calculated according to the following equations (Zeng and Ying, 2000a).

$$C_{P,k}^{max} = \frac{C_{F,k}}{\sum_{i=1}^{k} C_{F,i}} \tag{18}$$

$$C_{W,k}^{max} = \frac{C_{F,k}}{\sum_{i=k}^{N} C_{F,i}} \tag{19}$$

Using these relationships, the appropriate withdrawal of cascade for separation of the target isotope can be determined. Collecting the product of the first step (cascade) as the feed of the next step, the same procedure can be done for the second step and the separation map can be determined until the final concentration is reached. In this paper, four cascades are connected as Fig. 1 shows. The equations (18) and (19) can be used for approving the selected separation path.

## 3. Optimization method

### 3.1. Objective functions

In square cascade design, there are five types of parameters that can be optimized, including feed location, feed flow rate, cut of the first stage, cascade cut, and feed flow of GCs. Square cascade can be used in low feed flowrate and the wide range of cut; so, the performance of the square cascade is high. Also, it has high flexibility in separating isotopes. In this research, four square cascades are connected for $^{123}$Te separation. So, there are many parameters that need to be optimize simultaneously. In this article, two test cases are evaluated. In the first case, the cascade arrangement is fixed and in the second one, the number of stages and GCs added to the optimization parameters. In both cases attempts have been made to maximize the amount of product ($^{123}$Te 99.9%) for a given amount of natural Tellurium. The objective functions specified for these two cases are as follows:

$$OF_1 = \min\left(k_1 \frac{1}{Product} + k_2 \left|C_{i,calc} - C_{i,De}\right|\right) \tag{20a}$$

**Table 4**
The GWO-ANN results for $^{123}$Te separation (first test case).

| parameter | Run number | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $NF_1$ | 39 | 37 | 37 | 39 | 39 | 37 | 39 | 37 |
| $Feed_1$ | 0.51 | 0.54 | 0.74 | 0.53 | 0.52 | 0.52 | 0.53 | 0.61 |
| $\theta 1_1$ | 0.36 | 0.40 | 0.41 | 0.37 | 0.37 | 0.36 | 0.35 | 0.38 |
| $f_{GC1}$ | 2.01 | 1.76 | 2.42 | 1.00 | 1.00 | 1.67 | 1.00 | 1.32 |
| $\theta_{cas1}$ | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
| $NF_2$ | 38 | 37 | 27 | 29 | 29 | 33 | 29 | 32 |
| $Feed_2$ | 0.79 | 0.30 | 0.31 | 0.36 | 0.43 | 0.50 | 1.06 | 0.66 |
| $\theta 1_2$ | 0.39 | 0.52 | 0.54 | 0.55 | 0.53 | 0.50 | 0.55 | 0.54 |
| $f_{GC2}$ | 1.19 | 1.02 | 1.85 | 1.31 | 2.26 | 1.00 | 1.31 | 1.91 |
| $\theta_{cas2}$ | 0.60 | 0.59 | 0.57 | 0.54 | 0.55 | 0.63 | 0.61 | 0.59 |
| $NF_3$ | 2 | 5 | 2 | 2 | 2 | 3 | 2 | 3 |
| $Feed_3$ | 0.30 | 0.40 | 0.30 | 0.37 | 0.36 | 0.32 | 0.37 | 0.35 |
| $\theta 1_3$ | 0.41 | 0.55 | 0.49 | 0.52 | 0.52 | 0.50 | 0.52 | 0.52 |
| $f_{GC3}$ | 1.00 | 1.00 | 1.00 | 1.00 | 1.18 | 1.02 | 1.00 | 1.00 |
| $\theta_{cas3}$ | 0.18 | 0.14 | 0.15 | 0.11 | 0.14 | 0.12 | 0.14 | 0.14 |
| $NF_4$ | 35 | 36 | 35 | 39 | 37 | 38 | 39 | 37 |
| $Feed_4$ | 0.30 | 0.32 | 0.30 | 0.30 | 0.36 | 0.34 | 0.30 | 0.33 |
| $\theta 1_4$ | 0.54 | 0.53 | 0.51 | 0.49 | 0.48 | 0.55 | 0.49 | 0.46 |
| $f_{GC4}$ | 1.24 | 1.71 | 1.11 | 1.00 | 1.37 | 1.97 | 1.00 | 1.03 |
| $\theta_{cas4}$ | 0.11 | 0.10 | 0.11 | 0.12 | 0.10 | 0.08 | 0.12 | 0.13 |
| Concentration | 0.9990 | 0.9990 | 0.999 | 0.9991 | 0.9992 | 0.9992 | 0.9991 | 0.9990 |
| Product (gr/year) | 46.9 | 42.8 | 42.2 | 41.31 | 44.0 | 42.00 | 41.31 | 39.9 |

$$OF_2 = \min\left(k_1\frac{1}{Product} + k_2\left|C_{i,calc} - C_{i,De}\right| + \mathrm{N} * \mathrm{NumberofGCsinastage}\right) \tag{20b}$$

In the two objective functions, the value of the final concentration $C_i$ is applied as a constraint. The *De* index represents the desired value of the parameter (concentration) and the *calc* index represents the values calculated from the simulation. Also, the $k_1$ is equal to 100, and $k_2$ is equal to 10000. These are weighing factors. In the first objective function, the aim is to gain the maximum product from a specified amount of feed by a specific arrangement of GCs. In the second one, the aim is to gain the maximum product from a specified amount of feed by minimizing the total number of GCs. In these cases, the base septation factors are different and their effects are studied.

