Heat and Mass Transfer (2024) 60:329–361
https://doi.org/10.1007/s00231-023-03434-z

ORIGINAL ARTICLE

Optimization of MLP neural network for modeling effects of electric fields on bubble growth in pool boiling

Mahyar Ghazvini[1] · Seyyed Mojtaba Varedi-Koulaei[2] · Mohammad Hossein Ahmadi[2] · Myeongsub Kim[1]

Received: 11 May 2023 / Accepted: 9 October 2023 / Published online: 16 November 2023

Abstract

In this paper, a multilayer perceptron (MLP)-type artificial neural network model with a back-propagation training algorithm is utilized to model the bubble growth and bubble dynamics parameters in nucleate boiling with a non-uniform electric field. The influences of the electric field on different parameters that describe bubble's behaviors including bubble waiting time, bubble departure frequency, bubble growth time, and bubble departure diameter are considered. This study models single bubble dynamic behaviors of R113 created on a heater in an inconsistent electric field by utilizing a MLP neural network optimized by four different swarm-based optimization algorithms, namely: Salp Swarm Algorithm (SSA), Grey Wolf Optimizer (GWO), Artificial Bee Colony (ABC) algorithm, and Particle Swarm Optimization (PSO). For evaluating the model effectiveness, the MSE value (Mean-Square Error) of the artificial neural network model with various optimization algorithms is measured and compared. The results suggest that the optimal networks in the two-hidden layer and three-hidden layer models for the bubble departure diameter improve MSE by 33.85% and 35.27%, respectively, when compared with the best response in the one-hidden layer model. Additionally, for bubble growth time, the networks with two hidden layers and three hidden layers have the 44.51% and 45.85% reduction in error, when compared with the network with one hidden layer, respectively. For the departure frequency, the error reduction in the two-layer and three-layer networks is 46.85% and 62.32%, respectively. For bubble waiting time, the best networks in the two hidden-layer and three hidden-layer models improve MSE by 52.44% and 62.27% compared with the best 1HL model response, respectively. Also, the two algorithms of SSA and GWO are able to compete well (comparable MSE) with the PSO and ABC algorithms.

Abbreviation

HT	Heat Transfer
ABC	Artificial Bee Colony
MSE	Mean-Square error
BP	Back-Propagation
LSSVM	Least-Squares Support-Vector Machine
RBF	Radial Basis Function
GNN	Genetic Neural Network
ANN	Artificial Neural Network
MLP	Multilayer Perceptron
PSO	Partial Swarm Optimization
GWO	Grey Wolf Optimizer
SSA	Salp Swarm Algorithm
EHD	Electrohydrodynamic

Nomenclature

p_i	Probability of food source
fit	Fitness vakue
$\vec{X}$	Position vector of a grey wolf
$\vec{X}_p$	Position vector of prey
x_j^1	Position of leader
F_j	Position of food
l	Current iteration
L	Maximum number of iteration
ub	Upper bound
lb	Lower bound
N	Number of neuron
F	Transfer function

Greek letters

α	Leader
β	Subordinates

✉ Myeongsub Kim
kimm@fau.edu

[1] Department of Ocean and Mechanical Engineering, Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431, USA

[2] Faculty of Mechanical and Mechatronics Engineering, Shahrood University of Technology, Shahrood, Iran

δ	Subordinates
ω	Subordinates

Subscripts

i	i Th
j	j Th

1 Introduction

Boiling in power stations, distillation, coolers, and electronic devices has been employed as an excellent thermal management strategy because of its remarkable heat transfer performance [1, 2]. Unlike single-phase heat transfer, boiling heat transfer has the ability to dissipate a substantial amount of heat at a higher rate by means of latent heat absorbed by liquid. During boiling at microcavities, entrapped gases originate bubble formation called nucleation. These nucleated bubbles enlarge and detach from the surface so that heat is eliminated from the surface by phase change [3]. By augmenting the surface temperature, the generated bubbles are gradually increased, leading to enhancing the boiling heat transfer rate. Thus, the behaviors and characteristics of bubbles in the nucleate boiling play a crucial role in understanding and accordingly increasing the heat transfer performance. Various researchers have worked on quantifying the bubble growth dynamics including bubble departure diameter [4–9], growth rate [10–12], waiting time [2, 13, 14], and bubble departure frequency [4, 15–17]. One of the main findings of these studies is that the effects of these parameters on the nucleate boiling heat transfer (HT) is significant.

By manipulating the bubble dynamics parameters, the higher heat dissipation can be achieved, which enhances the HT in nucleate boiling. This HT enhancement is urgently needed with the emerge of new products with a higher heat flux, such as high-power electronic devices. Thus, numerous scholars focused on enhancing the HT coefficient during nucleate boiling [18–21]. Among them, using an electric field to the boiling fluid is appealing since superior boiling heat transfer has been achieved [22, 23]. An external electric field in the HT enhancement was first proposed by Chubb in 1916 [24]. After that, various scientists studied the effects of electrically charged fluids, called electrohydrodynamics (EHD), on boiling heat transfer and fluid flow [25, 26]. For instance, Ogata and Yabe [27] utilized a mesh electrode for studying the boiling heat transfer enhancement by applying EHD. It was concluded that the number of bubbles during nucleate boiling was significantly increased when an external electric field was used. It was also stated that the bubble departure diameter was reduced and the departure frequency was augmented. They reported that a dielectrophoresis force pushed the vapor bubble to the heated surface because of the inequality of the dielectric constants of vapor and liquid. In another experimental study, Karayiannis and Xu [28] examined the boiling HT of R-123 with applying a high DC voltage and rode electrodes. According to their results, the enhancement ratio of boiling HT was more than 4.9 times higher than that with no voltage. Also, Ahmad et al. [29] examined the influences of surface roughness and EHD (5 kV to 25 kV) on the boiling HT enhancement of R123. Based on the results, applying EHD increases heat transfer and make a delay in the critical heat flux. Zu and Yan [30] in a numerical study investigated the impacts of EHD on nucleate boiling enhancement and the bubble shape. It was stated that the bubble was pulled axially which leads to the bubble elongation axially. Furthermore, in an experiment performed by Kweon and Kim [23], a plate-shaped electrode was used to investigate the EHD influence on nucleate boiling around a heated wire. They stated which the average bubble departure diameter and the growth and waiting time were decreased with applying a voltage to the fluid. In contrast, the average bubble frequency and nucleation density were magnified by increasing the electric field strength, leading to the boiling HT enhancement. Pascual et al. [31] worked on the nucleate boiling of R123 with the EHD enhancement technique utilizing a platinum heating wire and a mesh electrode. According to a statistical evaluation, the average bubble departure diameter and the active nucleation sites were decreased by an increase in the electric field strength. Additionally, a heating wire and wire electrode were used by Madadnia and Koosha [32] to evaluate EHD impacts on nucleate boiling. In their results, once the electric voltage and heat flux exceeded 6 kV and 6 kW/m^2, respectively, the enhancement effects were observed with increasing the bubble frequency and the nucleation site density with respect to a zero-voltage condition. It was hypothesized that this could be as a consequence of polarization forces that push bubbles on the heated surface which keeps bubbles longer on it. Recently, Siedel et al. [33] investigated a single bubble formation, growth, and departure from a heated surface with and without applying an external electric field. They reported that the boiling HT was improved and the bubble was enlarged with applying the electric field. However, the departure frequency and growth time were not considerably modified. Chen et al. [34] achieved a conclusion that a bubble was enlarged and its departure frequency and growth time increased while its waiting time reduced under an electric field. However, any conclusion regarding boiling heat transfer enhancement with EHD was not provided. Based on the literature, some impacts of EHD on nucleate boiling have been reported, but discrepancies regarding nucleate boiling with the EHD enhancement still exist due to the complexity of this phenomenon [35–37]. Therefore, modern and accurate methods to predict the bubble behavior in nucleate boiling with the EHD enhancement are required.

Existing computational and experimental approaches to predict the bubble behaviors are commonly either high-priced or sometimes inaccurate [38–41]. Thus, accurate and effective methods to predict the bubble behaviors in nucleate boiling remain a substantial challenge. Some algorithms that learn from data, called machine learning, can be applied to obviate these challenges and this subject has gained significant attentions over the last decade [42–48]. For instance, Wei et al. [49] used an artificial neural network and genetic algorithm (GNN) for predicting the onset of nucleate boiling in a narrow channel contained distilled water. By comparing experimental data and the genetic neural network (GNN), they achieved the conclusion that the model is able to predict the experimental data accurately. Furthermore, different neural networks such as ANFIS (adaptive network-based fuzzy inference system), LSSVM (least-squares support vector machine), and RBF (radial basis function) were applied by Zendehboudi and Tatar [50] for modeling nucleate boiling HT of a refrigerant-oil/nanoparticles mixture. Also, the influences of using nanoparticles on the HT coefficient in various conditions were investigated. According to their results, RBF was the best network to estimate the nucleate boiling HT and the HT coefficient. Balcilar et al. [51] examined characteristics of the nucleate boiling HT with the presence of TiO_2 by using ANN. Their results show that the ANN could estimate the experimental data of the pool boiling HT coefficient with a deviation range of ±5%. Pare and Ghosh [52] used an ANN to model the effects of Al_2O_3/water nanofluid on pool boiling. According to the results, the most accurate model could be achieved by using Levenberg Marquardt training algorithm. Also, Liang et al. [53] estimated the boiling heat transfer in helical coils on the condition of high gravity by using an MLP neural network. Eight input parameters such as helical coil diameter, surface area, inlet pressure and temperature, thermal power, and so on were considered. It was mentioned that the suggested model could predict the heat transfer. Based on these studies, artificial neural networks were able to model heat transfer phenomena in a reasonable range of errors. However, to model the heat transfer with a better accuracy, artificial neural networks can be optimized by different optimization algorithms.

In this study, a multilayer perceptron (MLP) ANN with a back-propagation (BP) training algorithm is applied for modeling the bubble growth and bubble dynamics parameters in nucleate boiling under a non-uniform electric field. The influences of the electric field and heat flux on bubble dynamics parameters including bubble waiting time, bubble departure frequency, bubble growth time, and bubble departure diameter are considered. It is important to note that throughout the study, other influencing parameters such as heated surface orientation, surface roughness, and similar factors remained constant, ensuring a focused examination of the specific effects of voltage and heat flux on the bubble dynamics parameters in nucleate boiling. In addition, a single bubble dynamic behavior of R113 created on a heater is modeled in an inconsistent electric field with MLP neural network and then optimized by four various swarm-based optimization algorithms, namely: Salp Swarm Algorithm (SSA), Grey Wolf Optimizer (GWO), Artificial Bee Colony (ABC) algorithm, and Partial Swarm Optimization (PSO). A throughout literature review illustrates that there is no study focusing on using these algorithms to model nucleate boiling and the associated bubble dynamics parameters. Finally, for the aim of evaluating the accuracy of the models, the MSE value of the models with different optimization algorithms is measured and compared with each other.

2 Data gathering

The utilized data is gathered from the previous experimental studies related to nucleation and bubble dynamics behaviors of R113 formed on a heated surface under variable electric fields [27, 54, 55]. These data illustrate the impacts of the electric field on different bubble dynamics parameters that explain the behavior of a growing bubble, including bubble departure diameter (the bubble diameter while leaving the heated surface), bubble growth time (the time between bubble initiation and departure), bubble departure frequency (the bubble separation's frequency from a nucleation site), and bubble waiting time (the time between departure of the former bubble and creation of the succeeding bubble).

Figure 1 depicts an example experimental setup for the nucleate boiling on a heated surface, consisting of a high-voltage power supplier that generated 0 to 5000 DC voltage and a needle electrode that provides the electric field [27, 54, 55]. There is a distance between the heated surface and the needle so that the needle does not intervene the bubble growth. Additionally, 3-D plots of all 4 parameters with respect to various heat fluxes and voltages are shown in Fig. 2.

3 Structure of neural network

Artificial neural networks (ANN) are a major modeling technique of data in different engineering problems including heat transfer and boiling phenomena [56]. ANNs are information-processing tools developed founded upon the operation of the brain's neural network. The network's processing units are neurons which are connected to each other through communication links, each with an associated weight. A standard neural network has considerable amounts of neurons and their connections [57]. Network structure, transfer function, and learning algorithm make ANN methods distinct from other artificial neural networks.

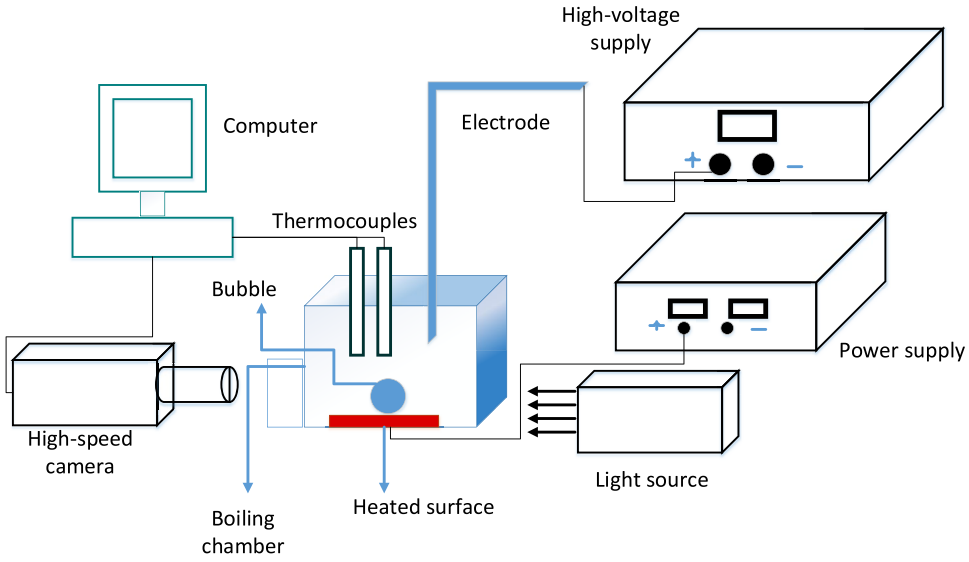


Fig. 1 Experimental setup for nucleate boiling

In this study, a multilayer perceptron (MLP) neural network with a Back-Propagation (BP) training algorithm is utilized to tackle the nucleate boiling problem. In the BP training algorithm on the basis of the supervised learning technique, the neuron connections' weights are adjusted founded upon the discrepancy between the desired network outputs and the predicted outputs [58]. An MLP network composes of one input layer, one output layer, and one or

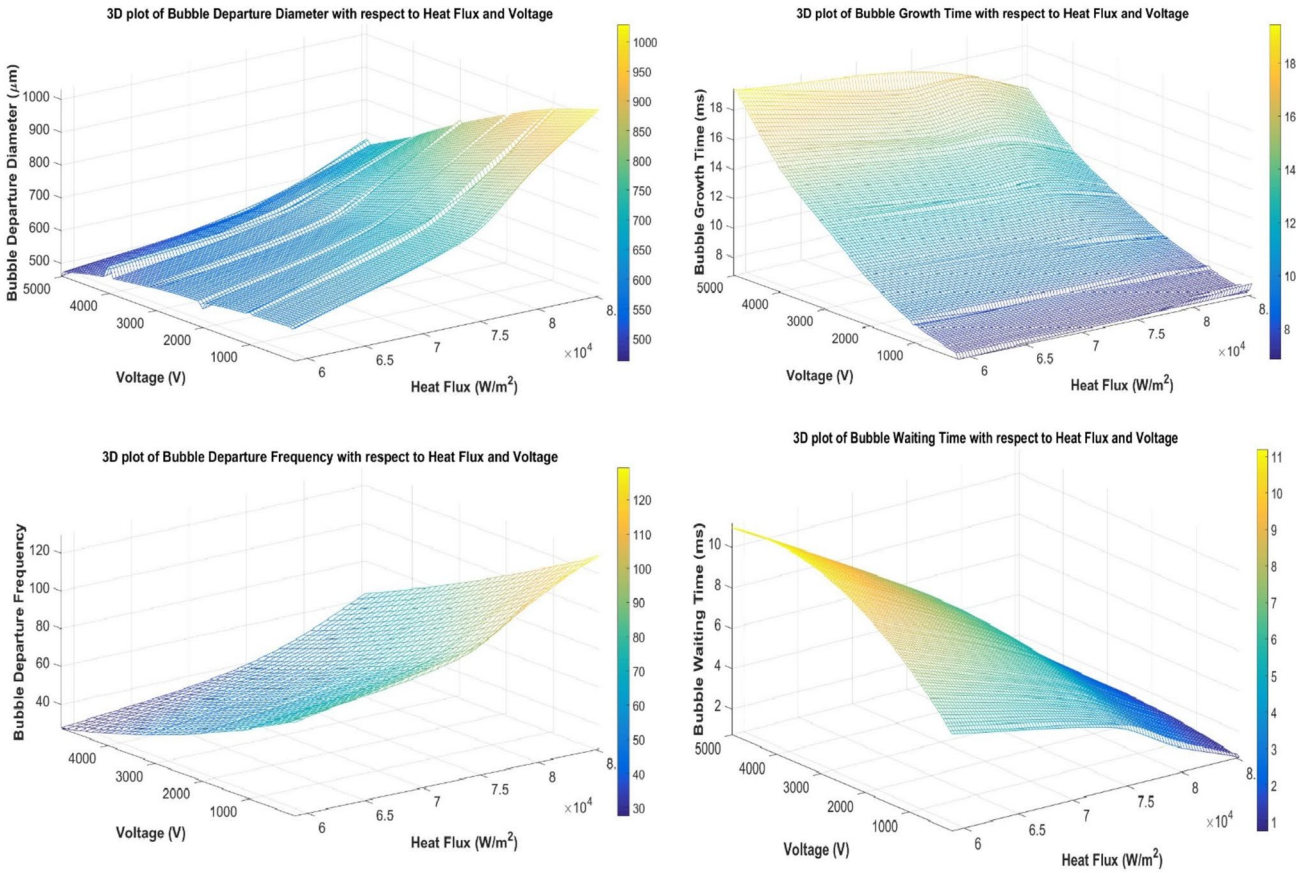


Fig. 2 3-D plots of **a** bubble departure diameter, **b** growth time, **c** bubble departure frequency, **c** bubblewaiting time with respect to heat flux and voltage [55]

more hidden layers (Fig. 3). The inputs' and outputs' number of the problem determines the number of neurons in the input and output layers, whereas the hidden layers' number and the neurons' number in each hidden layer can be selected by designer's choice. Moreover, the designer can choose the neurons' transfer function in the hidden and output layers. Thus, the performance of this method is susceptible to these selectable parameters, in particular the transfer function in each layer and the neurons' number in each hidden layer. The network's architecture can be optimized to produce the minimum error and the best performance.

This study initially models an MLP neural network with one hidden layer and then with two and three hidden layers for the optimization. The optimization problem for the two hidden layers network has four design variables $X = [N_1, N_2, F_1, F_2]^T$, and for the three hidden layers network has six design variables $X = [N_1, N_2, N_3, F_1, F_2, F_3]^T$, where *N*, *F,* and *T* are the number of neurons, the type of transfer function, and the transpose of the design variable matrix, respectively. To evaluate the effectiveness of the optimization, the value of the Mean Squared Error (MSE) is considered as the cost function. The optimization problem is shown in Eq. (1),

$$\begin{aligned} &\text{Min } MSE\,(X),\\ &X:\{N_j,,F_j\}\\ &\text{S.t (subject to): } <1<N_j<30\\ &\qquad F_j\in\{logsig,\ tansig,\ elliotsig\} \end{aligned} \tag{1}$$

$i = 1, 2$ For two hidden layers network

$i = 1, 2, 3$ For three hidden layers network

where N_i and F_i are the neuron number and the transfer function in the *i*th hidden layer, respectively. It is assumed that the transfer function is selected among three cases; Log-sigmoid transfer function (sigmoid), Hyperbolic tangent sigmoid transfer function (tansig), and Elliot symmetric sigmoid transfer function (elliotsig). Also, the lower band the neurons' number in each hidden layer is 1 and the upper bound is 30. The design parameters of this optimization problem are discrete variables, and the optimization algorithms must search for optimal variables to be integer values.