### 3.2. Gray Wolf optimization algorithm

In this work, to optimize the square cascades, the Gray Wolf *meta*-heuristic algorithm is used. The Gray Wolf algorithm was proposed by Mirjalili et al., in 2014 based on their group hunting (Mirjalili et al., 2014). The method of hunting gray wolves is mathematically modeled and the best solution is called *alpha*, the second optimal solution is *beta,* and the third optimal solution is called *delta*; the rest of the answers that do not matter are called *omega* (Naserbegi et al., 2020b). The hunting process is monitored by the alpha, beta and delta wolves. Omega wolves are looking for alpha, beta, and delta wolves. Gray wolves basically follow three stages of hunting. The first stage is tracking and chasing the prey, the second stage is immobilizing the prey by encircling it, and the third stage to attack the prey. The second stage is introduced by Equations (21) and (22) (Nadimi-Shahraki et al., 2021):

$$\vec{\mathrm{D}} = \left|\vec{C}.\vec{X_p}(t) - \vec{X_p}(t)\right| \tag{21}$$

$$\vec{\mathrm{X}}(\mathrm{t}+1) = \vec{X_p}(t) - \vec{A}.\vec{D} \tag{22}$$

Where $\vec{A}$ and $\vec{C}$ denote the coefficient vectors, $\vec{X_p}$ and $\vec{X}$ are the prey and the gray wolf's position, respectively, and the $\vec{D}$ determines the new position of the wolves. In addition, the parameter t represents the number of iterative loops, and $\vec{\mathrm{X}}(\mathrm{t}+1)$ indicates the position of the wolves in the next iteration. $\vec{A}$ and $\vec{C}$ vectors are defined in the following equations (Mirjalili et al., 2014):

$$\vec{\mathrm{A}} = 2\vec{\mathrm{a}}.\vec{r_1} - \vec{\mathrm{a}} \tag{23}$$

$$\vec{C} = 2\vec{r_2} \tag{24}$$

Where, the vector $\vec{\mathrm{a}}$ changes linearly from zero to 2 along the iteration loop. In the first iteration, it is equal to 2, and in the last iteration, it is equal to zero. The vectors $\vec{r_1}$ and $\vec{r_2}$ are random vectors in the range of zero to 1. In modeling the hunting behavior of gray wolves, alpha, beta and delta wolves are used assuming that they have more information about the hunting position; therefore, the position of omega wolves based on these three wolves (optimal answers) is updated according to the following relationships (Bian et al., 2017):

$$\vec{\mathrm{X}}(\mathrm{t}+1) = \frac{\vec{X_1} + \vec{X_2} + \vec{X_3}}{3} \tag{25}$$

$$\begin{aligned}\vec{X_1} &= \left|\vec{X_\alpha} - \vec{A_1}\vec{D_\alpha}\right|, \vec{X_2} = \left|\vec{X_\beta} - \vec{A_2}\vec{D_\beta}\right|, \vec{X_3} \\ &= \left|\vec{X_\delta} - \vec{A_3}\vec{D_\delta}\right|\end{aligned} \tag{26}$$

$$\begin{aligned}\vec{D_\alpha} &= \left|\vec{C_1}.\vec{X_\alpha} - \vec{X}\right|, \vec{D_\beta} = \left|\vec{C_2}.\vec{X_\beta} - \vec{X}\right|, \vec{D_\delta} \\ &= \left|\vec{C_3}.\vec{X_\delta} - \vec{X}\right|\end{aligned} \tag{27}$$

The vectors $\vec{X_1}$, $\vec{X_2}$ and $\vec{X_3}$ are the three optimal answers of the algorithm in the iteration of t. The vector $\vec{A}$ has a random value between −2a and 2a. When $\left|\vec{A}\right| < 1$,the wolves attack and when $\left|\vec{A}\right| > 1$, the wolves have to change their position to find the prey (see for more details [17]).

To find the optimal cascade, the GWO algorithm first generates a random population of wolves (answers). For separation in four steps, each cascade will have five parameters for optimization, and a total of 20 random parameters will be generated. The values