4 Optimization algorithms

Various optimization algorithms are implemented, making network outputs and actual outputs (experimental data) get closer, in order to reduce modeling errors of the nucleate boiling heat transfer. These include four different swarm-based optimization algorithms, namely: Salp Swarm Algorithm, Grey Wolf Optimizer, Artificial Bee Colony, Partial Swarm Optimization. During the optimization, the MSE value reaches its minimum by manipulating the design variables. The swarm-based methods imitate groups of animals and their social behaviors which are one of meta-heuristic optimization methods [59]. Table 1 shows the applied optimization algorithms with the corresponding year of development.

4.1 Particle swarm optimization

Particle swarm optimization (PSO) was established in accordance with social behaviors of birds, bees, and fish and their dynamic movements and each species is considered "particle" [60]. The particle path is modified based on the particle's knowledge (individual) and the swarm's knowledge (group) which is adapted during iterations. The final solution of this algorithm is the position of the swarm. If a swarm has *P* particles, a position vector $X_j^t = (x_{j1}, x_{j2}, x_{j3}, \ldots x_{jn})^T$ and a velocity vector $V_j^t = (v_{j1}, v_{j2}, v_{j3}, \ldots v_{jn})^T$ at *t* iteration for each one of the *j* particle ($j = 1,2, \ldots, P$) can be considered [61]. Also, the *j*th particle with the best location is called $P_{best,j}$, and its corresponding location for the swarm is represented by G_{best} in each iteration. For determining the *j*th

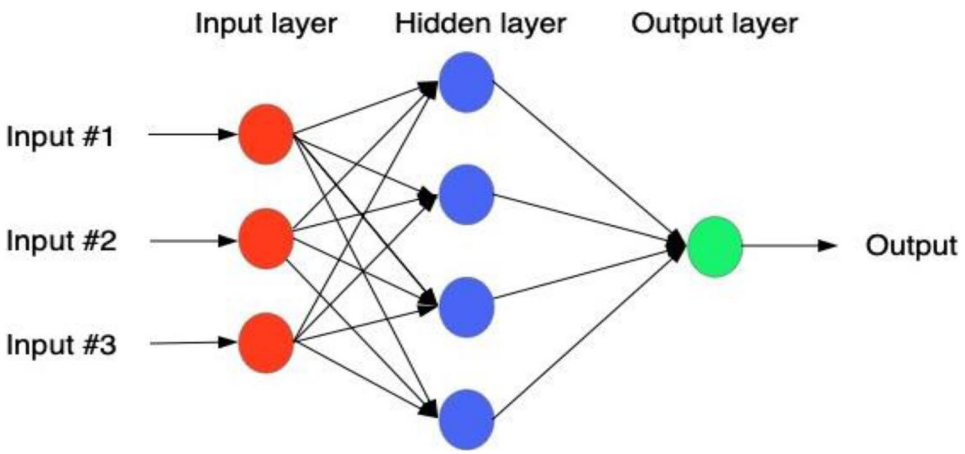


Fig. 3 Schematic of a neural network structure

Table 1 Optimization algorithms used in this study

No.	Applied Optimization Algorithm	Proposed Year
1	Particle Swarm Optimization	1995
2	Artificial Bee Colony	2005
3	Grey Wolf Optimizer	2014
4	Salp Swarm Algorithm	2017

particle position and velocity, the following equations can be utilized [62, 63],

$$X_j(i) = X_j(i-1) + V_j(i), \qquad j = 1, 2, \ldots, P. \tag{2}$$

$$V_j(i) = \gamma V_j(i-1) + c_1 u_1 \left[P_{best,j} - X_j(i-1)\right] + c_2 u_2 \left[G_{best} - X_j(i-1)\right], \quad j = 1, 2, \ldots, P. \tag{3}$$

where c_1 and c_2 are constants defined as learning coefficients of cognition for an individual and group, respectively. u_1 and u_2 are two random numbers in [0, 1]. Finally, $\gamma(i)$ is a weighting factor which has a value between 0.4 and 0.9. This factor can be kept constant or reduced with rising the iterations' number.

4.2 Artificial bee colony algorithm

Artificial bee colony (ABC) algorithm is a popular optimization algorithm due to its simplicity having few control parameters. The solution of this algorithm is a food source's position based on bees' search [64, 65]. In ABC, three kinds of bees are considered in the bee colony: employed bees, onlooker bees, and scout bees when 50% of the colony includes the employed and scout bees and the rest are the onlookers. An employed bee turns into scout when a food source by it. It is assumed that solution 's quality (fitness) is indicated by the food source's nectar amount. In ABC, an initial solution is created, and the algorithm iteratively goes through three phases: the employed bee, onlooker bee, and scout bee phases till a satisfaction criterion is met.

The steps of ABC algorithms are:

1. Initialization of parameters including the food source number, a satisfaction criterion, and limit (the number of trials before abandoning a food source).
2. Population initialization of solutions which is randomly created by [66],

$$x_{ij} = x_j^{min} + rand\,(0,1) \times \left(x_j^{max} - x_j^{min}\right) \tag{4}$$

where x_{ij} and *rand* are the food source's position and a random number between [0, 1]. *min* and *max* indicate the minimum and maximum values of x_{ij}.

3. In the employed bee phase, the bee is in charge of exploiting possible food sources (new) around its initially designated food sources (old). The old food source is substituted by the new one once the old one's nectar quantity is lower than that of the new one.
4. In the next phase, an onlooker bee assesses all the information gathered by the employed bees and picks out a food source relying upon the probability contributed to the food source (p_i), that is measured by the following equation [66],

$$p_i = \frac{fit_i}{\sum_{n=1}^{SN} fit_n} \tag{5}$$

where fit_i represents the fitness value of the solution and *SN* is referred to the food sources' number. In this step, the solutions are modified by the onlooker bee to generate new food sources near the initialized food source position by using,

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj}) \tag{6}$$

where $k \in \{1, 2, \ldots, SN\}$ and v is the posision of the newly generated food source. $\phi_{i,j}$ is a random number in [-1, 1] which regulates the generation of a neighbor food source adjacent to $x_{i,j}$. According to Eq. (6), the perturbation associated with the position $x_{i,j}$ decreases while the discrepancy between the parameters of $x_{i,j}$ and $x_{k,j}$ is reduced. Therefore, the step length (change in the target value at each step) is appropriately decreased while the search reaches the optimal solution in the search space.
5. In the scout bee phase, if an employed bee cannot further improve a food source (based on the limit), the bee leaves the food source and turns into a scout bee. Next, the bee arbitrarily searches a new food source to replace the abandoned food.
6. Steps 3 to 5 are repeated till a satisfaction criterion is fulfilled.

4.3 Grey wolf optimizer (GWO)

A grey wolf optimization algorithm (GWO) provides competitive results in comparison with renowned algorithms like PSO [67]. In this algorithm, grey wolves are regarded as predators that are mainly in a group that follow an extremely severe social dominant hierarchy as depicted in Fig. 4. Alpha (α) is located on the top as the leader of the group including a male and a female. The alpha is mainly in charge of determining hunting. Other members of the group are commitment to the alpha's decisions [68]. Betas are the second level of the grey wolves' hierarchy. The beta can be considered as the representatives helping alphas in making decision or other group activities. The beta orders the lower-level wolves and follow the alpha. The lowest level in the hierarchy of the grey wolf is omega. The omega has to adhere to all the other predominant wolves and these wolves have the

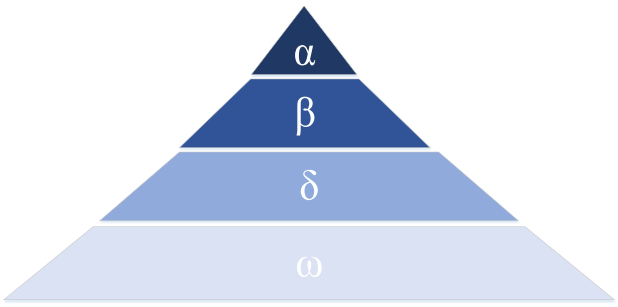


Fig. 4 Hierarchy of grey wolf (dominance decreases from top to down) [67]

last permission to eat among others. A wolf is called subordinate (delta) that is not an alpha, beta, or omega. Delta has to obey alpha and beta, however, they govern the omega.

In GWO, the most suitable solution is the alpha (α) while beta (β) and delta (δ) are the second and third best solutions. α, β, and δ direct the hunting (optimization) and the ω wolves obey these three wolves. The succeeding relation have been introduced for modeling the wolves' encircling conduct mathematically [67],

$$\vec{D} = \left|\vec{C}.\vec{X}_p(t) - \vec{X}(t)\right| \tag{7}$$

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A}.\vec{D} \tag{8}$$

where t is referred to the current iteration. $\vec{A}$ and $\vec{C}$ indicate coefficient vectors. $\vec{X}$ and $\vec{X}_p$ indicate a grey wolf's position vector and the position vector of the prey, respectively.

The following equations are used to calculate $\vec{A}$ and $\vec{C}$ [67],

$$\vec{A} = 2\vec{a}.\vec{r_1} - \vec{a} \tag{9}$$

$$\vec{C} = 2.\vec{r_2} \tag{10}$$

where $\vec{a}$'s components are linearly reduced from 2 to 0 during iterations. Also, r_1 and r_2 are random vectors in [0, 1]. By utilizing Eqs. (7) and (8), the position of a grey wolf is upgraded inside the space near the prey in any random location [67]. Next, for mathematically simulating the hunting attitude of grey wolves, the following equations have been introduced [67],

$$\begin{aligned}\overrightarrow{D_\alpha} &= \left|\overrightarrow{C_1}.\vec{X}_\alpha - \vec{X}\right|,\\ \overrightarrow{D_\beta} &= \left|\overrightarrow{C_2}.\vec{X}_\beta - \vec{X}\right|,\end{aligned} \tag{11}$$

$$\begin{aligned}\overrightarrow{D_\delta} &= \left|\overrightarrow{C_3}.\vec{X}_\delta - \vec{X}\right|\\ \overrightarrow{X_1} &= \vec{X}_\alpha - \overrightarrow{A_1}.\left(\overrightarrow{D_\alpha}\right),\\ \overrightarrow{X_2} &= \vec{X}_\beta - \overrightarrow{A_2}.\left(\overrightarrow{D_\beta}\right),\end{aligned} \tag{12}$$

$$\begin{aligned}\overrightarrow{X_3} &= \vec{X}_\delta - \overrightarrow{A_3}.\left(\overrightarrow{D_\delta}\right)\\ \vec{X}(t+1) &= \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3}\end{aligned} \tag{13}$$

where D_α, D_β, and D_δ are the updated positions of search agents. Equation (13) is used to upgrade the current search agent's position for the next iteration.

4.4 Salp swarm algorithm (SSA)

A salp swarm algorithm (SSA) is a recently developed algorithm which shows effective behaviors in finding the optimum solutions. This algorithm is generally inspired by the swarm conduct of salps, called a salp chain (Fig. 5).

For modeling the salp chain mathematically, the population is initially split into two parts: followers and leader. The salp located at the front of the chain is regarded the leader whilst the others are followers. The guidance of the swarm is the leader's responsibility and the followers obey each other. The salps' position is determined in an n-dimensional search space in which n represents the variables' number in a specific problem. In modeling, a 2-D matrix named x is used to store the position of all salps. Also, it is presumed that the swarm's target is a food source named F in the search space. The following equation has been introduced for updating the leader's position [70],

$$x_j^1 = \begin{cases} F_j + c_1\left(\left(ub_j - lb_j\right)c_2 + lb_j\right) & c_3 \geq 0\\ F_j - c_1\left(\left(ub_j - lb_j\right)c_2 + lb_j\right) & c_3 < 0\end{cases} \tag{14}$$

where F_j and x_j^1 indicate the food source's position and the leader's position in the jth dimension, respectively. ub_j and lb_j represent the upper and lower bounds of jth dimension, respectively. c_2, and c_3 are random numbers in [0, 1], and c_1 is determined by the succeeding equation [70],

$$c_1 = 2e^{-\left(\frac{4l}{L}\right)^2} \tag{15}$$

where l and L represent the current iteration and the maximum number of iterations, respectively.

Additionally, the subsequent equation has been used for updating the followers' position [70],

$$x_j^i = \frac{1}{2}(x_j^i + x_j^{i-1}) \tag{16}$$

In summary, the steps of the SSA algorithm are first, SSA parameters such as the salp population, upper and lower bounds, and maximum number of iterations are initiated. In the next step, the fitness of each salp is measured and the salp with the best fitness is found out. Then, the best salp's position is assigned to F which is the food source. Meantime, c_1 can be upgraded by Eq. (15). Also, the positions of

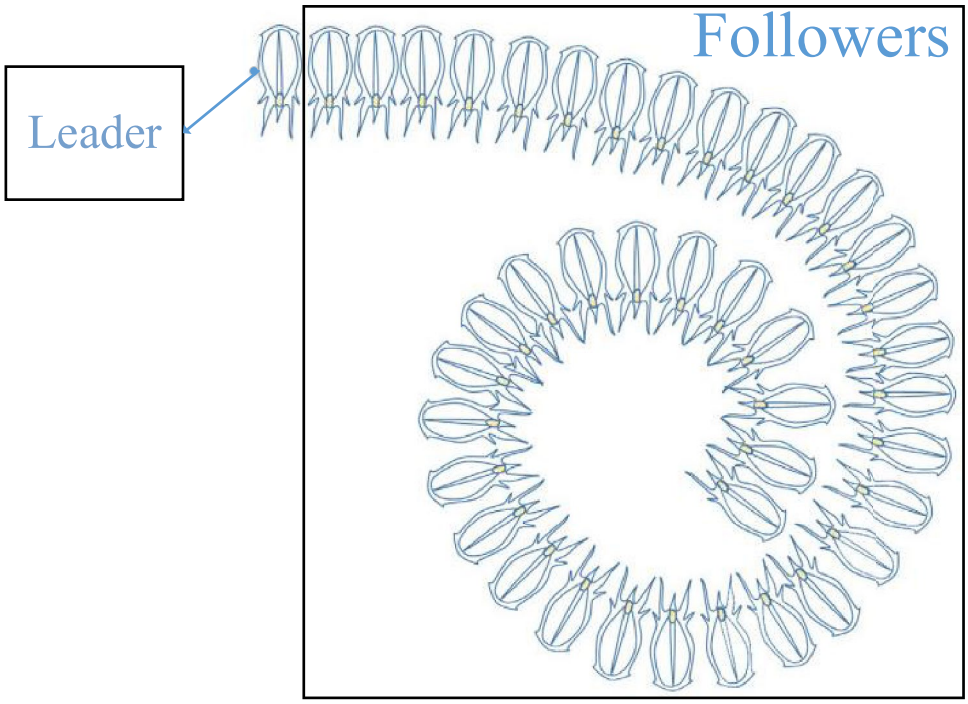


Fig. 5 Schematic of a swarm of salp (salp chain) [69]

the leader and the followers can be upgraded by Eqs. (14) and (16) in each dimension, respectively. Then, if any salp passes the boundaries of the search space, it is brought back on the basis of the upper and lower bounds. All the mentioned steps are iteratively performed except the initialization until an end criterion is satisfied.

5 Results and discussion

In this section, the optimization for four bubble dynamics parameters including bubble waiting time, bubble departure frequency, bubble growth time, bubble departure diameter are initially provided using the one hidden-layer network. Then, these parameters are modeled and compared by optimizing two- and three-hidden layer networks with four optimization algorithms including PSO, ABC, GWO, and SSA. Also, the best neural network for each bubble dynamics parameter is provided by comparison of MSE values in four different optimization algorithms. For modeling the boiling HT with the neural networks, the obtained dataset is split into three datasets randomly, so that 70% of the dataset is utilized for training, 15% for testing, and the last 15% for validation. To train the network, the Levenberg–Marquardt algorithm is applied [71]. The network's performance was determined based on the MSE values between the desired output and the predicted output from the networks. The input variables and the number of samples for each case are presented in Table 2. All datasets are normalized before using in the neural network.

5.1 One hidden layer (1HL) network

For the neural network with one hidden layer, only two design variables $X=[N_1, F_1]^T$ are considered for the optimization problem. Figure 6 shows the typical structure of the 1HL network. By considering the N_1 and F_1 values with 30 and 3 levels, respectively, the problem has 90 various cases. By testing all cases, the optimal design variables that have a minimum MSE value were chosen. Table 3 shows the best results for four different bubble dynamics parameters.

Figures 7, 8, 9 and 10 show the regression diagrams for the best MLP network with one hidden layer (in accordance with the specifications in Table 3) for all four parameters including bubble waiting time, bubble departure frequency, bubble departure diameter, bubble growth time. In these diagrams, the MLP network with one hidden layer can approximate the desired outputs based on the values of coefficient of determination. The coefficient

Table 2 Inputs and the number of samples for each output parameter

Output	Inputs		Number of samples
Bubble departure diameter	Heat Flux (W/m^2)	Voltage (V	740
Bubble growth time	Heat Flux (W/m^2)	Voltage (V)	724
Bubble departure frequency	Heat Flux (W/m^2)	Voltage (V)	668
Bubble waiting time	Heat Flux (W/m^2)	Voltage (V)	684

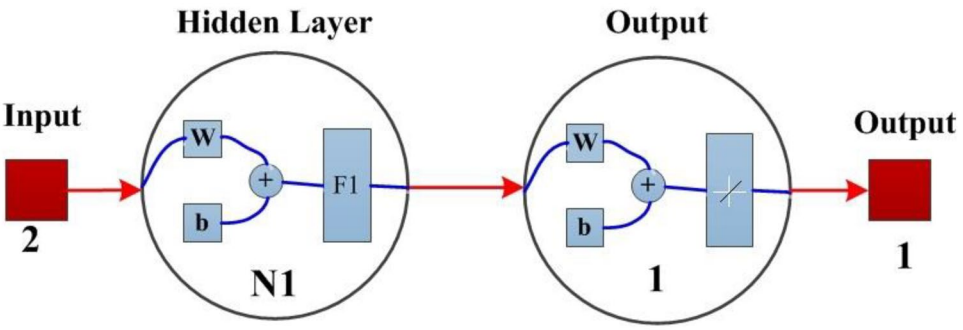


Fig. 6 A typical structure of the 1HL network

of determination (R^2) proves the ability of the suggested models in predicting the bubble dynamics parameters data once it is close to 1. Based on Fig. 7, the regression diagrams of experimental and estimated values for the bubble departure diameter show the values of R^2 for training and all datasets are 0.99914 and 0.99842, respectively. Most of the data points for both test and training datasets are condensed near the line that indicates the precise estimation of the suggested models. Also, similar results can be obtained from Figs. 8, 9 and 10 for the bubble growth time, bubble departure frequency, and bubble waiting time, respectively. The coefficients of determination (R^2) for the bubble growth time for the training and all datasets are 0.9995 and 0.99938, respectively. Based on these measured values, the model is considerably powerful in predicting the bubble dynamics parameters. Also, Figs. 7, 8, 9 and 10 also verify the precision and the prediction ability of the MLP neural network.