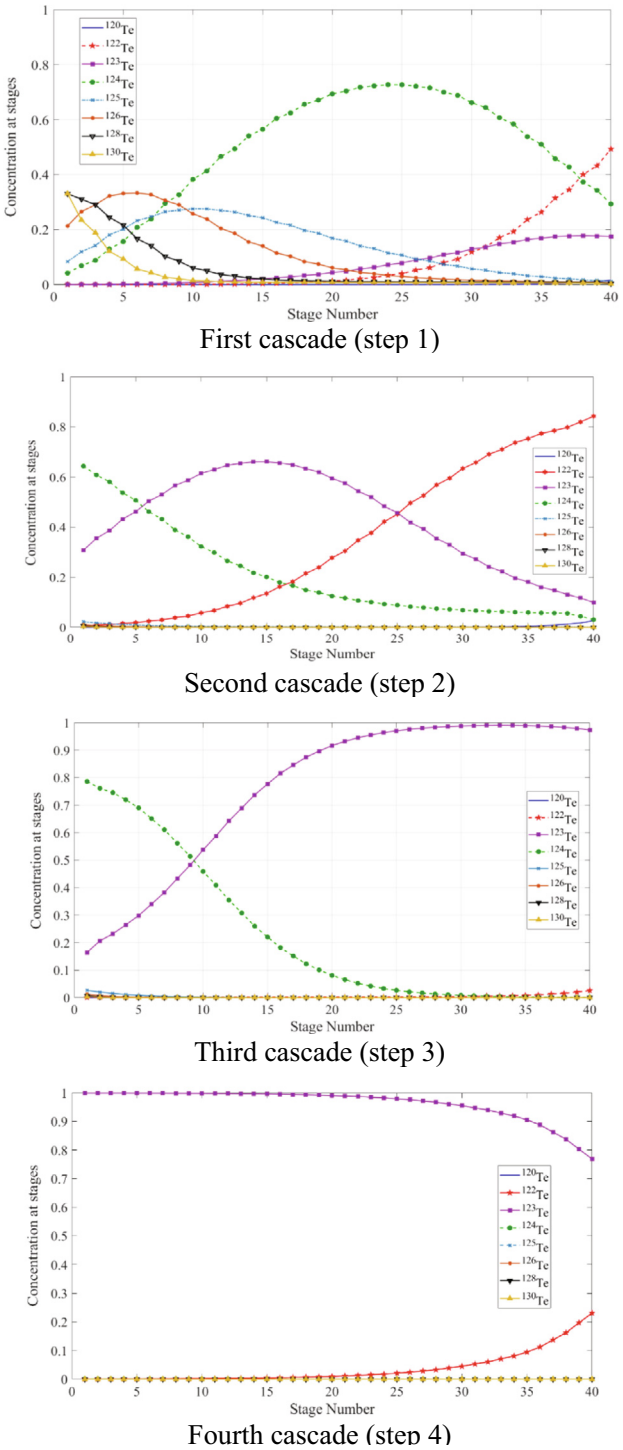


**Fig. 8.** Concentration distribution of stages during separation steps for the test case 1.

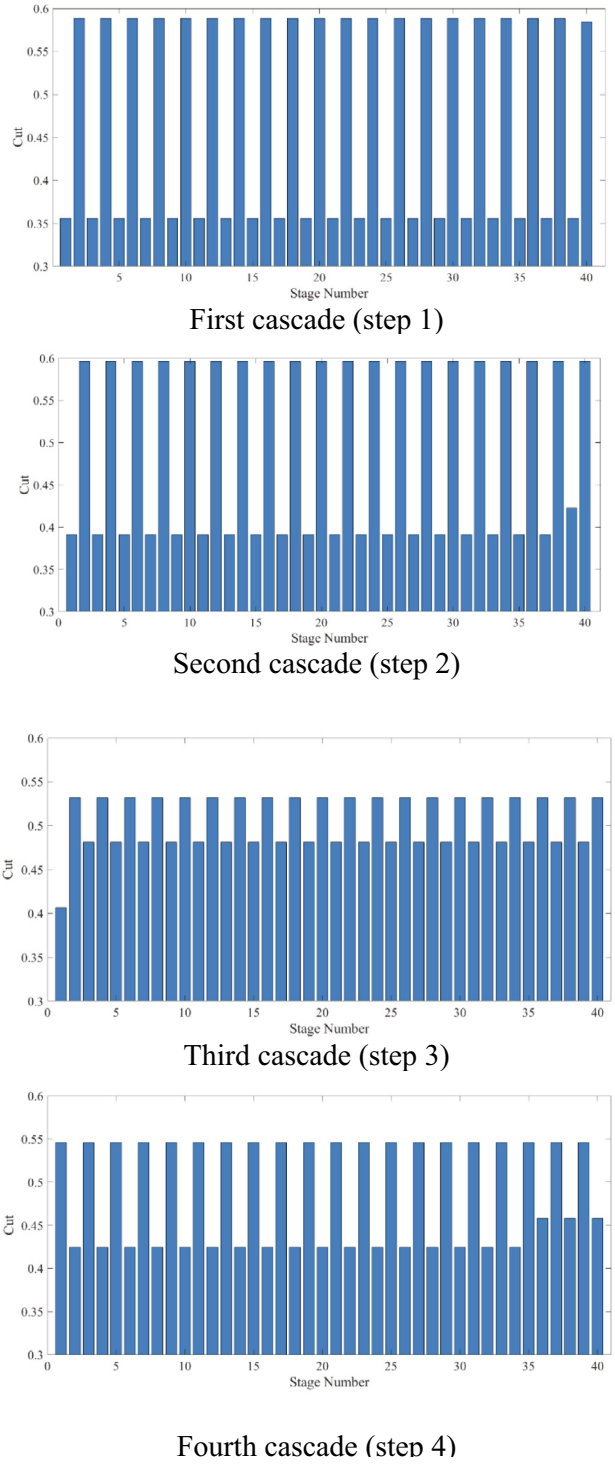


**Fig. 9.** Interstage cuts for the test case.

of these defined parameters are between the defined maximum and minimum values. The four-step separation, for instance, has random parameters as follows:

$$T = \left\{ \begin{matrix} NF_1 & Feed_1 & \theta 1_1 & f_{GC1} & \theta_{cas1} & NF_2 & Feed_2 & \theta 1_2 & f_{GC2} & \theta_{cas2} \\ NF_3 & Feed_3 & \theta 1_3 & f_{GC3} & \theta_{cas3} & NF_4 & Feed_4 & \theta 1_4 & f_{GC4} & \theta_{cas4} \end{matrix} \right\} \quad (28)$$

In this vector, all parameters belong to one cascade in different separation steps, *NF* is the feed location, *Feed* denotes the input feed flow rate, $\theta_1$ represents the cut of the first stage, $f_{GC}$ is the feed flow rate of a centrifuge and $\theta_{cas}$ stands for the cascade cut. Fig. 3 illustrates the GWO algorithm for finding the optimal answer. The value of objective function is calculated by a trained neural network. For second test case in this research, N (number of stages) and number of GCs in a stage is also added to above vector.

**Table 5**
The concentration of isotopes in light and heavy sides for test case 1.

| Isotope | Light side | | | | Heavy side | | | |
|---|---|---|---|---|---|---|---|---|
| | Step 1 | Step 2 | Step 3 | Step 4 | Step 1 | Step 2 | Step 3 | Step 4 |
| $^{120}$Te | 0.0180 | 0.0301 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| $^{122}$Te | 0.5098 | 0.8494 | 0.0288 | 0.2523 | 0.0000 | 0.0057 | 0.0008 | 0.0003 |
| $^{123}$Te | 0.1733 | 0.0933 | 0.9706 | 0.7477 | 0.0002 | 0.2920 | 0.1478 | 0.9990 |
| $^{124}$Te | 0.2793 | 0.0266 | 0.0006 | 0.0000 | 0.0352 | 0.6545 | 0.7934 | 0.0007 |
| $^{125}$Te | 0.0101 | 0.0004 | 0.0000 | 0.0000 | 0.0739 | 0.0244 | 0.0296 | 0.0000 |
| $^{126}$Te | 0.0049 | 0.0001 | 0.0000 | 0.0000 | 0.1976 | 0.0120 | 0.0145 | 0.0000 |
| $^{128}$Te | 0.0030 | 0.0000 | 0.0000 | 0.0000 | 0.3335 | 0.0076 | 0.0092 | 0.0000 |
| $^{130}$Te | 0.0016 | 0.0000 | 0.0000 | 0.0000 | 0.3595 | 0.0039 | 0.0048 | 0.0000 |