5.2 Two hidden layers (2HL) network

In this section, two hidden layers is considered in the neural network. The optimization problem for the 2HL network has four design variables $X = [N_1, N_2, F_1, F_2]^T$. Figure 11 shows a typical structure of the 2HL network. For the optimization process by using different algorithms, the population sizes and the numbers of maximum iterations are selected as 10 and 100, respectively.

a) Bubble departure diameter

For bubble departure diameter, the results of optimizing the MLP network with 2 hidden layers applying 4 various algorithms including PSO, ABC, GWO, and SSA are provided in Table 4 and Fig. 12. It can be observed that the lowest MSE value belongs to the ABC algorithm and then followed by the SSA algorithm. The lowest MSE value of the bubble departure diameter is 4.31×10^{-5} and the second lowest MSE belongs to SSA algorithm which is 4.37×10^{-5}. It should be noted that this small difference could be significant if the number of experimental data points is less than the number of data points used in this study. Under this condition, to have a precise prediction, the best model should be selected. Also, in comparison with different transfer functions, the tansig function is seen more than others in optimal populations. Additionally, the trend of changes in the error (MSE) in four optimization algorithms are plotted according to the number of iterations in Fig. 13. As it is clear, the MSE value decreases in each iteration for all of the optimization algorithms. It should be noted that the initial population in these methods is randomly created, making the starting points of these algorithms in Fig. 13 different.

b) Bubble growth time

Table 5 and Fig. 14 show the results of MLP network optimization with 2 hidden layers for bubble growth time output using four optimization algorithms. Based on these results, the ABC method has the least error, followed by the GWO and PSO algorithms that have less MSE values than the SSA algorithm. The lowest MSE value of the bubble growth time is 7.1×10^{-6} so it can be considered as the most accurate model in predicting the bubble growth time in the 2HL network. The percentage of error increase between ABC and GWO is 3.5%. Also, in comparison of different transfer functions, tansig is still more common than the others in optimal populations. Figure 15 shows the trend of error changes according to the iterations for 4 optimization methods. The figure suggests that it is possible to obtain the optimal value of the ABC algorithm before reaching 50 iterations, which indicates the strength of this algorithm.

c) Bubble departure frequency

Table 3 The best results for four bubble dynamics parameters using the 1HL network

Output	N_1	F_1	MSE	R^2
Bubble departure diameter	29	logsig	6.5166e-05	0.99894
Bubble growth time	24	logsig	1.2802e-05	0.99938
Bubble departure frequency	27	tansig	3.861e-06	0.99994
Bubble waiting time	27	logsig	4.0665e-06	0.99996

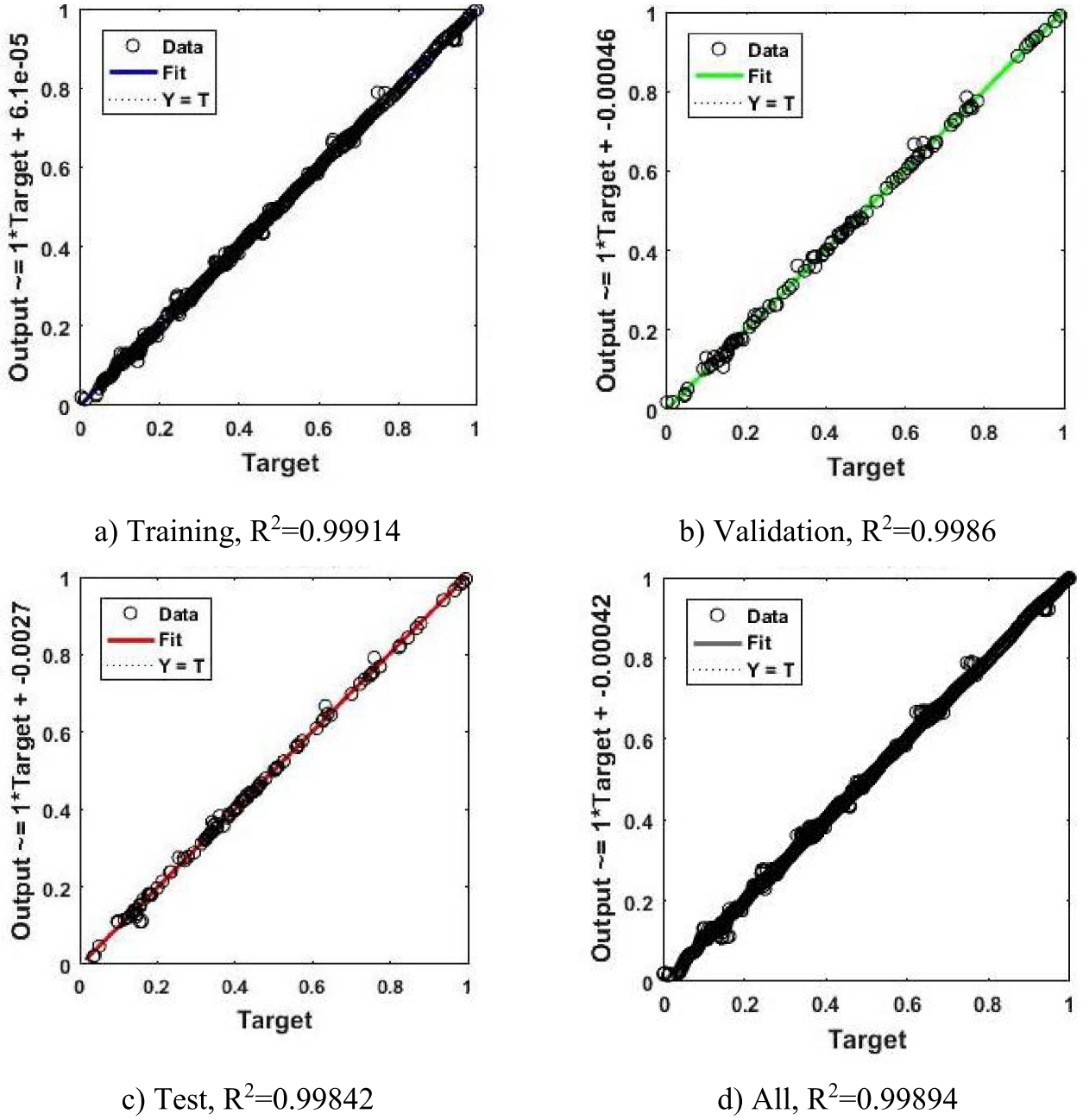


Fig. 7 Regression diagrams of the best MLP network for bubble departure diameter with the 1HL network

The results of bubble departure frequency optimized by the MLP networks with the 2HL network are presented in Table 6 and Fig. 16. Based on these results, the ABC method has the lowest error and the other three algorithms follow it. The lowest MSE value for the bubble departure frequency is 2.05×10^{-6}. Also, the logsig function can be observed more than the rest in the optimal populations. The trend of changing the MSE values (performance) in Fig. 17 shows that the ABC algorithm reaches the optimal value much faster at 41 iterations. As it is clear, the MSE value decreases in each iteration for all of the optimization algorithms which is highly favorable.

d) Bubble waiting time

As the final output, the bubble waiting time is optimized by the same technique. Similar to the previous three outputs, Table 7 and Fig. 18 show that the ABC algorithm performs best, followed by the PSO algorithm with less error than the remaining two algorithms. To

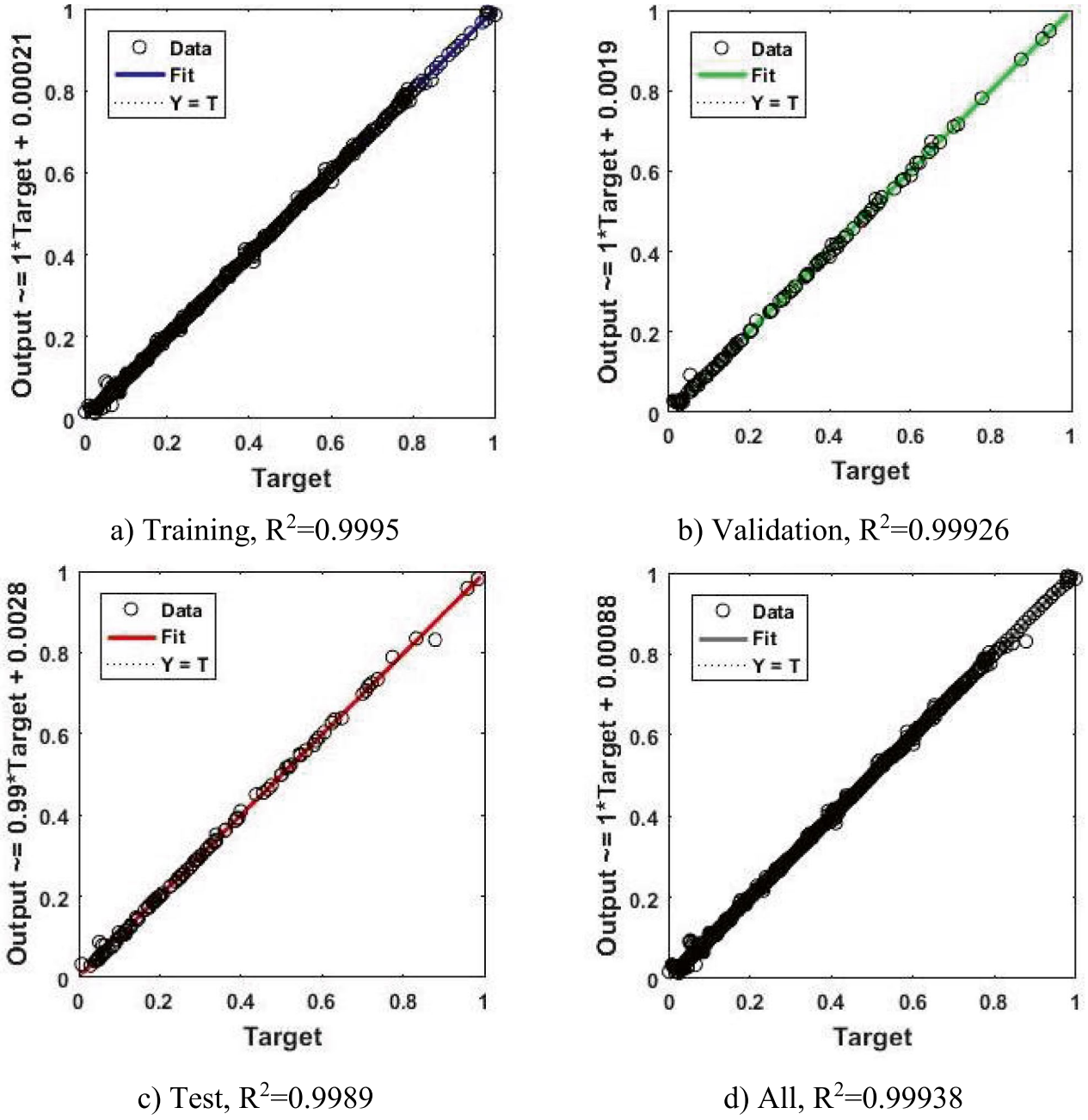


Fig. 8 Regression diagrams of the best MLP network for bubble growth time with the 1HL network

be specific, the lowest MSE value of the bubble waiting time in the 2HL network is 1.93×10^{-6}. Also, in comparison of different transfer functions, tansig and elliotsig are more common than logsig in optimal populations. Figure 19 also shows the trend of error changes for all four utilized algorithms. This figure shows that the ABC algorithm reaches the optimal value at iterations less than 100.

5.3 Three hidden layers (3HL) network

In this section, we consider three hidden layers (3HL) in the neural network. The optimization problem for the 3HL network has six design variables $X=[N1, N2, N3, F1, F2, F3]^T$. Figure 20 shows the typical structure of the 3HL network. Similar to the 2HL network optimization, the population sizes and the numbers of maximum iterations are considered as 10 and 100, respectively.

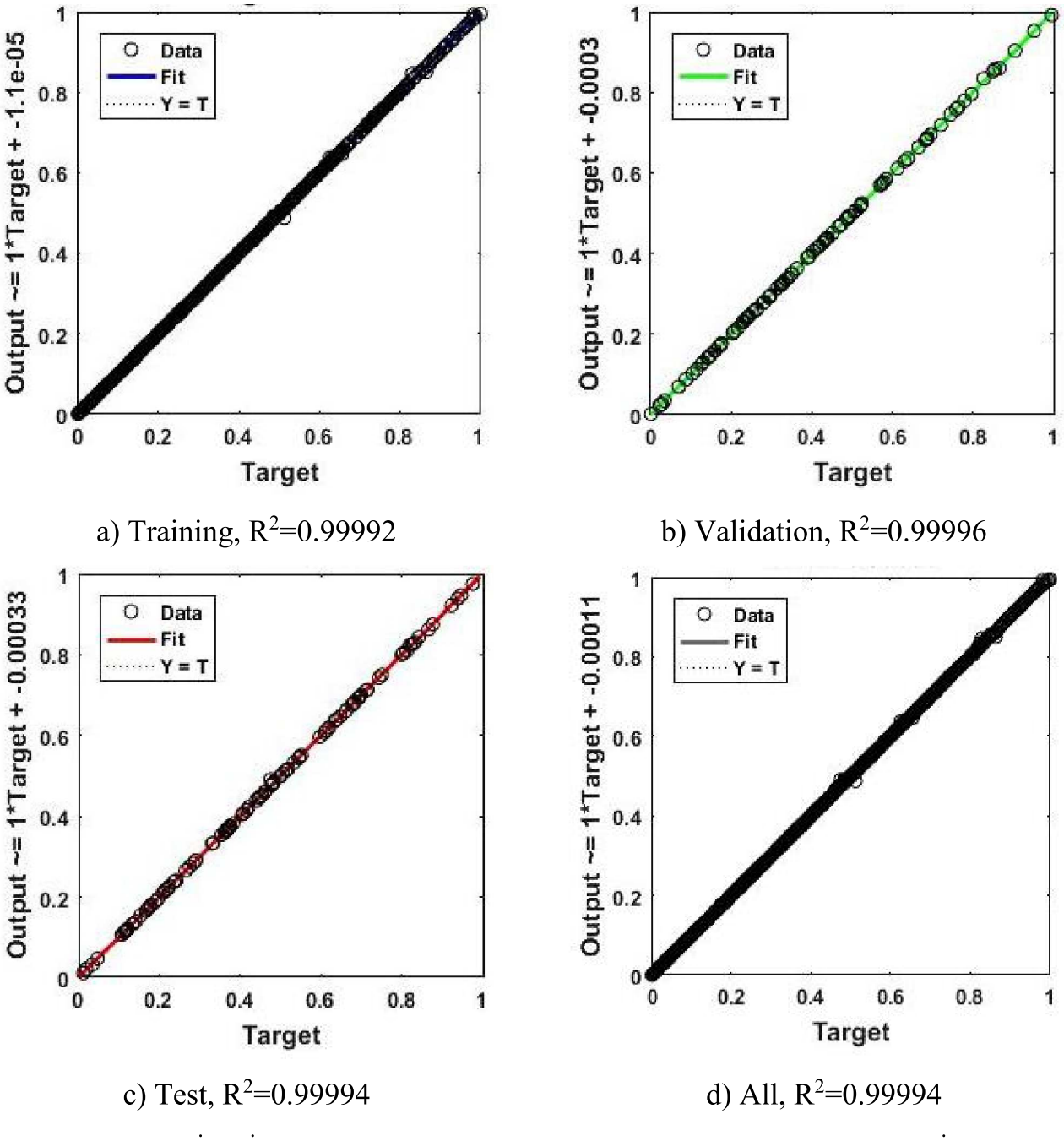


Fig. 9 Regression diagrams of the best MLP network for bubble departure frequency with the 1HL network

a) Bubble departure diameter

The results for bubble departure diameter with the 3HL network using four optimization algorithms are provided in Table 8. Figure 21 illustrates the error comparison between the four optimization methods. Based on these results, the PSO algorithm has the lowest MSE value and the SSA and GWO algorithms are in the next ranks. The lowest MSE value of the bubble departure diameter for the 3HL network is 4.22×10^{-5}. Interestingly, the ABC algorithm did not work properly in the 3HL network because its error for this output is much greater than that in the 2HL network. The trend of error changes in Fig. 22 shows that the PSO algorithm reaches the optimal value in 15 iterations. Also, in comparison

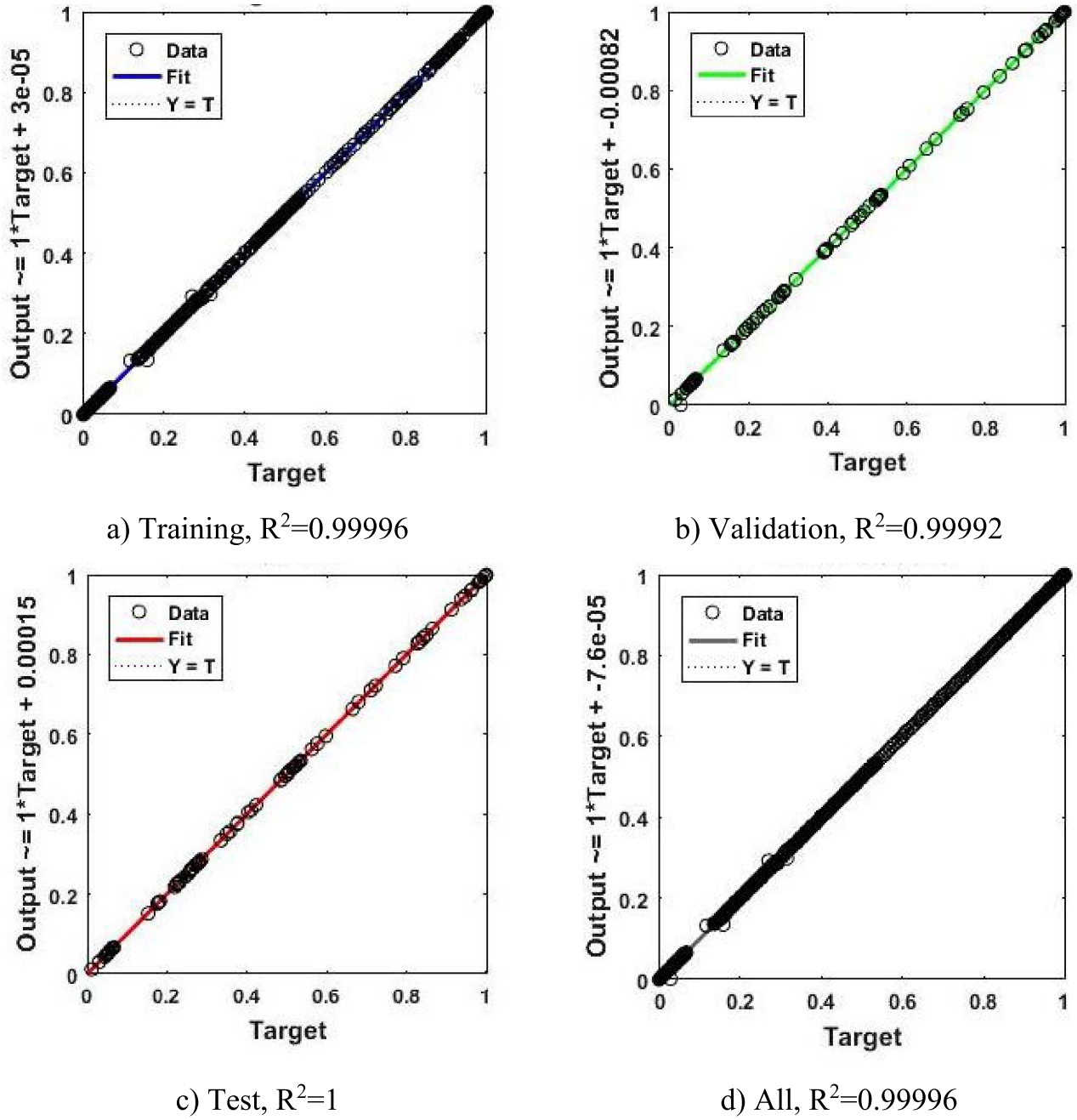


Fig. 10 Regression diagrams of the best MLP network for bubble waiting time with one 1HL network

with different transfer functions, logsig and elliotsig are more common than tansig in optimal populations of the PSO algorithm which shows the best performance among the algorithms.

b) Bubble growth time

The results for bubble growth time are presented in Table 9 and Figs. 23 and 24 based on the 3HL network optimized by four different algorithms. According to these results, the GWO method has the lowest MSE value and then followed by the SSA method when compared with the remaining two algorithms. The lowest MSE value of the bubble growth time in the 3HL network is 6.93×10^{-6} and the second lowest MSE is 7.32×10^{-6}. It should be noticed that the SSA error is almost 5% higher than GWO algorithm for the bubble growth time in the 3HL network. In terms of transfer functions, elliotsig can be observed less in optimal populations than the other transfer functions.