**Table 6**
Maximum concentration can be obtained from light and heavy sides for test case 1.

| No. | $C_{P,De}^{max}$ (Light side) | $C_{W,De}^{max}$ (Heavy side) | Product side |
|---|---|---|---|
| Step 1 | 0.2521 | 0.0091 | Light |
| Step 2 | 0.2471 | 0.3669 | Heavy |
| Step 3 | 0.9808 | 0.2936 | Light |
| Step 4 | 0.9711 | 0.9993 | Heavy |

### *3.3. Prediction of objective function by an artificial neural network (ANN)*

ANNs are approaches of machine learning operation that teach the computers to perform several tasks by taking into account the training data and examples. ANNs generate the recognizing features from the learning data which they process, automatically. In fitting problems, neural network plans between a data set of numeric inputs and a set of numeric targets. The networks can be trained with Levenberg-Marquardt back propagation, Bayesian Regularization, and Scaled Conjugate Gradient algorithms. An ANN is a strong Toolbox in MATLAB which has been used in this research (Naserbegi et al., 2020a). In order to develop the predictions to use as an objective function in GWO, two feed-forward multi-layer neural network with a 40 hidden layer has been used. The first network considers four cascades simultaneously with 20 inputs (see Fig. 3). The second network applies on single cascade with 13 inputs (see Fig. 3). The square cascade was simulated to apply the input data for the neural network. Moreover, the concentration and amount of the product have calculated for 5000 randomly produced data for square cascades. All data have been divided into three categories by network including a) training (3500 of the data), b) validation to estimate the extension of the network (750 of the data) and c) test data to estimate the performance of the network (750 of the data), randomly. The Levenberg-Marquardt algorithm has been chosen for the both networks due to taking less time. Training automatically stops when generalization stops improving, as represented by a rise in the mean square error (*MSE*) of the validation data. Fig. 3 shows the schematic diagram of the networks. The first network has an out-

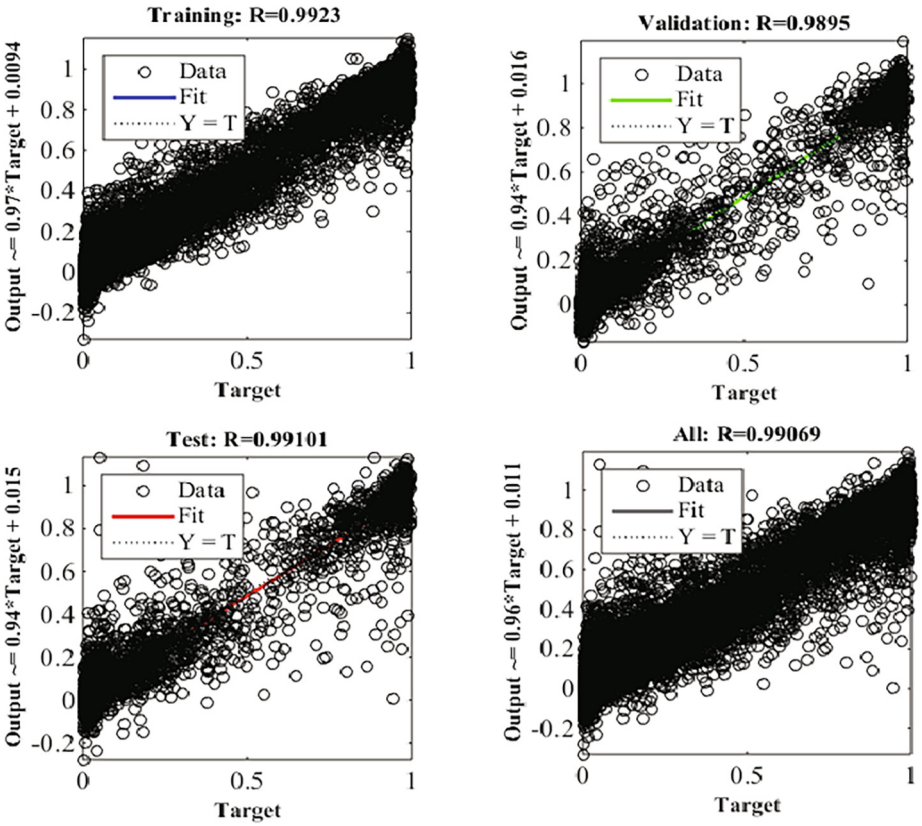


**Fig. 10.** The neural network training regression for the neural network in test case 2.

**Table 7**
The GWO-ANN results for the test case 2.

| parameter | Run number | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $NF_1$ | 27 | 32 | 25 | 26 | 31 | 30 | 28 | 27 |
| $Feed_1$ | 0.496 | 0.51 | 0.487 | 0.492 | 0.449 | 0.506 | 0.501 | 0.482 |
| $\theta 1_1$ | 0.549 | 0.525 | 0.549 | 0.544 | 0.537 | 0.495 | 0.522 | 0.487 |
| $f_{GC1}$ | 18.05 | 17.29 | 18.47 | 20.39 | 19.35 | 18.45 | 19.23 | 20.66 |
| $\theta_{cas1}$ | 0.052 | 0.051 | 0.05 | 0.051 | 0.051 | 0.052 | 0.05 | 0.05 |
| $NF_2$ | 32 | 33 | 28 | 30 | 33 | 33 | 31 | 30 |
| $Feed_2$ | 1.47 | 1.5 | 1.47 | 1.52 | 1.51 | 1.48 | 1.47 | 1.49 |
| $\theta 1_2$ | 0.501 | 0.551 | 0.5 | 0.542 | 0.526 | 0.533 | 0.519 | 0.543 |
| $f_{GC2}$ | 11.12 | 10.51 | 10.84 | 10.91 | 11.38 | 12.05 | 11.61 | 11.45 |
| $\theta_{cas2}$ | 0.529 | 0.538 | 0.516 | 0.545 | 0.551 | 0.576 | 0.593 | 0.599 |
| $NF_3$ | 4 | 5 | 4 | 4 | 3 | 4 | 5 | 3 |
| $Feed_3$ | 0.69 | 0.71 | 0.71 | 0.68 | 0.69 | 0.77 | 0.74 | 0.69 |
| $\theta 1_3$ | 0.492 | 0.444 | 0.461 | 0.502 | 0.512 | 0.469 | 0.505 | 0.477 |
| $f_{GC3}$ | 9.97 | 10.97 | 10.2 | 10.11 | 9.95 | 11.27 | 11.45 | 10.38 |
| $\theta_{cas3}$ | 0.19 | 0.195 | 0.187 | 0.192 | 0.192 | 0.199 | 0.215 | 0.22 |
| $NF_4$ | 31 | 34 | 33 | 30 | 34 | 32 | 31 | 31 |
| $Feed_4$ | 0.3 | 0.301 | 0.309 | 0.301 | 0.311 | 0.319 | 0.306 | 0.314 |
| $\theta 1_4$ | 0.428 | 0.428 | 0.428 | 0.428 | 0.428 | 0.428 | 0.428 | 0.428 |
| $f_{GC4}$ | 13.98 | 12.2 | 13.29 | 11.64 | 14.48 | 12.35 | 13.31 | 13.18 |
| $\theta_{cas4}$ | 0.088 | 0.083 | 0.091 | 0.098 | 0.09 | 0.09 | 0.089 | 0.102 |
| $N$ | 65 | 65 | 65 | 65 | 65 | 65 | 65 | 65 |
| $NGC$ | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| Concentration | 0.9990 | 0.9990 | 0.9990 | 0.9990 | 0.9990 | 0.9991 | 0.9990 | 0.9991 |
| Product (gr/year) | 63.66 | 63.20 | 61.70 | 60.28 | 60.01 | 59.89 | 59.79 | 59.42 |

put which is the objective function defined in section 3–1. The second network can calculate the output parameter for a single cascade. By 4-times using this network, the objective function for a 4-step separation plant can be predicted. Fig. 4 shows the major steps of GWO-ANN optimization approach.

## 4. Results and discussion

For first test case, to separate the $^{123}$Te from 15 kg of natural feed in order to reach 99.9% concentration, four square cascades with 40 stages and 4 centrifuges in each stage are used. In this case, the cascades configuration and number of GCs are fixed. In second case, the total number of the CGs in four connected cascades is also minimized. In this case the number of the stages (N) and number of GCs in a stage are included into the optimization parameters. For two cases, the base septation factors are different. It means the separation capacities are different. Table 1 shows the natural feed concentrations of Tellurium.

### 4.1. Test case 1

In this case, the objective function is evaluated according to the Eq. (20a). Two networks (ANN1 and ANN2) are trained and the efficient one is chosen. By several times training of networks and modeling the data, two networks with the low errors have been obtained. Figs. 5 and 6 illustrate the validation data, error regression of networks for training data, and test data for ANN1 and ANN2, respectively. Based on Figs. 5 and 6, the networks are going to operate properly while the calculated values approaching 1.

The proficiency of the proposed networks can be investigated by the criteria of

mean absolute error (*MAE*), root mean square error (*RMSE*), and mean absolute percentage error (*MAPE*).

$$MAE = \frac{1}{n}\sum_{1}^{n}|X_{Model} - X_{Real}| \tag{29}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{1}^{n}(X_{Model} - X_{Real})^2} \tag{30}$$

$$MAPE = \left(\frac{1}{n}\sum_{1}^{n}\left(\frac{|X_{Model} - X_{Real}|}{X_{Real}}\right)\right) * 100 \tag{31}$$

Where $n$ is the number of models, $X_{Real}$ is the value obtained by simulation of square cascades and $X_{Model}$ is the objective functions obtained by the trained network. The real values and the prediction values of the model for ANN1 and ANN2 are shown in Fig. 7. This comparison is carried out for 30 random samples. The value of *MAE*, *RMSE* and *MAPE* are presented in Table 2. The error calculation results confirm that the predicted values are in good agreement with the real values. The *RMSE* values for the two networks indicate that the GWO-ANN can be a proper tool in the optimization process.

In Table 3, the time consumption of the networks for four-step separation is presented. The results are the average of 10-times repeating of the test. It can be seen; the first network has the lower time consumption. In the simulation, the non-linear conservation equations for multi-component system are solved by iteration method. So, the two networks reduce the time consumption significantly. For example, the optimization operation with 1000 iterations and 30 initial search agents by GWO takes approximately 10 h on a personal computer (Intel (R) Core (TM) i7-9750H CPU @ 4.40 GHz). It takes about 130 s by replacing the first network with the objective function.

The first network is used in the GWO algorithm evaluations, regarding its time consumption. All GWO-ANN parameters for the 8 best runs and the top objective values are given in Table 4. In this case, for simplicity the separation factor of the GCs is considered as a function of the GC feed flowrate ($f_{GC}$) (Eq. (32)). The minimum allowable flow rate to a GC is 1 mg/s and the maximum allowable flow rate to a GC is 20 mg/s.

$$\alpha_0 = 1.3 f_{GC}^{-0.05} \tag{32}$$

For the best run (Run No.1), the four-step separation of cascades produce 46.9 gr of $^{123}$Te with 99.9% concentration per year. In this evaluation, the results and the process parameters used for simulation are presented in Figs. 8–9, respectively. Fig. 8 shows the concentration of isotopes along the cascades at each step.

It can be seen, the $^{123}$Te concentration increases in each step and desired concentration is obtained in fourth cascade. Fig. 9