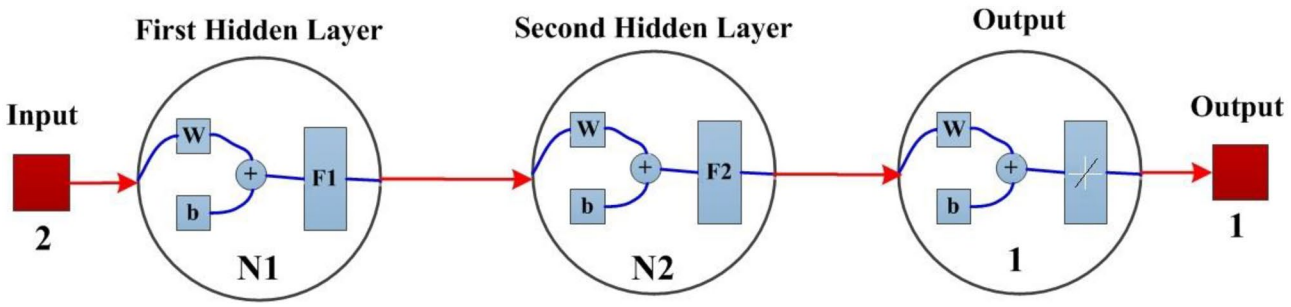


Fig. 11 A structure of the 2HL network

Table 4 Optimum values of the design parameters for bubble departure diameter with the 2HL network

	N_i (number of neurons)		F_i (transfer function)		MSE
	N_1	N_2	F_1	F_2	
PSO	30	25	tansig	logsig	4.4664e-05
ABC	27	27	tansig	tansig	4.3087e-05
GWO	27	13	tansig	logsig	4.4623e-05
SSA	23	24	tansig	tansig	4.3660e-05

Table 5 Optimum values of the design parameters for bubble growth time with the 2HL network

	N_i (number of neurons)		F_i (transfer function)		MSE
	N_1	N_2	F_1	F_2	
PSO	26	20	tansig	tansig	7.3738e-06
ABC	25	29	tansig	logsig	7.1041e-06
GWO	30	21	tansig	tansig	7.3454e-06
SSA	29	21	tansig	tansig	7.9099e-06

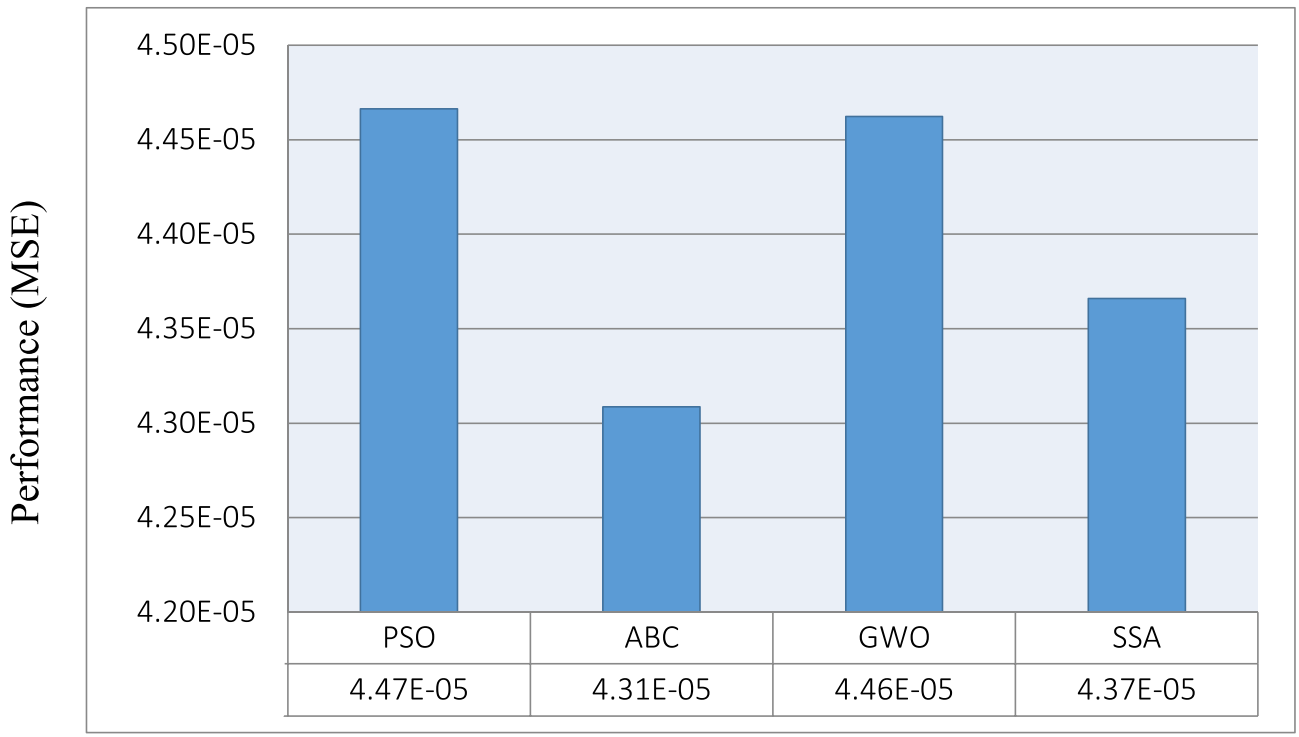


Fig. 12 Comparison of MSE values in four different optimization algorithms for bubble departure diameter with the 2HL network

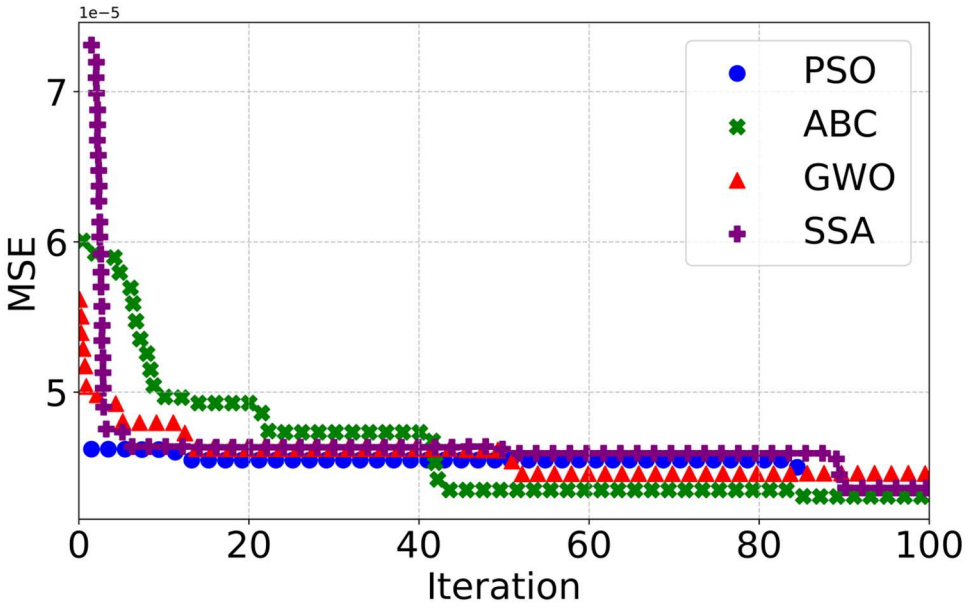


Fig. 13 Performance evolution of four optimization algorithms for bubble departure diameter with the 2HL network

Figure 24 shows that the two algorithms, GWO and SSA, have higher errors in the initial population, but are eventually able to achieve lower errors than ABC and PSO, which indicates that the capabilities of these new optimization methods are considerable in predicting experimental data.

c) Bubble departure frequency

The results of bubble departure frequency are provided in Table 10 using four different algorithms. Figure 25 shows a comparison of MSE values for all implemented algorithms. The results reveal that the PSO algorithm has the lowest MSE value which

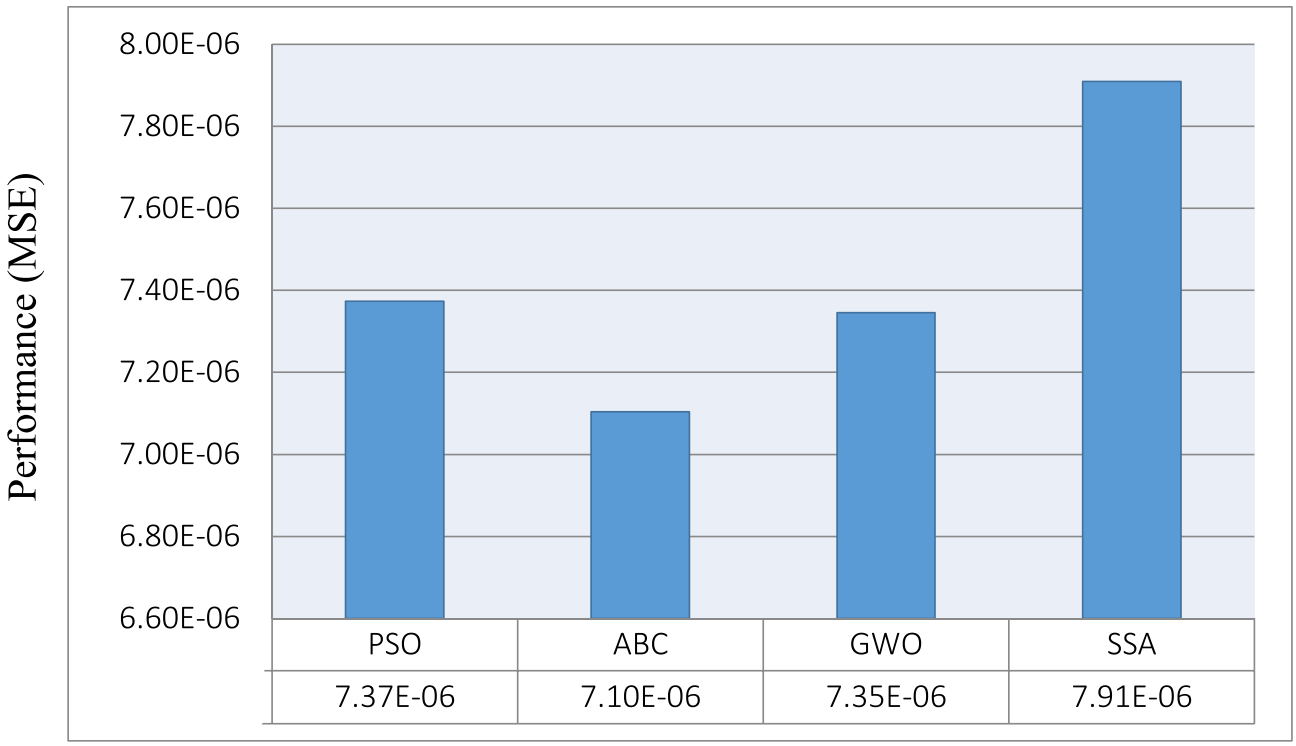


Fig. 14 Comparison of MSE values in four different optimization algorithms for bubble growth time with the 2HL network

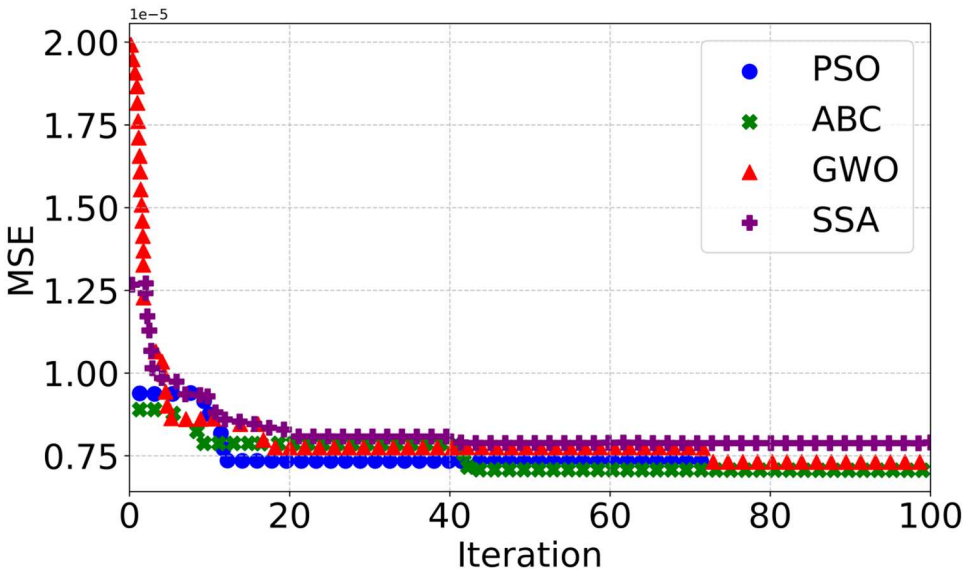


Fig. 15 Performance evolution of four optimization algorithms for bubble growth time with the 2HL network

Table 6 Optimum values of the design parameters for bubble departure frequency with the 2HL network

	N_i (number of neurons)		F_i (transfer function)		MSE
	N_1	N_2	F_1	F_2	
PSO	23	27	tansig	logsig	2.5135e-06
ABC	22	16	elliotsig	tansig	2.0521e-06
GWO	24	27	elliotsig	logsig	2.5522e-06
SSA	27	20	logsig	logsig	2.5314e-06

is 1.46×10^{-6} and then followed by the SSA method. The highest MSE value for the bubble departure frequency of the 3HL belongs to ABC algorithm which is 2.34×10^{-6}. The error increase between PSO and ABC is 60%, which is significant. For transfer functions, elliotsig is seen more than others in the optimal populations. Additionally, the trend of error changes in Fig. 26 indicates that the PSO algorithm reaches the optimal value in the 50th iteration.

d) Bubble waiting time

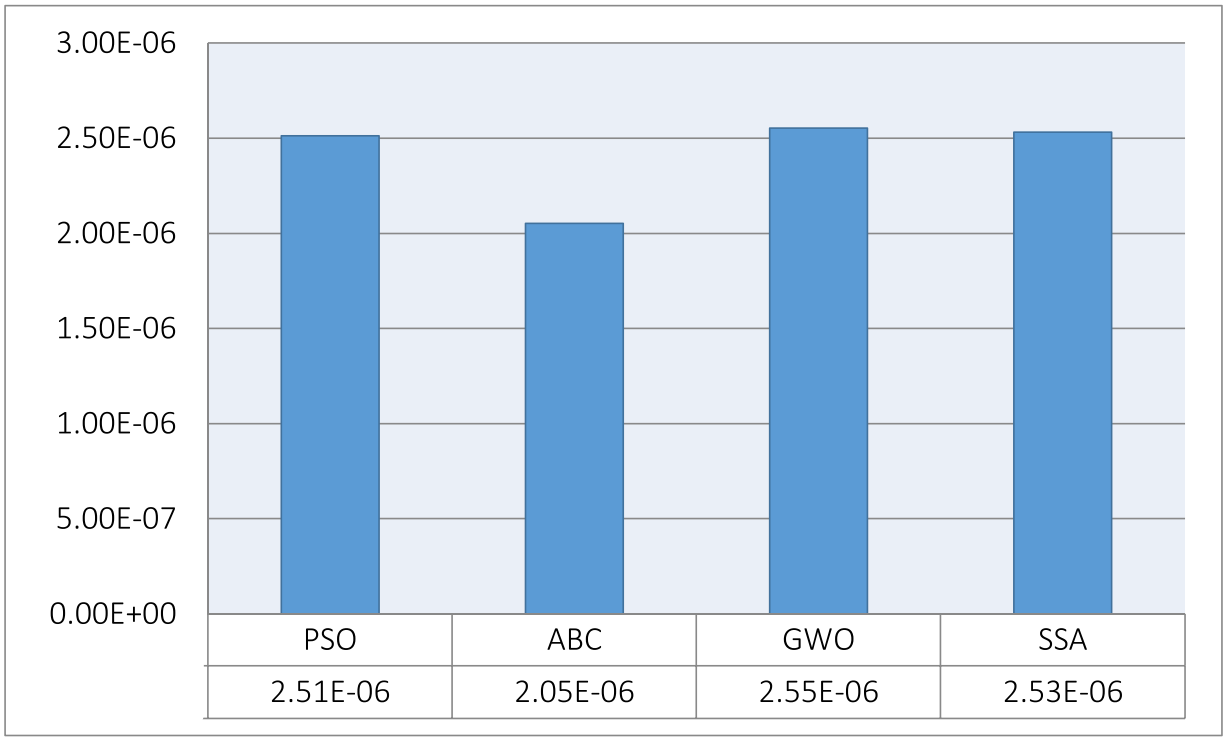


Fig. 16 Comparison of MSE values in four different optimization algorithms for bubble departure frequency with the 2HL network

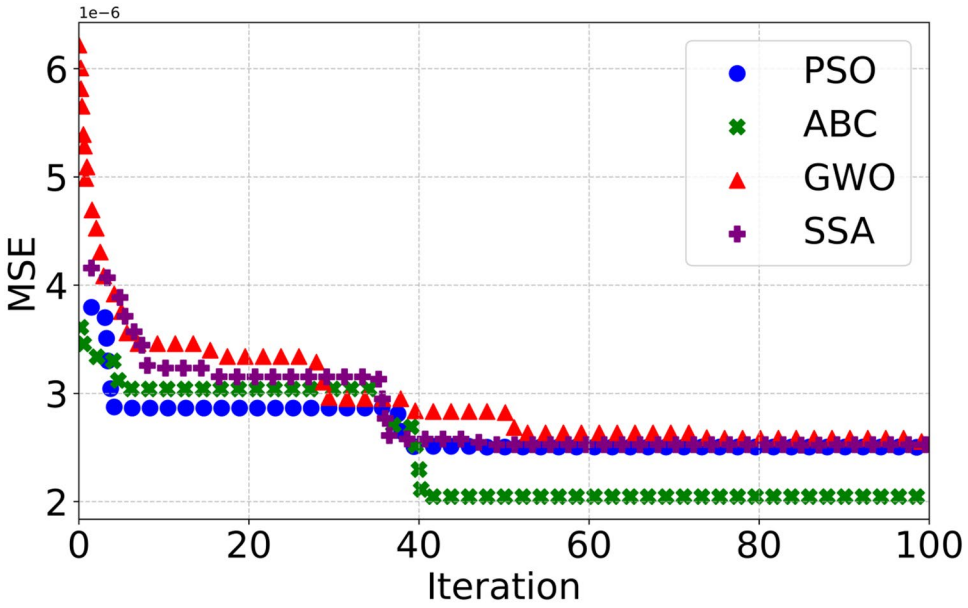


Fig. 17 Performance evolution of four optimization algorithms for bubble departure frequency with the 2HL network

As the final output, the bubble waiting time is optimized, and the results are provided in Table 11 and Figs. 27 and 28. According to the results, the SSA algorithm which is followed by the GWO algorithm has the lowest MSE value as compared to ABC and PSO methods. Also, by comparing different transfer functions, the tansig function is less than the other two transfer functions in optimal populations. Additionally, by considering the trend of error changes in Figure 28, it can be mentioned that the SSA algorithm reaches the optimal value in the 55th iteration. As it is clear, the MSE value decreases in each iteration for all of the optimization algorithms which is highly favorable.

5.4 Comparison of results between different hidden layers

This section examines the influences of the number of hidden layers in the MLP network on the reduction of MSE value. The results of four bubble dynamics parameters optimized by four optimization algorithms in the 2HL and 3HL networks are compared. Figure 29 shows that in the SSA and GWO algorithms, the optimal network error in the 3HL model is always less than the optimal network in the 2HL model. In the PSO algorithm, the error in the 3HL model is less than that in the 2HL model for the bubble departure diameter and departure frequency while the error in the 3HL model is more than that in the 2HL model for the bubble growth time and bubble waiting time. In all cases of the ABC algorithm, the error in the 2HL model is less than that in the 3HL model.

In Tables 12, 13, 14 and 15, the best values of the least MSE are compared in all algorithms according to the number of hidden layers. Table 12 illustrates that the optimal networks in the 2HL and 3HL models for the bubble departure diameter improve by 33.85% and 35.27%, respectively, when compared with the best response in the 1HL model. Additionally, Table 13 shows that networks with 2HL and 3HL have the 44.51% and 45.85% reduction in error for the bubble growth time when compared with the network with 1HL, respectively. Furthermore, Tables 12 and 13 show that increasing the number of layers from 1 to 2 is significant in bubble departure diameter and bubble growth time, but changing the number of layers from 2 to 3 has a minimal effect on reducing errors. Table 14 shows that the error in 2HL and 3HL networks for departure frequency is reduced by 46.85% and 62.32%, respectively, when compared with the best 1HL network. For the bubble waiting time, Table 15 shows that the best 2HL and 3HL networks reduce the MSE values by 52.44% and 62.27% when compared with the best 1HL model response, respectively. Consequently, for bubble departure frequency and bubble waiting time, raising the number of hidden layers from 1 to 2 and then from 2 to 3 in the best networks could have a significant effect on reducing network errors.

Table 7 Optimum values of the design parameters for bubble waiting time with the 2HL network

	N_i (number of neurons)		F_i (transfer function)		MSE
	N_1	N_2	F_1	F_2	
PSO	19	19	elliotsig	tansig	2.0735e-06
ABC	23	28	elliotsig	logsig	1.9339e-06
GWO	25	8	elliotsig	tansig	2.8184e-06
SSA	28	29	elliotsig	tansig	2.4132e-06

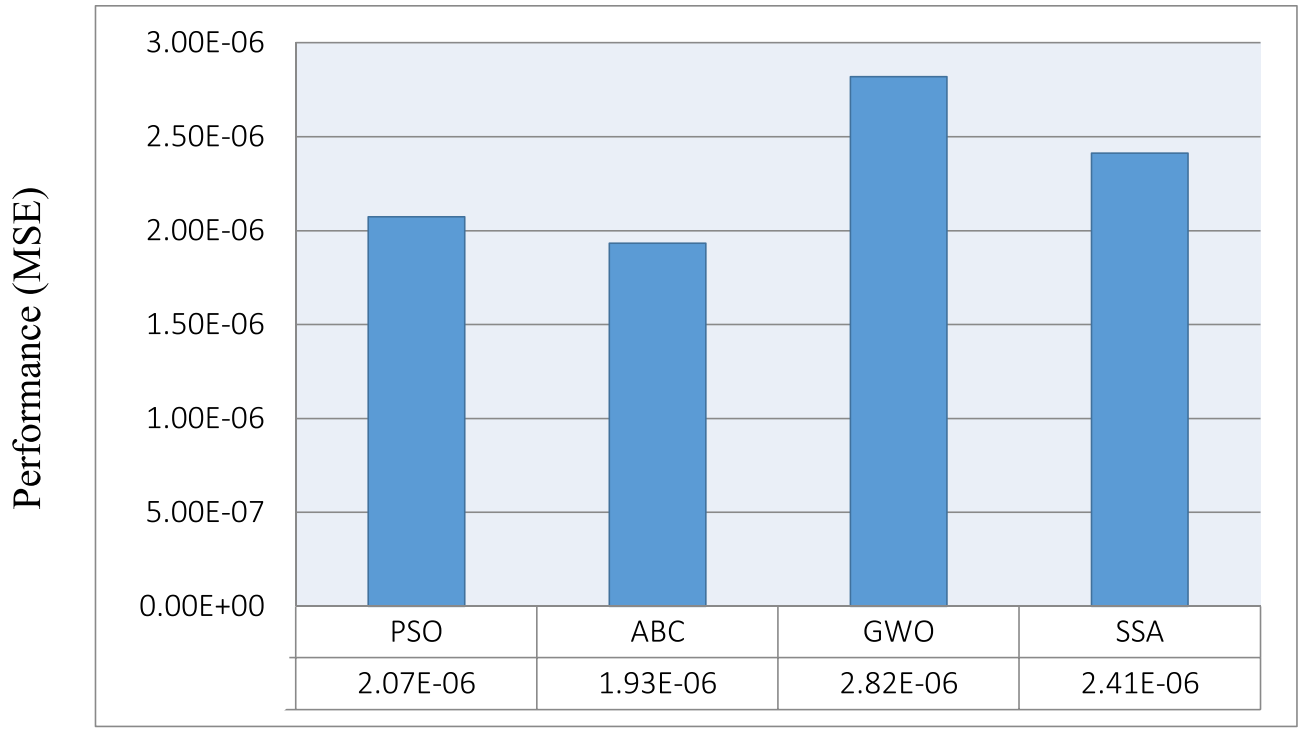


Fig. 18 Comparison of MSE values in four different optimization algorithms for bubble waiting time with the 2HL network

Additionally, Figs. 30, 31, 32 and 33 show the error analysis for different outputs in terms of the percentage of the relative deviation and the error histogram using the best 3HL network. The relative deviation shows how well the individual numbers agree with each other. In other words, it can tell how precise the average of results is. It can be seen that the relative deviation for the bubble departure diameter, the bubble growth time, the bubble departure frequency, and the bubble waiting time is in the range of ± 4%, approximately. Additionally, the error histogram is a graph of the errors between predicted values from the neural network and target values. This shows how the target values and predicted values are close and how the errors from the neural network are spread. For achieving a precise model, the error distribution diagram is needed to obey the normal distribution diagram. Based on Figs. 31, 32 and 33, the errors

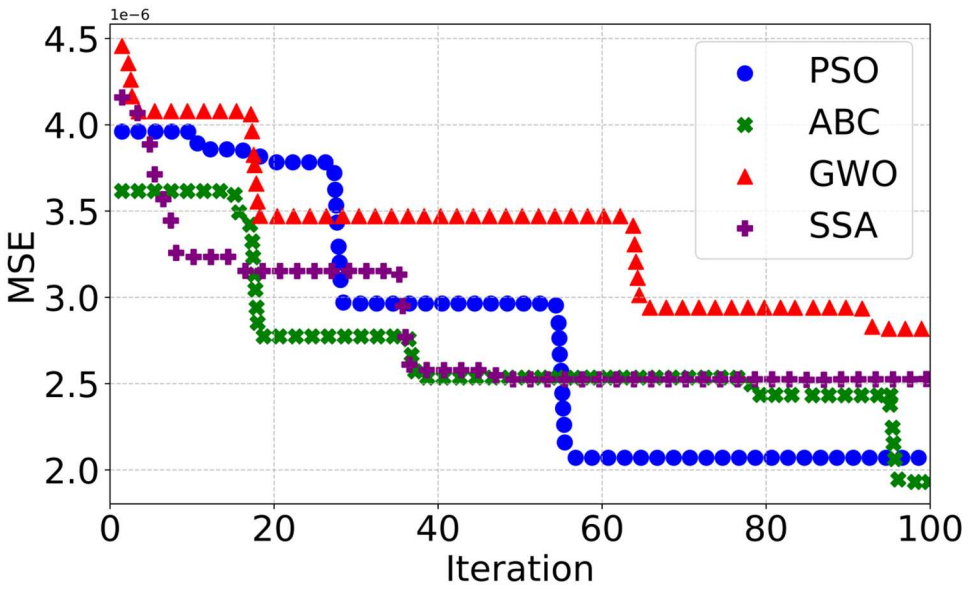


Fig. 19 Performance evolution of four optimization algorithms for bubble waiting time with the 2HL network

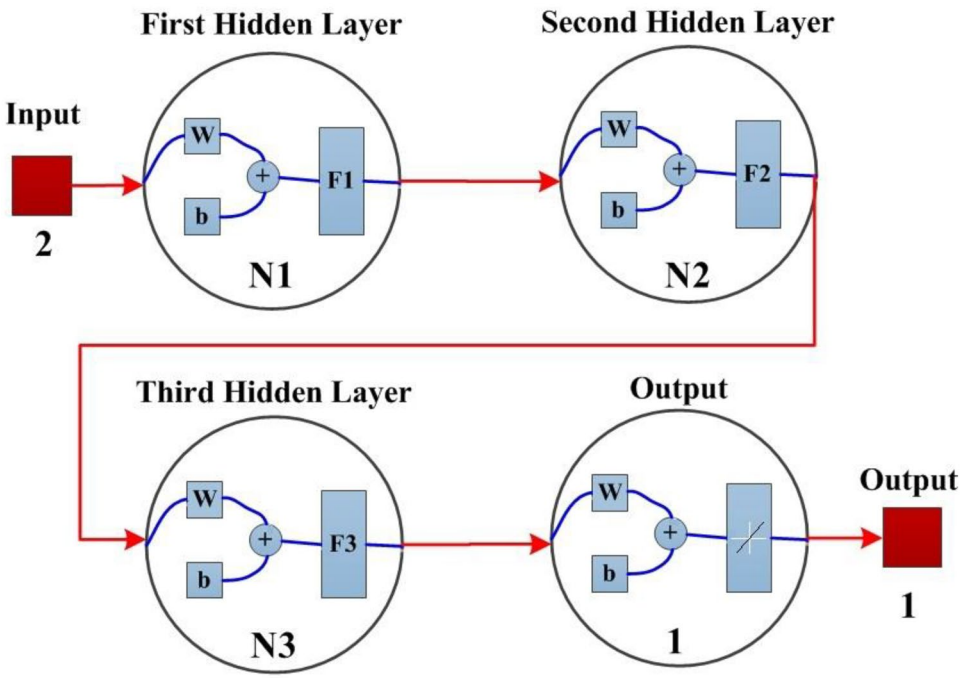


Fig. 20 A structure of the 3HL network

Table 8 Optimum values of the design parameters for bubble departure diameter with the 3HL network

	N_i (number of neurons)			F_i (transfer function)			MSE
	N_1	N_2	N_3	F_1	F_2	F_3	
PSO	27	18	9	elliotsig	logsig	logsig	4.2180e-05
ABC	23	14	18	elliotsig	elliotsig	logsig	4.3914e-05
GWO	25	21	16	tansig	tansig	tansig	4.2541e-05
SSA	28	28	19	elliotsig	logsig	tansig	4.2537e-05

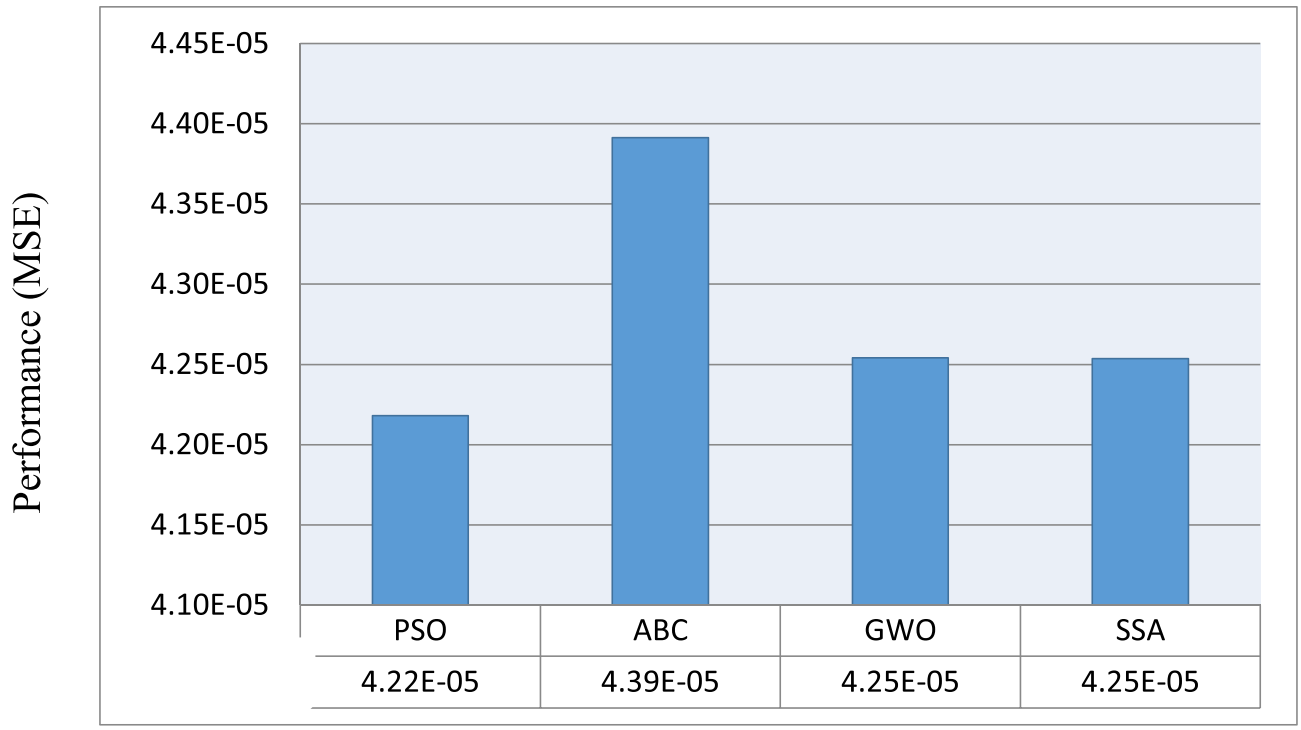


Fig. 21 Comparison of MSE values in four different optimization algorithms for bubble departure diameter with the 3HL network

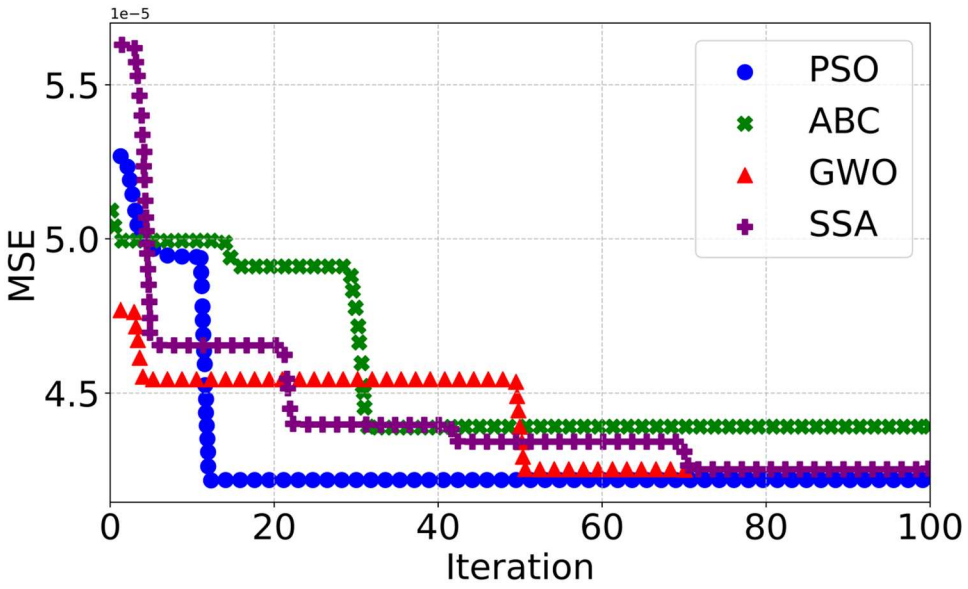


Fig. 22 Performance evolution of four optimization algorithms for bubble departure diameter with the 3HL network

Table 9 Optimum values of the design parameters for bubble growth time with the 3HL network

	N_i (number of neurons)			F_i (transfer function)			MSE
	N_1	N_2	N_3	F_1	F_2	F_3	
PSO	20	20	22	elliotsig	logsig	tansig	8.1514e-06
ABC	15	18	11	tansig	logsig	logsig	8.1041e-06
GWO	30	25	22	logsig	elliotsig	tansig	6.9321e-06
SSA	11	30	12	tansig	tansig	logsig	7.3150e-06

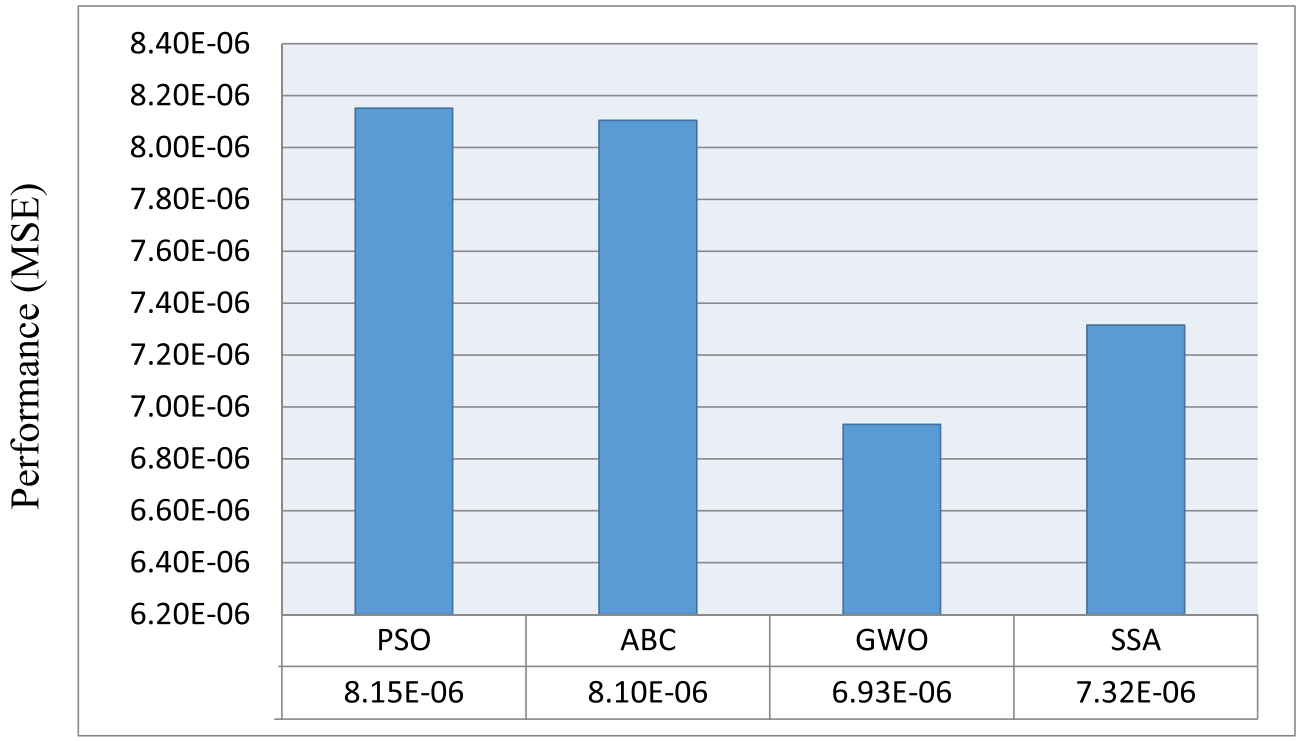


Fig. 23 Comparison of MSE values in four different optimization algorithms for bubble growth time with the 3HL network