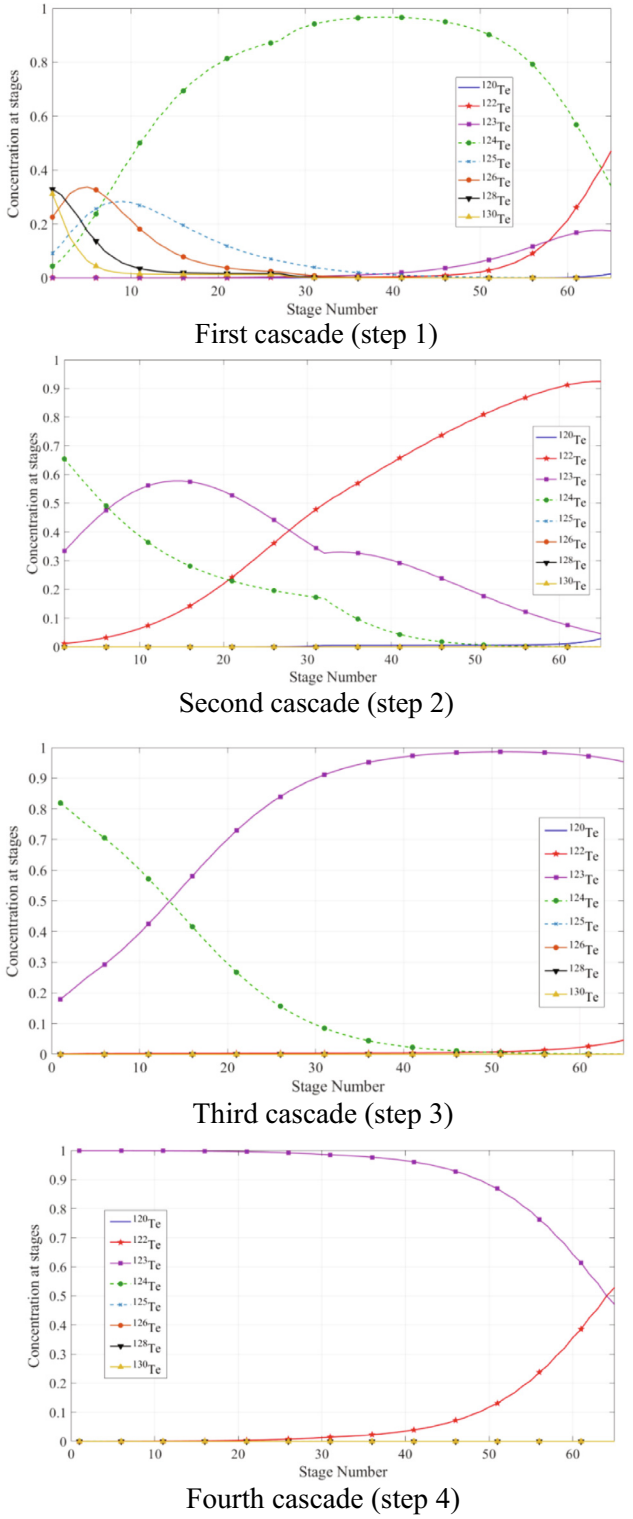


**Fig. 11.** Concentration distribution of stages during separation steps for the test case 2.

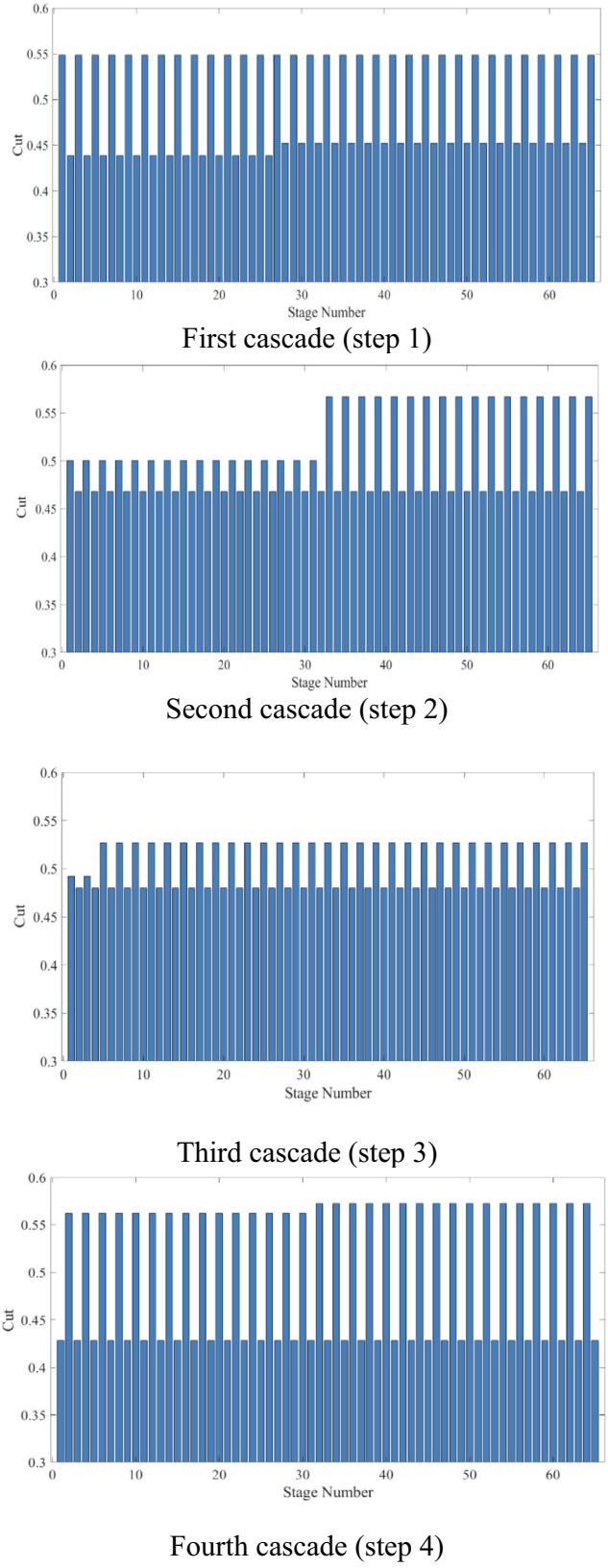


**Fig. 12.** Interstage cuts for the test case 2.

shows the changes in the interstage cuts of the cascades in different steps. The values of interstage cuts vary between 0.35 and 0.6. To apply these cuts to the cascades, the control valves and pressure functions can be used to operate the cascades.

Table 5 shows the concentration of isotopes in light and heavy sides of each step. Using Eqs. (18) and (19) the maximum concentration of the desired isotope in each step calculated and presented in Table 6. It can be seen, the selected path for the separation of $^{123}$Te is correct.