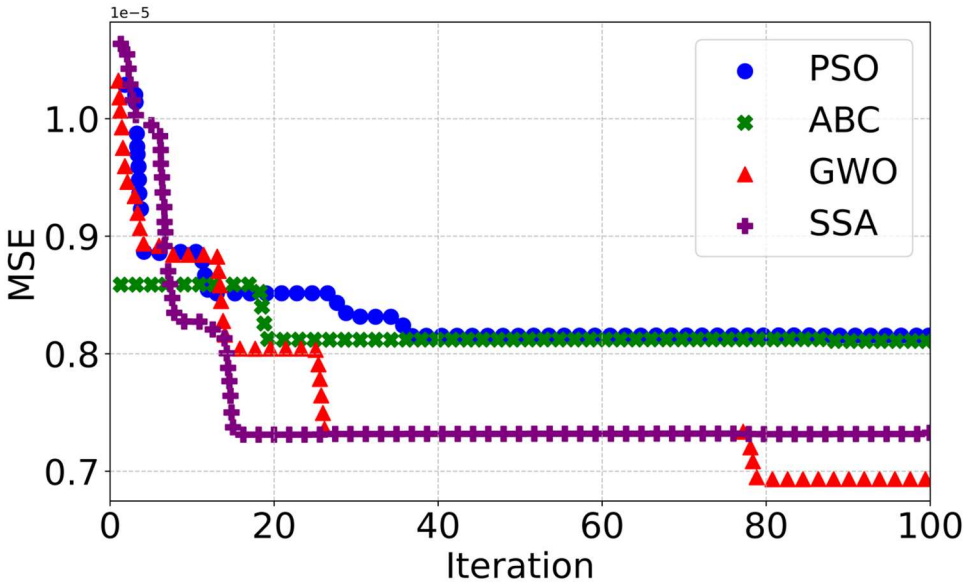


Fig. 24 Performance evolution of four optimization algorithms for bubble growth time with the 3HL network

Table 10 Optimum values of the design parameters for bubble departure frequency with the 3HL network

	N_i (number of neurons)			F_i (transfer function)			MSE
	N_1	N_2	N_3	F_1	F_2	F_3	
PSO	22	18	9	elliotsig	tansig	elliotsig	1.4550e-06
ABC	10	22	23	elliotsig	elliotsig	elliotsig	2.3375e-06
GWO	23	21	11	elliotsig	tansig	elliotsig	2.1460e-06
SSA	27	30	1	elliotsig	tansig	logsig	2.1192e-06

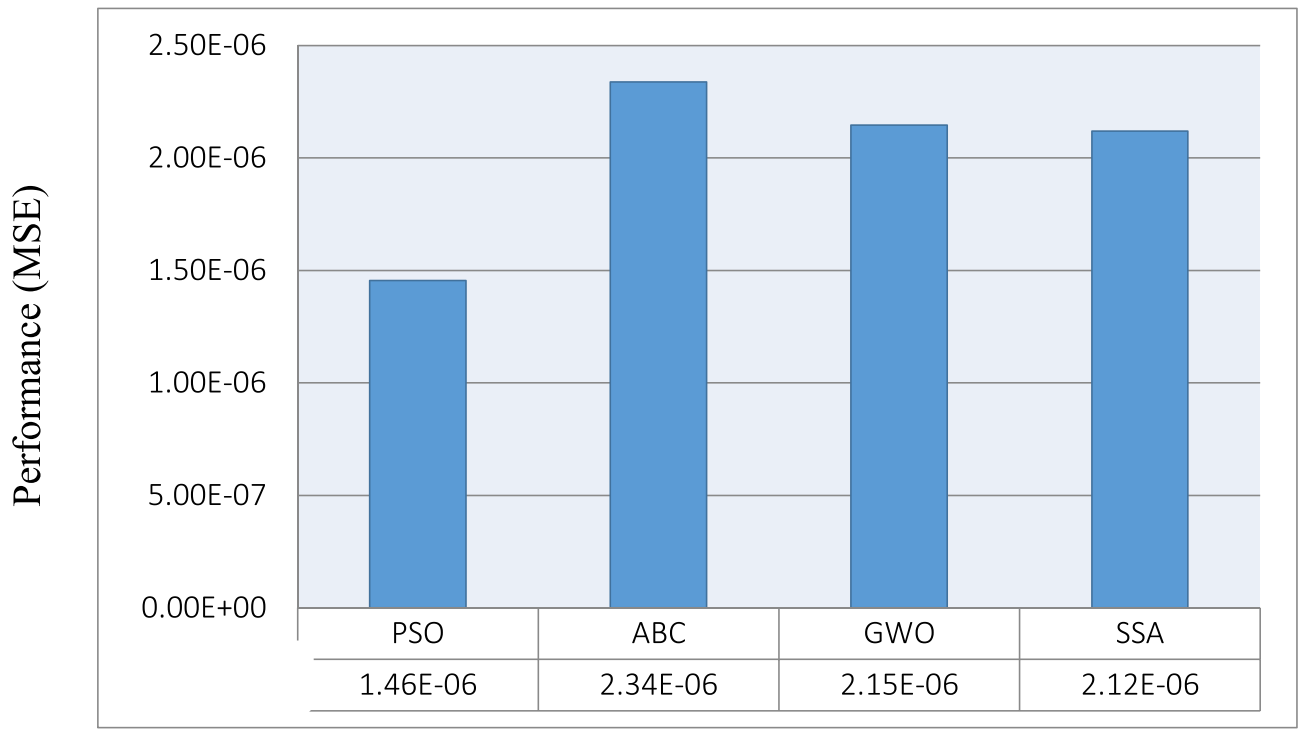


Fig. 25 Comparison of MSE values in four different optimization algorithms for bubble departure frequency with the 3HL network

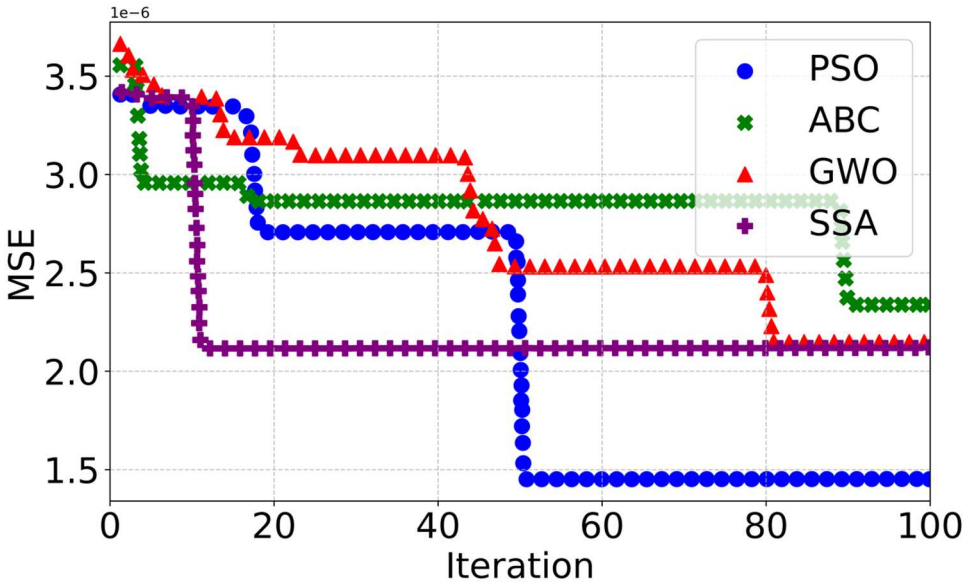


Fig. 26 Performance evolution of four optimization algorithms for bubble departure frequency with the 3HL network

Table 11 Optimum values of the design parameters for bubble waiting time with the 3HL network

	N_I (number of neurons)			F_I (transfer function)			MSE
	N_1	N_2	N_3	F_1	F_2	F_3	
PSO	13	28	4	elliotsig	logsig	logsig	2.7252e-06
ABC	22	9	27	elliotsig	logsig	elliotsig	2.7031e-06
GWO	17	23	30	tansig	elliotsig	elliotsig	2.1545e-06
SSA	29	30	19	logsig	logsig	tansig	1.4528e-06

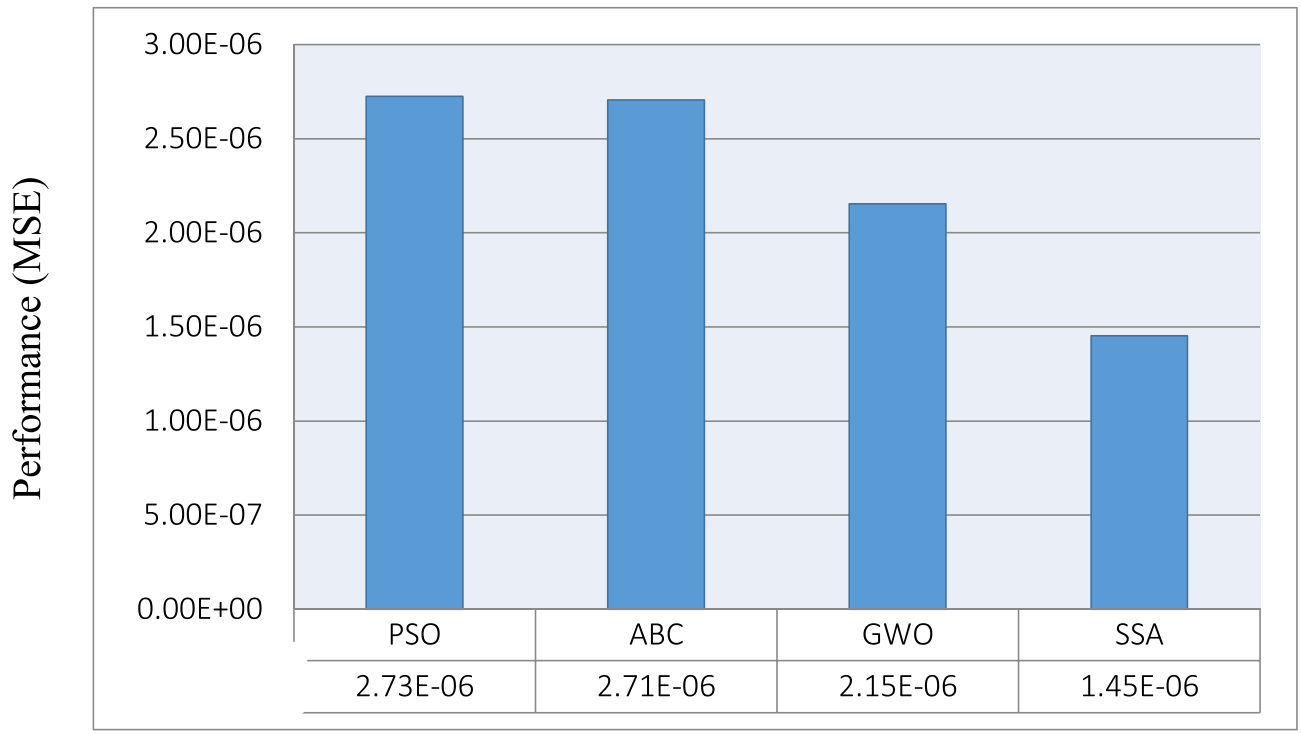


Fig. 27 Comparison of MSE values in four different optimization algorithms for bubble waiting time with the 3HL network

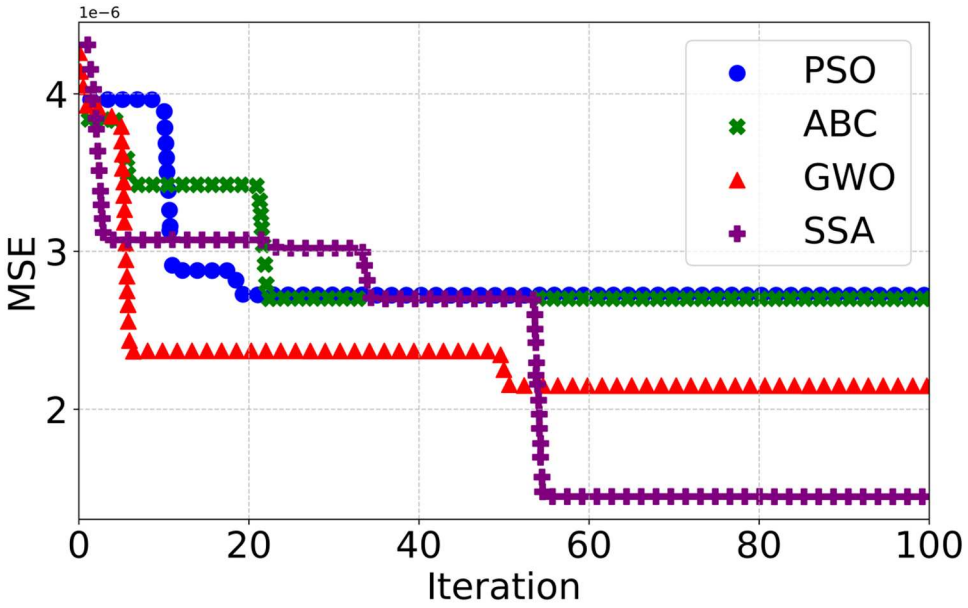


Fig. 28 Performance evolution of four optimization algorithms for bubble waiting time with the 3HL network

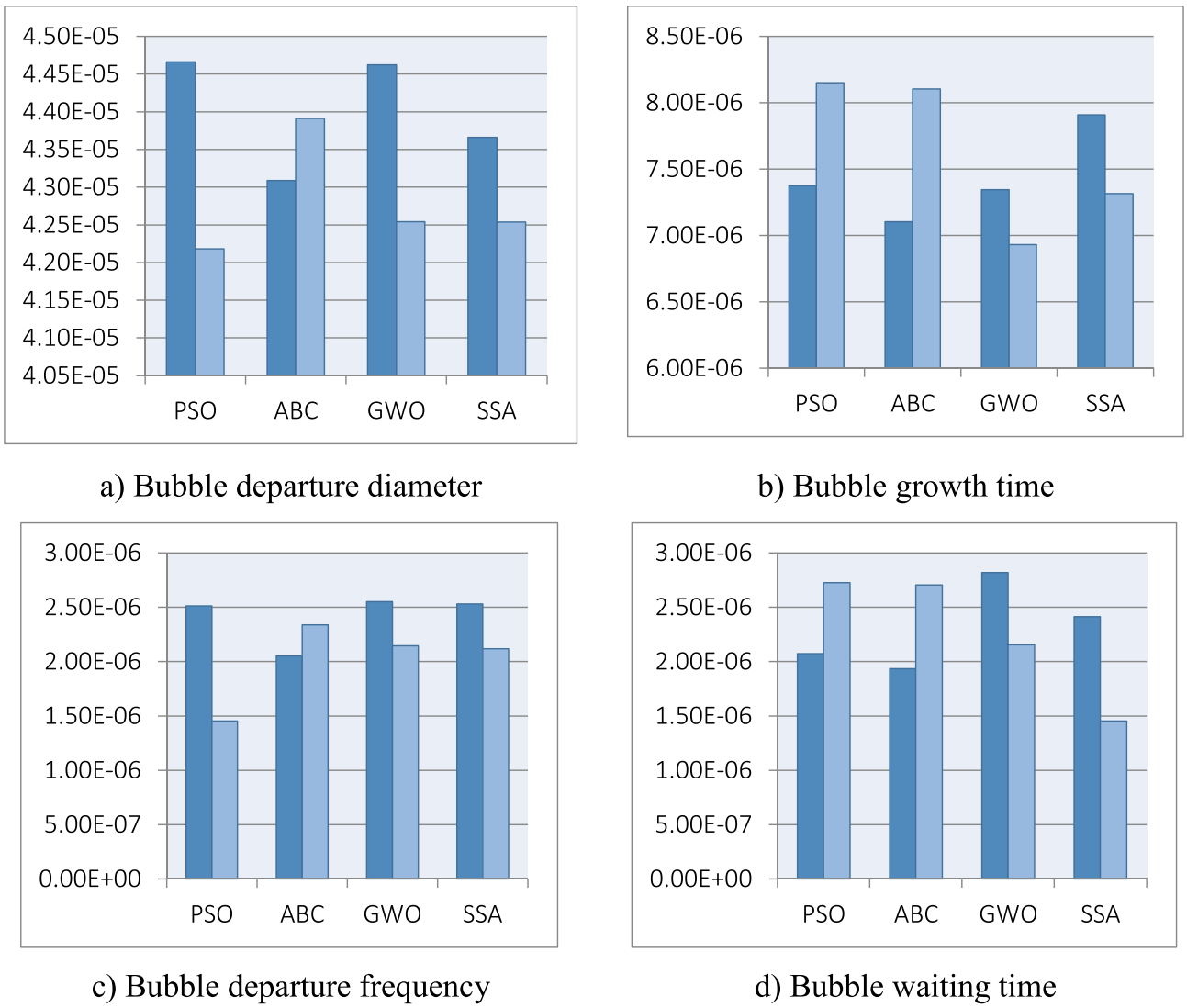


Fig. 29 Performance comparison of four bubble dynamics parameters using the 2HL and 3HL networks (2HL 3HL)

Table 12 Comparison of the best results for the bubble departure diameter output

	N_i (number of neurons)	F_i (transfer function)	MSE	Improvement compared to 1HL network (%)
1HL network	[29]	[logsig]	6.5166e-05	––
2HL network	[27]	[tansig, tansig]	4.3087e-05	33.88
3HL network	[9, 18, 27]	[elliotsig, logsig, logsig]	4.2180e-05	35.27

Table 13 Comparison of the best results for the bubble growth time output

	N_i (number of neurons)	F_i (transfer function)	MSE	Improvement compared to 1HL network (%)
1HL network	[24]	[logsig]	12.8020e-06	––
2HL network	[25, 29]	[tansig, logsig]	7.1041e-06	44.51
3HL network	[22, 25, 30]	[logsig, elliotsig, tansig]	6.9321e-06	45.85

Table 14 Comparison of the best results for the bubble departure frequency output

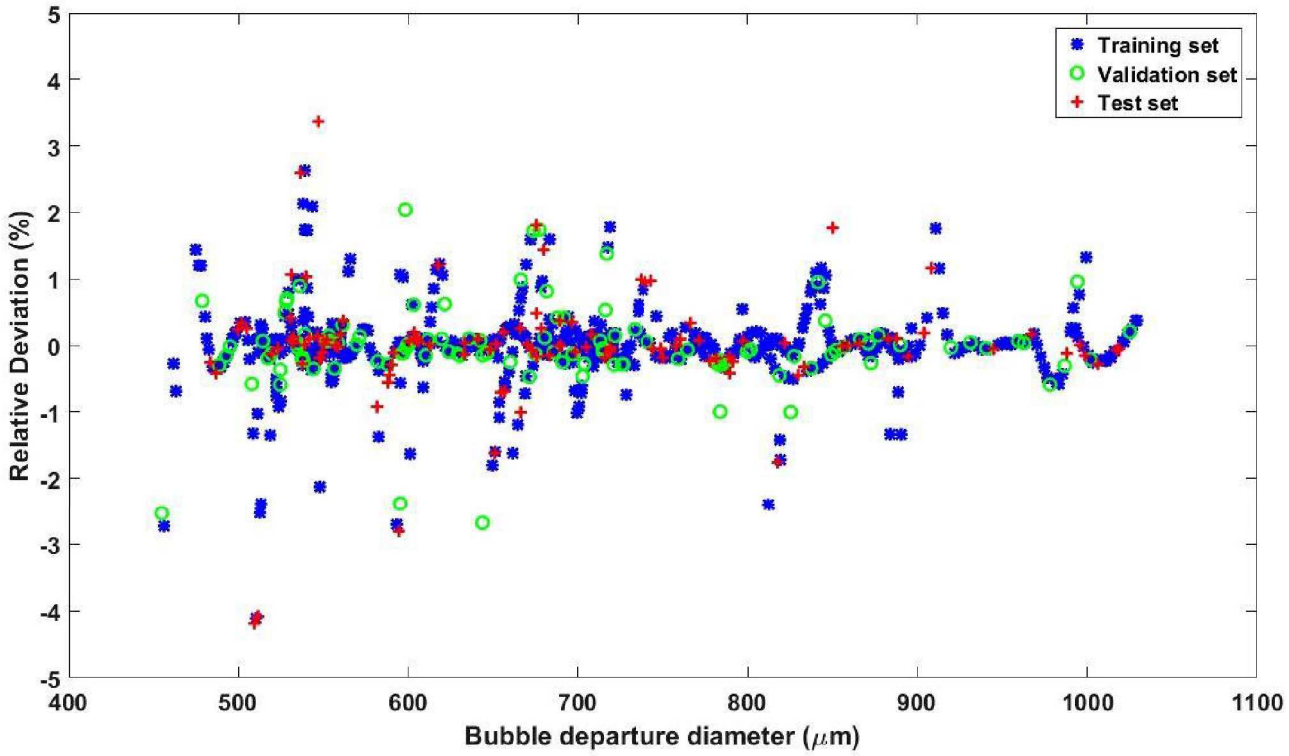
	N_i (number of neurons)	F_i (transfer function)	MSE	Improvement compared to 1HL network (%)
1HL network	[27]	[tansig]	3.8610e-06	––
2HL network	[16, 22]	[elliotsig, tansig]	2.0521e-06	46.85
3HL network	[9, 18, 22]	[elliotsig, tansig, elliotsig]	1.4550e-06	62.32

are normally distributed and considerably low in all bubble dynamics parameters. In other words, the error frequency accumulates mainly in the zero-axis error range that creates a symmetric graph and illustrates which the models can estimate the behavior of the bubble dynamics parameters excellent. Note that the "zero error" line divides positive and negative values. The direction of the bias can be identified by the sign of the error. Positive error indicates the outputs are smaller than the targets and negative error indicates that the targets are larger than the outputs.

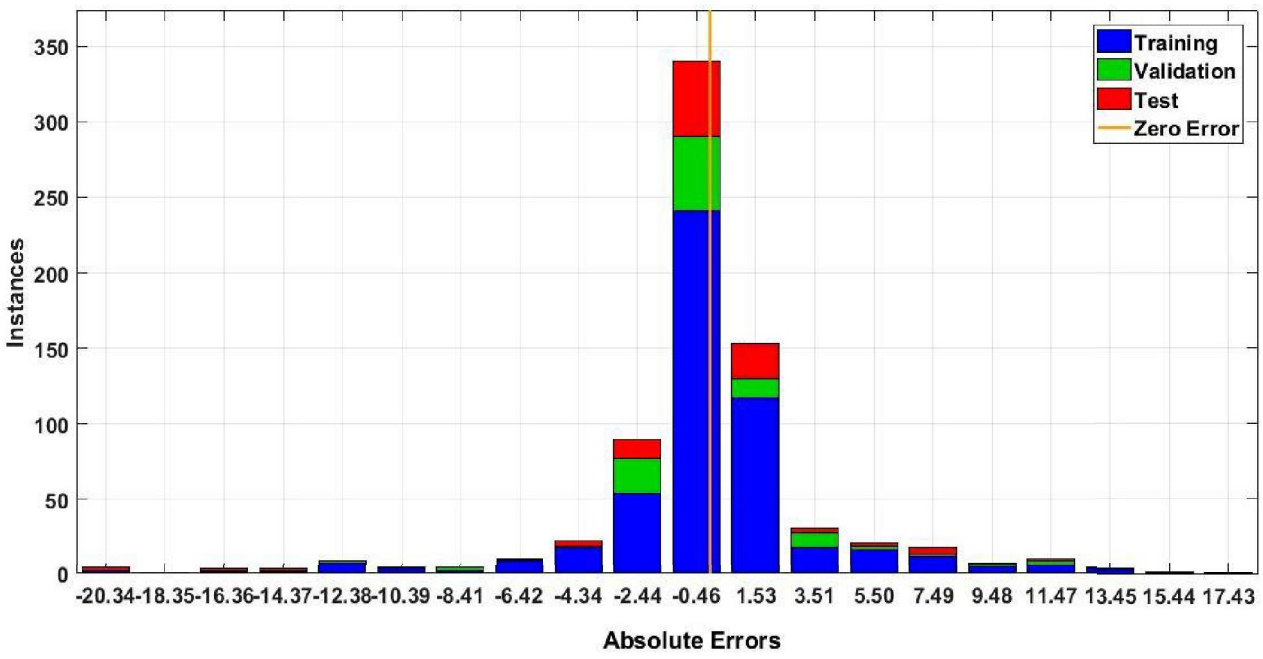
By increasing the number of hidden layers more than three, the network error may decrease further. However, augmenting the hidden layers and the neurons' number in each hidden layer could lead to a higher number of connections and thus increases the computational costs. It happens when the number of hidden layers and the number of neurons in each layer are increased simultaneously. In Figs. 34, 35, 36 and 37, the effects of the number of connections on the network's MSE are analyzed. It should be noted that the number of connections can be independent of the number of hidden layers in some cases. For instance, the 2HL network has greater connections than the 3HL network because it relies upon the neurons' number in each hidden layer. Figures 34, 35, and 37 show that, by increasing the number of connections in the network, its MSE is decreased for bubble departure diameter, bubble growth time, and bubble waiting time. However, Fig. 36 shows that this behavior is reversed for bubble departure frequency, that suggests that in a greater number of connections, MSE could be increased.

Table 15 Comparison of the best results for the bubble waiting time output

	N_i (number of neurons)	F_i (transfer function)	MSE	Improvement compared to 1HL network (%)
1HL network	[27]	[logsig]	4.0665e-06	––
2HL network	[23, 28]	[elliotsig, logsig]	1.9339e-06	52.44
3HL network	[19, 29, 30]	[logsig, logsig, tansig]	1.4528e-06	64.27