**Table 8**
The optimal parameters for test 2 (limited feed flow rate).

| Step No. | NF | Feed | $\theta 1$ | $f_{GC}$ | $\theta_{cas}$ |
|---|---|---|---|---|---|
| Step 1 | 40 | 0.519 | 0.374 | 1.970 | 0.050 |
| Step 2 | 36 | 0.506 | 0.351 | 2.499 | 0.573 |
| Step 3 | 11 | 0.303 | 0.350 | 1.656 | 0.173 |
| Step 4 | 13 | 0.456 | 0.416 | 2.141 | 0.081 |

**Table 9**
The concentration of isotopes in light and heavy sides for test case 1.

| Isotope | Light side | | | | Heavy side | | | |
|---|---|---|---|---|---|---|---|---|
| | Step 1 | Step 2 | Step 3 | Step 4 | Step 1 | Step 2 | Step 3 | Step 4 |
| $^{120}$Te | 0.0001 | 0.0068 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| $^{122}$Te | 0.0045 | 0.6608 | 0.0050 | 0.0371 | 0.0000 | 0.0101 | 0.0010 | 0.0003 |
| $^{123}$Te | 0.0182 | 0.2871 | 0.9712 | 0.9628 | 0.0000 | 0.3173 | 0.1684 | 0.9990 |
| $^{124}$Te | 0.9660 | 0.0452 | 0.0238 | 0.0000 | 0.0327 | 0.6724 | 0.8303 | 0.0007 |
| $^{125}$Te | 0.0106 | 0.0000 | 0.0000 | 0.0000 | 0.0745 | 0.0002 | 0.0003 | 0.0000 |
| $^{126}$Te | 0.0006 | 0.0000 | 0.0000 | 0.0000 | 0.1982 | 0.0000 | 0.0000 | 0.0000 |
| $^{128}$Te | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3343 | 0.0000 | 0.0000 | 0.0000 |
| $^{130}$Te | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.3602 | 0.0000 | 0.0000 | 0.0000 |

**Table 10**
Maximum concentration can be obtained from light and heavy sides for test case 2.

| No. | $C_{P,De}^{max}$ (Light side) | $C_{W,De}^{max}$ (Heavy side) | Product side |
|---|---|---|---|
| Step 1 | 0.2521 | 0.0091 | Light |
| Step 2 | 0.2520 | 0.3520 | Heavy |
| Step 3 | 0.9690 | 0.3205 | Light |
| Step 4 | 0.9509 | 0.9993 | Heavy |

### 4.2. Test case 2

In this case, the number of stages and the number of GCs in each stage added to the optimization parameters, and the obtained parameters are used for all steps. The unit separation factor of GCs is considered as a function of feed flow rate and cut according to Eq. (33). It is used in Eq. (15) for multi component separation factor calculation in q-iteration method.

$$\alpha_0 = (0.5 - 0.5\theta + 0.246\theta^2)(f_{GC} + 13.79)^{-0.08} \quad (33)$$

Where $\theta$ and $f_{GC}$ are the cut and feed flow rate of a single GC, respectively. Due to using 2 more optimization parameters (n umber of stages and number of GCs), 10,000 randomly generated data have been used for training the network. In this case using previous case experience proper network is trained. The Levenberg-Marquardt algorithm is selected to train the networks with 40 hidden layers. Fig. 10 illustrate the validation data, the error regression of network for training data. Based on Fig. 10, the network has high accuracy (R = 0.99069) for using in the optimization algorithm.

All the parameters for the 8 best runs and the lowest objective values are given in Table 7. In these two cases 20 search agents are considered.

For the best run, the cascade has 65 stages with 2 GCs in each stage. This four-step configuration produces 63.6 gr of $^{123}$Te with 99.9% concentration per year. The results and the parameters used for simulation are presented in Figs. 11–12, respectively. Fig. 11 shows the concentration of isotopes along the cascades for each step and Fig. 12 shows the interstage cuts for each step. This case uses different separation factor (Eq. (33)). It means machines used in this test has different separation capacity. Evaluating Eq. (33) shows, it has higher values in high $f_{GC}$ in comparison of Eq. (32). This drives the optimization algorithm to find higher $f_{GC}$ in calculations. The different selected flowrates for two test cases are obviously reported in Tables 4 and 7. Although the number of machines has decreased in this case, but the length of the cascade has increased and due to having a higher separation factor at high feed flows the production rate has increased. For more justification, in this case the feed flow rate of GCs is limited to 1–3 (mg/s). In this variant, the optimal parameters are reported in Table 8. The number of the stages is 60; the number of GCs in a stage is 4 and 51 gr/year of product (99.9%) is obtained. It is clear that the number of GCs increases and the product rate decreases.

Table 9 shows the concentration of isotopes in light and heavy sides of steps. Using equations (18) and (19) the maximum concentration of the desired isotope in each step are presented in Table 10. As can be seen, the selected path for the separation of $^{123}$Te is correct.

## 5. Conclusion

In this study, optimization of four connected square cascades for $^{123}$Te separation to 99.9% concentration is presented. In this regard, by applying the GWO algorithm the square cascades parameters have been optimized. Two different test cases with different trained neural networks have been also used for predicting the objective functions. The results evaluation show, in both cases the networks have proper accuracy. They reduce computing time up to 98%. In test case 1, the network which has 20 inputs and considers the four cascades simultaneously, reduces the time consumption more than the network which can be used for single cascade; as a result, this network used to objective function evaluations in GWO optimization. The amount of product with the desired concentration (99.9%) for the 40 stages with 4 GCs in each stage is determined. In this case the cascade configuration is fixed and the parameters are optimized. In optimal design, by consuming 15 kg natural Tellurium, 46.9 gr/year $^{123}$Te with 99.9% concentration is obtained. In the second case, the number of stages and GCs in each stage added to the optimization parameter, and a new network with 22 inputs is trained for the GWO optimization algorithm. In this case, machines with different separation capacity are used. The results show that square cascades with 65 stages and 2 GCs in each stage can produce 63.6 gr/year $^{123}$Te with 99.9% concentration. Applying limitation in feed flow of a GC, changes the configuration to 60 stages and 4 GCs in each stage and the product rate is decreased to 51 gr/year.

## CRediT authorship contribution statement

**Morteza Imani:** Software, Validation, Data curation, Writing - original draft. **M. Aghaie:** Conceptualization, Methodology, Writing - review & editing, Supervision, Visualization. **Mohammademad Adelikhah:** Software, Validation, Data curation.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References


Aghaie, M., Mahmoudi, S.M., 2016. A novel multi objective Loading Pattern Optimization by Gravitational Search Algorithm (GSA) for WWER1000 core. Prog. Nucl. Energy 93, 1–11. https://doi.org/10.1016/j.pnucene.2016.07.014.

Bian, X.Q., Zhang, Q., Zhang, L., Chen, J., 2017. A grey wolf optimizer-based support vector machine for the solubility of aromatic compounds in supercritical carbon dioxide. Chem. Eng. Res. Des. 123, 284–294. https://doi.org/10.1016/j.cherd.2017.05.008.

Imani, M., Keshtkar, A.R., Rashidi, A., Sabet, J.K., Noroozi, A., 2020. Investigation on the effect of holdup and cascade shape in NFSW cascades. Prog. Nucl. Energy 119, 103182. https://doi.org/10.1016/j.pnucene.2019.103182.

Mahmoudi, S.M., Aghaie, M., 2019. Evaluation of fuzzy based HS and GSA on reloading cycle length optimization of PWR nuclear power plant. Ann. Nucl. Energy 134, 1–10. https://doi.org/10.1016/j.anucene.2019.05.058.

Mansourzadeh, F., Safdari, J., Khamseh, A.GH., Norouzi, A., Khajenouri, M., 2018. Comparison of optimum tapered cascade and optimal square cascade for separation of xenon isotopes using enhanced TLBO algorithm. Sep. Sci. Technol. 53 (13), 2074–2087. https://doi.org/10.1080/01496395.2018.1443135.

Mirjalili, S., Mirjalili, S.M., Lewis, A., 2014. Grey Wolf Optimizer. Adv. Eng. Softw. 69, 46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007.

Nadimi-Shahraki, M.H., Taghian, S., Mirjalili, S., 2021. An improved grey wolf optimizer for solving engineering problems. Expert Syst. Appl. 166, 113917. https://doi.org/10.1016/j.eswa.2020.113917.

Naserbegi, A., Aghaie, M., Mahmoudi, S.M., 2020a. PWR core pattern optimization using grey wolf algorithm based on artificial neural network. Prog. Nucl. Energy 129, 103505. https://doi.org/10.1016/j.pnucene.2020.103505.

Naserbegi, A., Aghaie, M., Zolfaghari, A., 2020b. Implementation of Grey Wolf Optimization (GWO) algorithm to multi-objective loading pattern optimization of a PWR reactor. Ann. Nucl. Energy 148, 107703. https://doi.org/10.1016/j.anucene.2020.107703.

Palkin, V.A., 2014. Centrifuge number optimization of a cascade with one or several types of steps. At. Energy 116 (3), 206–212. https://doi.org/10.1007/s10512-014-9842-z.

Palkin, V.A., 2013. Optimization of a centrifuge cascade for separating a multicomponent mixture of isotopes. At. Energy 115 (2), 109–115. https://doi.org/10.1007/s10512-013-9757-0.

Sahin, I., Koyuncu, I., 2012. Design and implementation of neural networks neurons with radbas, logsig, and tansig activation functions on FPGA. Elektron. ir Elektrotechnika 51–54. https://doi.org/10.5755/j01.eee.120.4.1452.

Smirnov, A.Y., Sulaberidze, G.A., 2014. Features of mass transfer of intermediate components in square gas centrifuge cascade for separating multicomponent mixtures. Theor. Found. Chem. Eng. 48 (5), 629–636. https://doi.org/10.1134/S0040579514050248.

Sulaberidze, G.A., Mustafin, A.P., Smirnov, A.Y., Borisevich, V.D., Zeng, S., Jiang, D., 2018. Optimization of cascades with variable overall separation factors by various efficiency criteria. J. Phys. Conf. Ser. Inst. Phys. Publ. 1099, 012009. https://doi.org/10.1088/1742-6596/1099/1/012009.

Ying, Z.G., Zeng, S., 2015. Separation of multi-component isotope mixtures by matched abundance ratio net cascade. Yuanzineng Kexue Jishu/Atomic Energy Sci. Technol. 49, 6–12. https://doi.org/10.7538/yzk.2015.49.01.0006.

Zeng, S., Cheng, L.u., Jiang, D., Borisevich, V.D., Sulaberidze, G.A., 2014. A numerical method of cascade analysis and design for multi-component isotope separation. Chem. Eng. Res. Des. 92 (11), 2649–2658. https://doi.org/10.1016/j.cherd.2013.12.016.

Zeng, S., Jiang, D., Borisevich, V., Sulaberidze, G., 2011. Use of the Q-cascade in calculation and optimization of multi-isotope separation. Chem. Eng. Sci. 66, 2997–3002. https://doi.org/10.1016/j.ces.2011.03.056.

Zeng, S., Ying, C., 2002. Separating isotope components of small abundance. Sep. Sci. Technol. 37 (15), 3577–3598. https://doi.org/10.1081/SS-120014445.

Zeng, S., Ying, C., 2001. Transient process in gas centrifuge cascades for separation of multicomponent isotope mixtures. Sep. Sci. Technol. 36 (15), 3439–3457. https://doi.org/10.1081/SS-100107913.

Zeng, Shi, Ying, Chuntong, 2000a. A method of separating a middle component in multicomponent isotope mixtures by gas centrifuge cascades. Sep. Sci. Technol. 35 (14), 2173–2186. https://doi.org/10.1081/SS-100102096.

Zeng, Shi, Ying, Chuntong, 2000b. A second-order time-accurate method for determination of concentration distribution of multicomponent mixtures in separation cascades. Sep. Sci. Technol. 35 (5), 729–741. https://doi.org/10.1081/SS-100100187.

Zeng, Shi, Ying, Chuntong, 2000c. A robust and efficient calculation procedure for determining concentration distribution of multicomponent mixtures. Sep. Sci. Technol. 35 (4), 613–622. https://doi.org/10.1081/SS-100100179.

Zeng, S., Zhou, M., Ying, C., 2003. Theoretical and experimental study of a non-stationary isotope separation process in a gas centrifuge cascade. Sep. Sci. Technol. 38 (11), 2375–2394. https://doi.org/10.1081/SS-120022278.