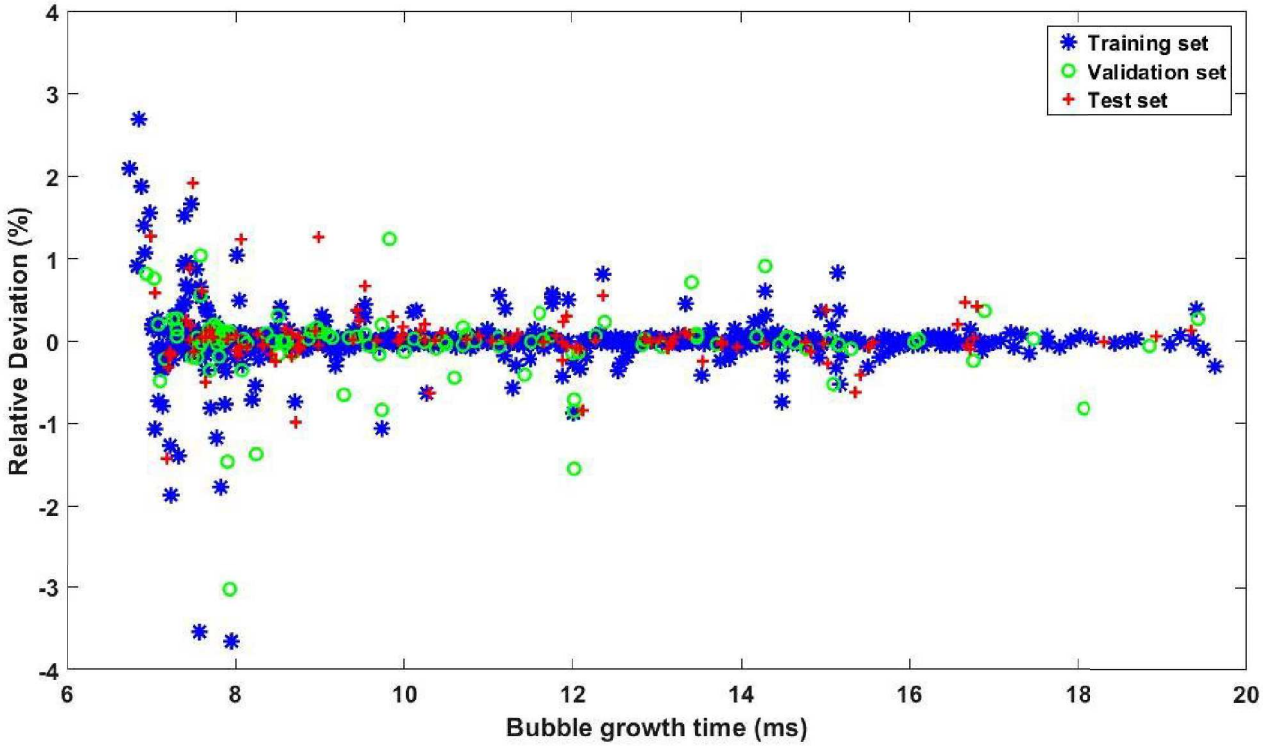


a) Relative deviation

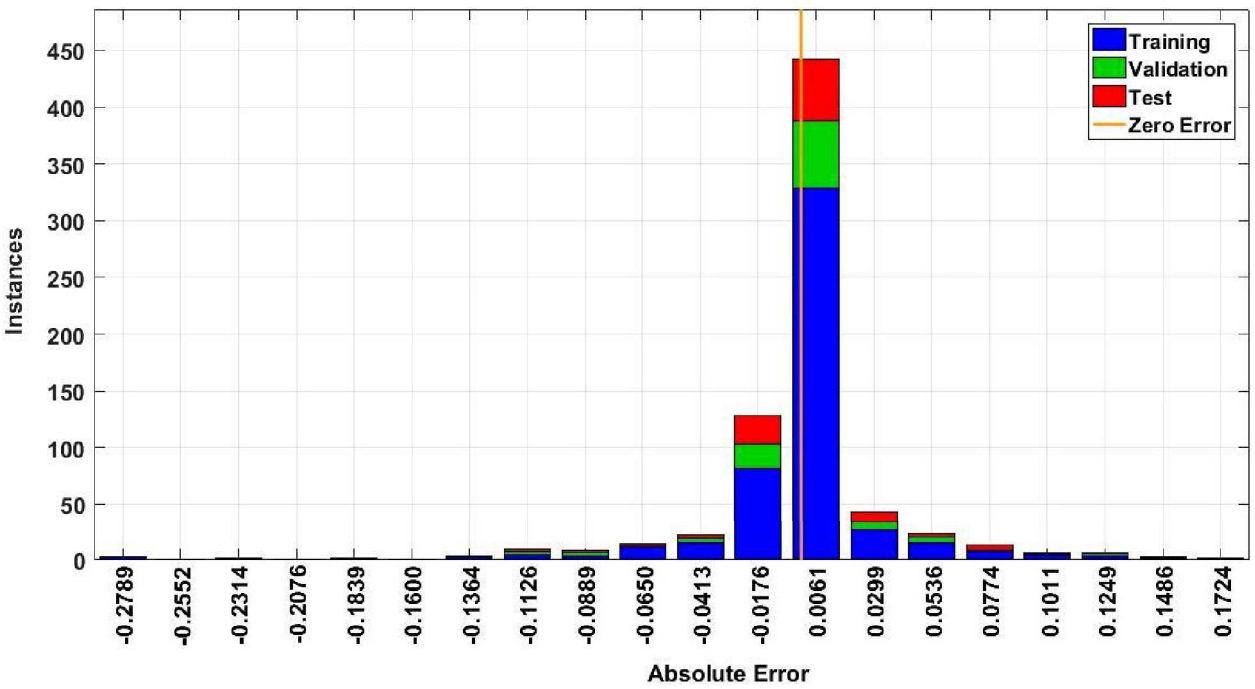


b) Error histogram

Fig. 30 Error analysis for bubble departure diameter with the best 3HL network

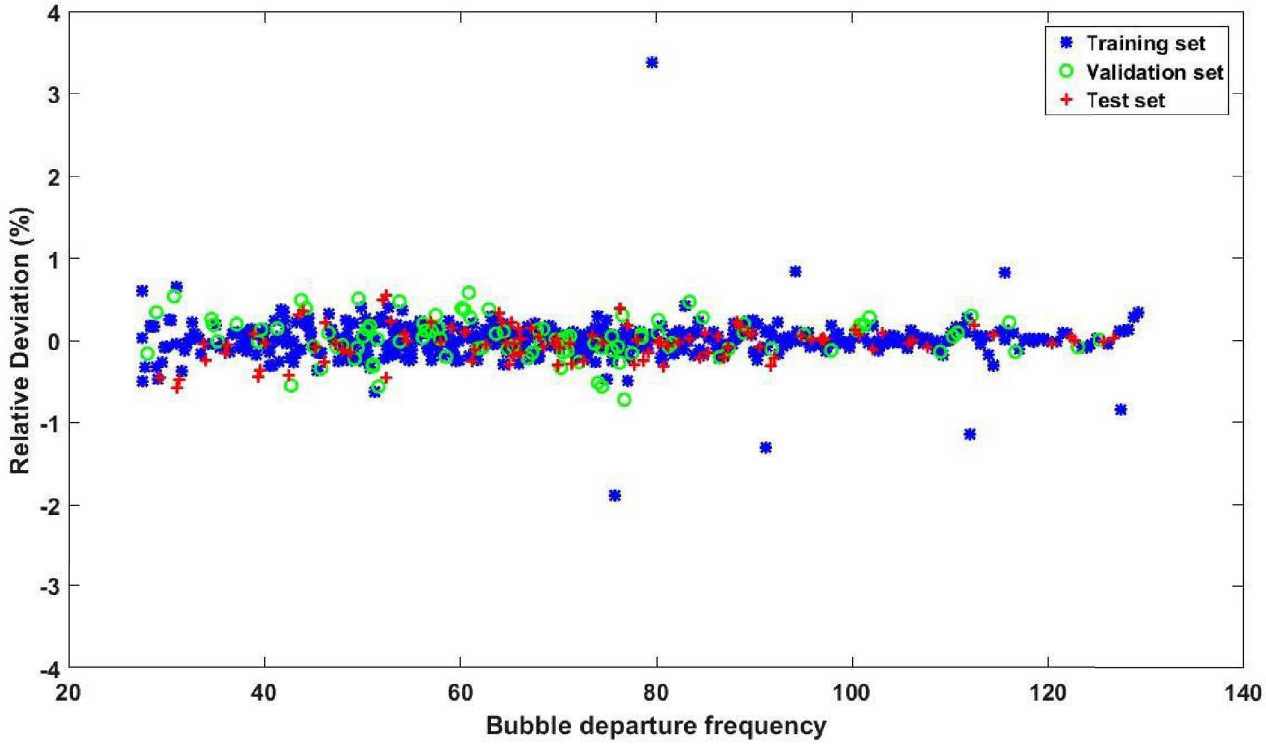


a) Relative deviation

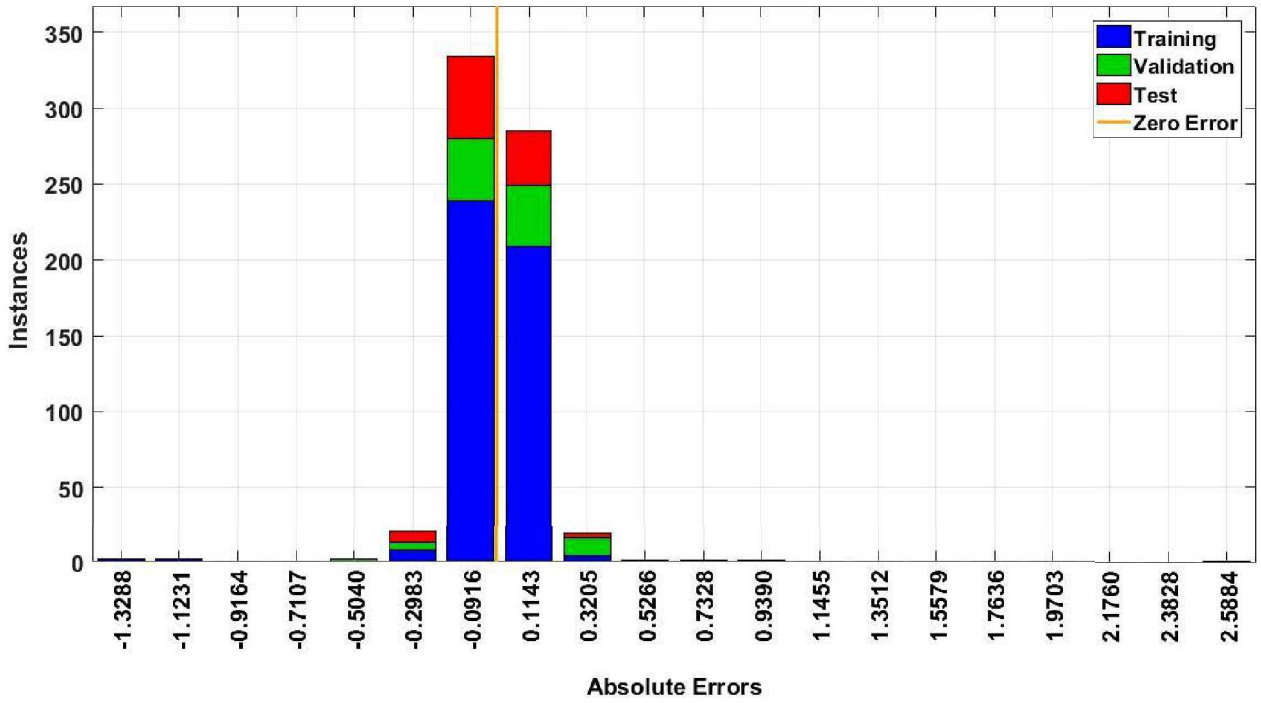


b) Error histogram

Fig. 31 Error analysis for bubble growth time with the best 3HL network



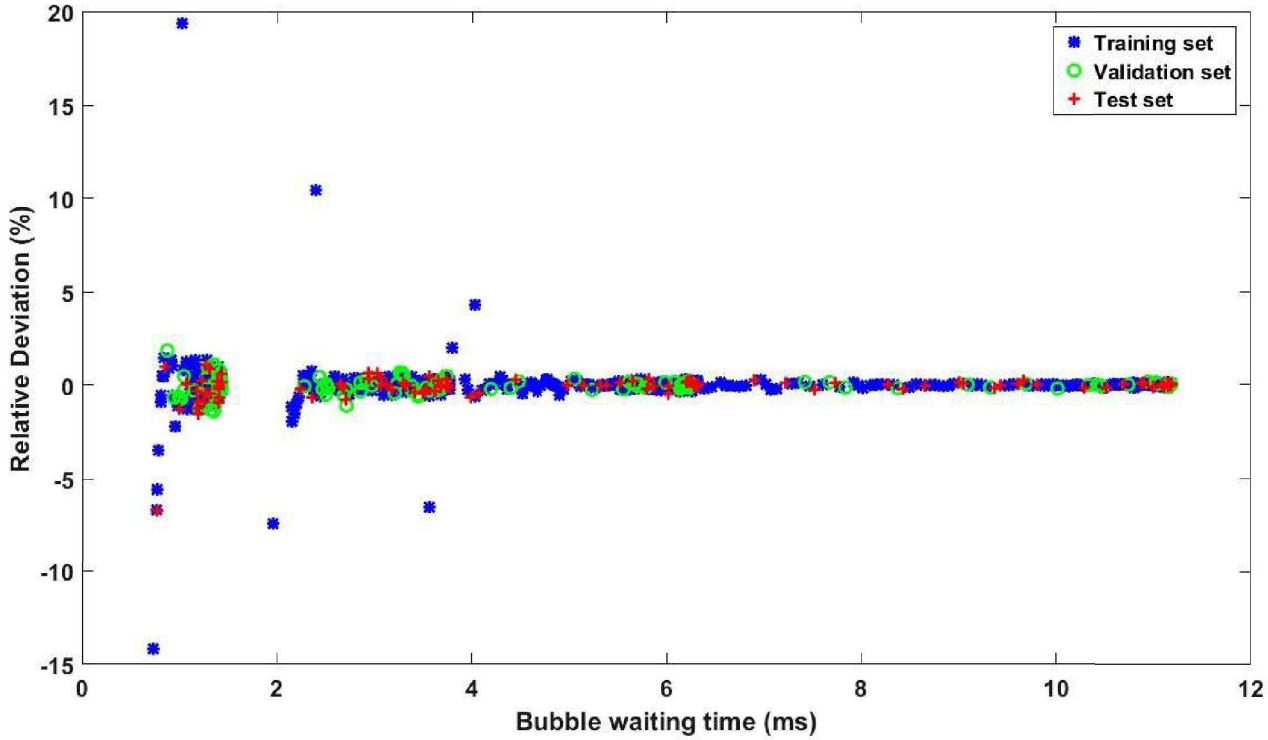
a) Relative deviation



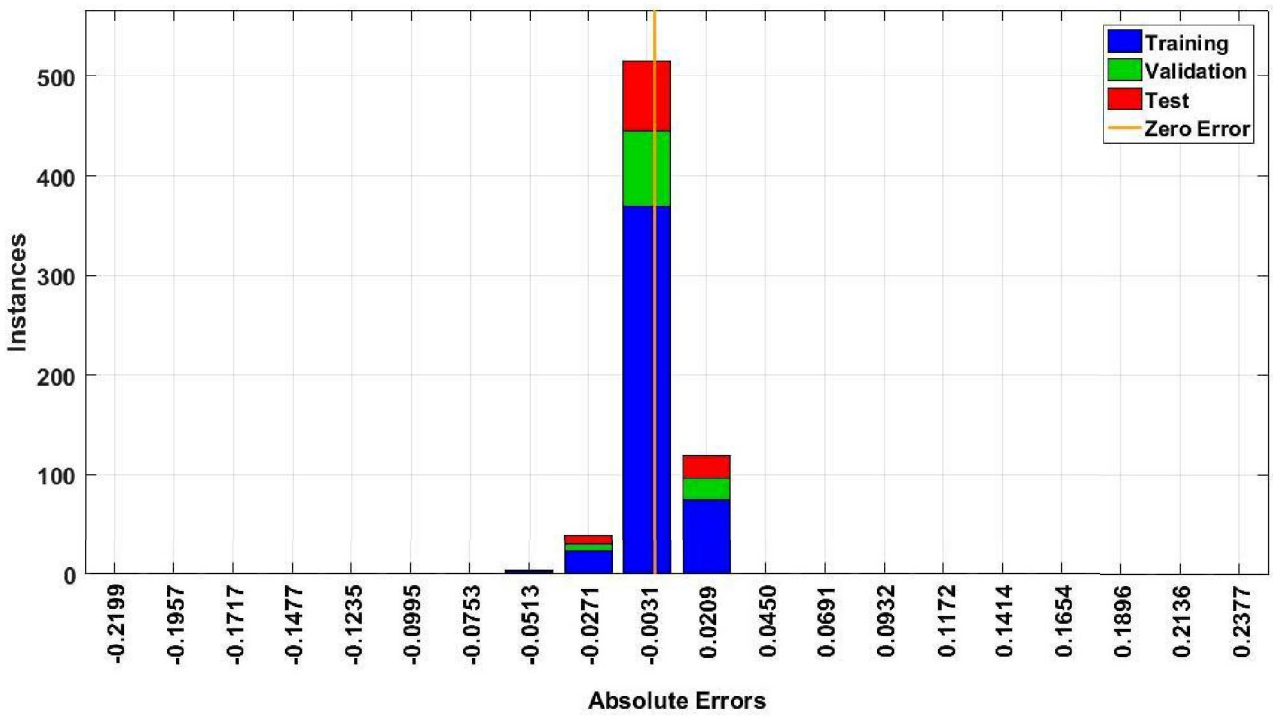
b) Error histogram

etwork.

Fig. 32 Error analysis for bubble departure frequency with the best 3HL network



a) Relative deviation



b) Error histogram

Fig. 33 Error analysis for bubble waiting time with the best 3HL network

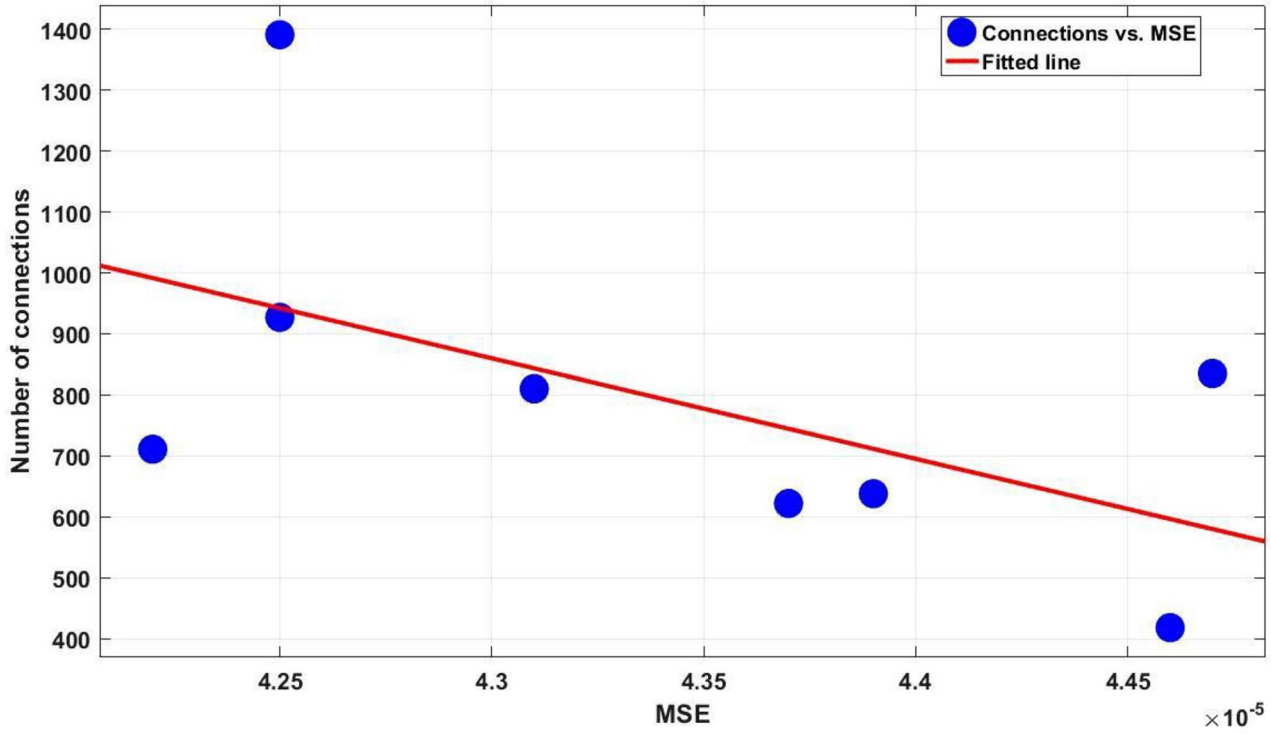


Fig. 34 Number of connections versus Network's MSE for bubble departure diameter with the best 2HL and 3HL networks

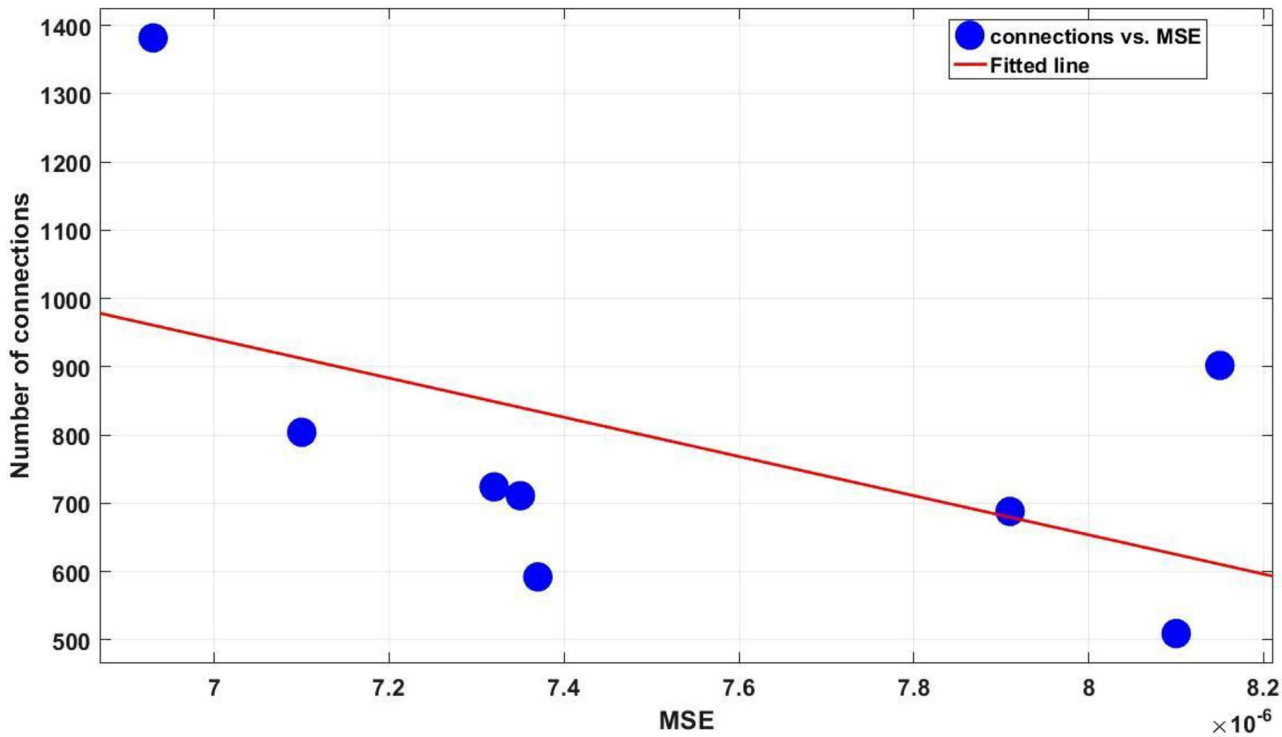


Fig. 35 Number of connections versus network's MSE for bubble growth time with the best 2HL and 3HL networks

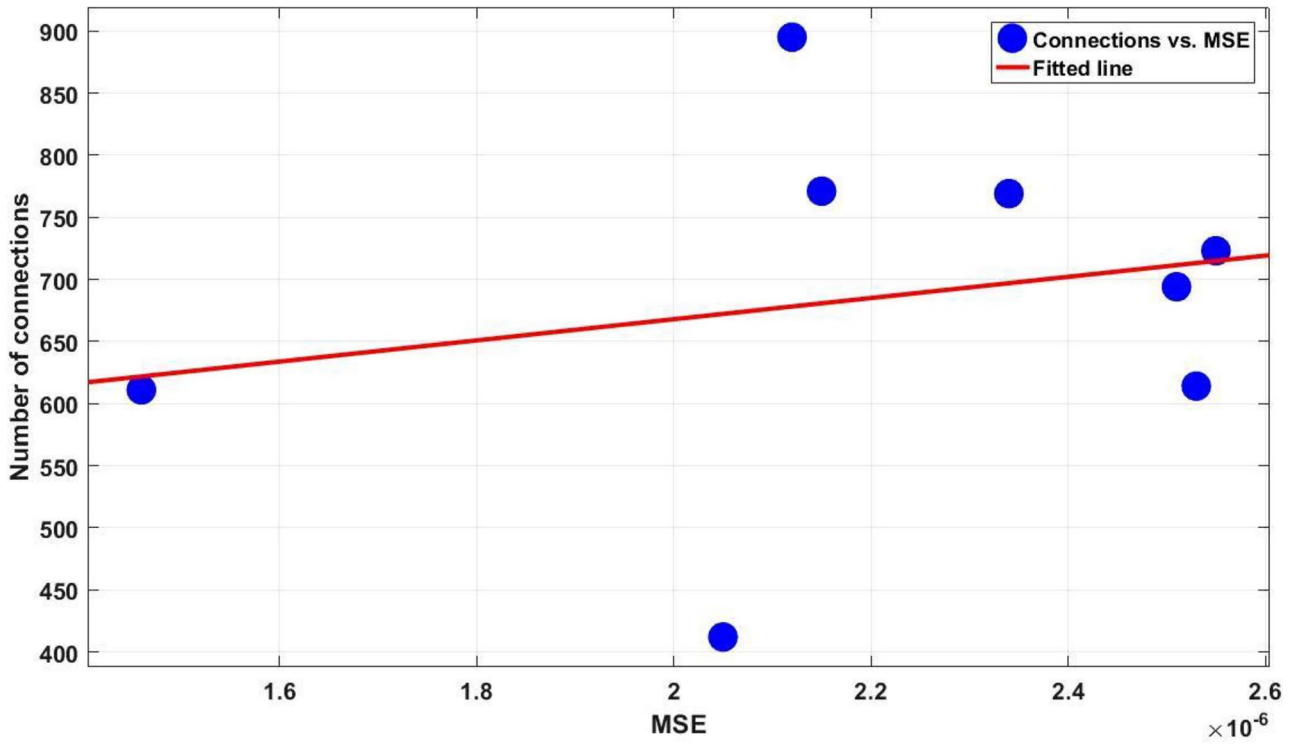


Fig. 36 Number of connections versus network's MSE for bubble departure frequency with the best 2HL and 3HL networks

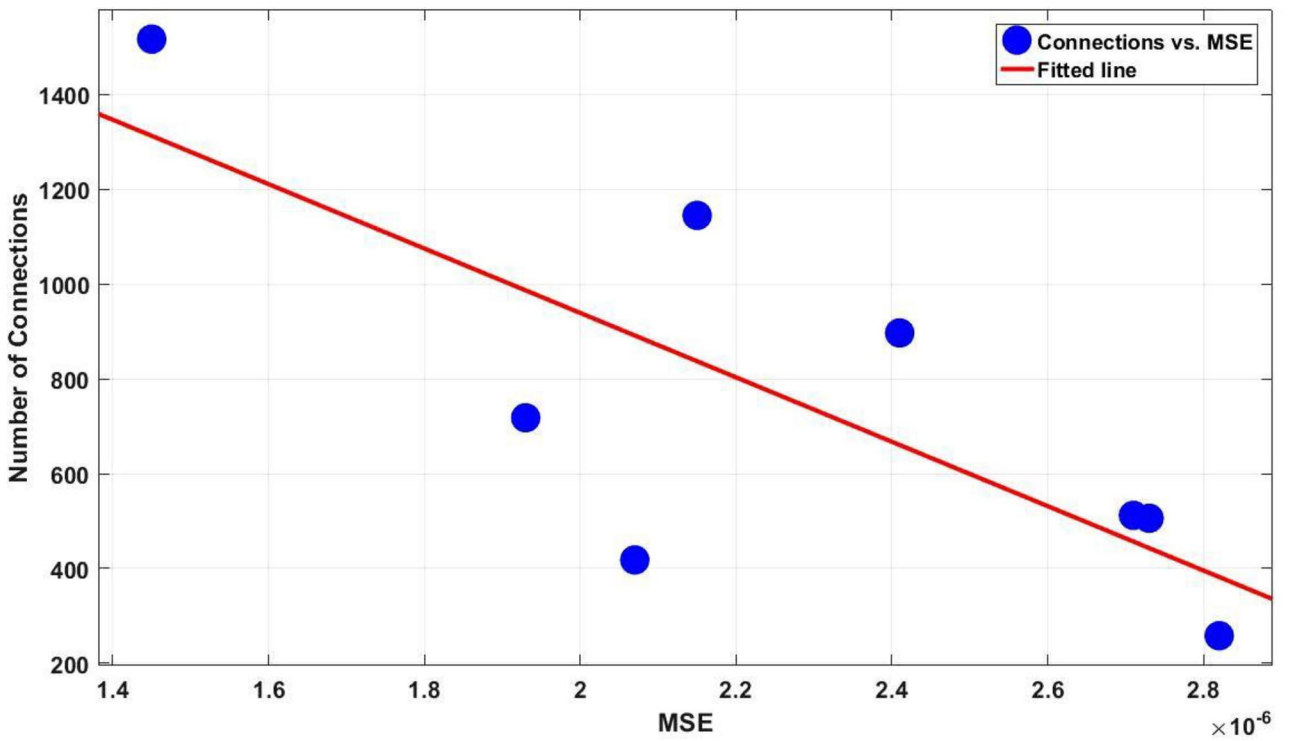


Fig. 37 Number of connections versus network's MSE for bubble waiting time with the best 2HL and 3HL networks

6 Conclusion

In this study, a MLP neural network with a BP training algorithm is used to model the heat transfer enhancement of nucleate boiling with a non-uniform electric field. The influences of the electric field on different parameters describing bubble's behaviors including bubble waiting time, bubble departure frequency, bubble growth time, bubble departure diameter are considered. Specifically, the objective of the investigation is to model a single bubble dynamic behavior of R113 created on a heater in an inconsistent electric field with the MLP-NN optimized by four different swarm-based optimization algorithms, namely: Partial Swarm Optimization (PSO), Artificial Bee Colony (ABC) algorithm, Grey Wolf Optimizer (GWO) and Salp Swarm Algorithm (SSA). For evaluating the accuracy of various models, the MSE value of the ANN model with different optimization algorithms is measured and compared. Also, the optimal network for each parameter is defined. The results of MLP network optimization with four algorithms (PSO, ABC, GWO, and SSA) can be summarized as follows:

- In optimization of the network with 2HL, the ABC algorithm has the least error.
- For network optimization in 3HL model, the PSO algorithm has the least error for the bubble departure diameter and bubble departure frequency and the SSA and GWO algorithms have the best performance for the bubble waiting time and bubble growth time, respectively.
- Comparing the results of the ABC algorithm, the error of the optimal networks in the 2HL model is always less than that of the optimal networks in the 3HL model. This suggests that the ABC algorithm performs better in a smaller number of design variables.
- In the results of the SSA and GWO algorithms, the error of optimal networks in the 3HL model is always less than that of optimal networks in the 2HL model. This suggests that the SSA and GWO algorithms perform better in larger design variables.
- The SSA and GWO algorithms compete properly with the PSO and ABC algorithms and obtain remarkable outcomes with low MSE values. Considering the antiquity and power of PSO and ABC algorithms, the capabilities of these fledgling algorithms are remarkable.
- The number of neurons in the middle layers in optimal networks shows that the number of neurons in the initial middle layer is higher than the neurons' number in the final middle layer.
- The relative deviations of the optimized networks for the bubble departure diameter, the bubble growth time, the bubble departure frequency, and the bubble waiting time are in the range of $\pm 4\%$, approximately.
- For bubble departure diameter, bubble growth time, and bubble waiting time, by increasing the number of connections in the network, its MSE is decreased.

Acknowledgements This work has been financially supported by the National Science Foundation (Grant #: 1917272).

Funding National Science Foundation, 1917272, Myeongsub Kim.

Data availability The data used in this study is available upon request.

Declarations

Conflict of interest The authors have no conflicts of interest to declare.

References

1. Bi J, Christopher DM, Zhao D, Xu J, Huang Y (2019) Numerical study of bubble growth and merger characteristics during nucleate boiling. Prog Nucl Energy 112:7–19. https://doi.org/10.1016/j.pnucene.2018.12.001
2. Kangude P, Srivastava A (2020) Understanding the growth mechanism of single vapor bubble on a hydrophobic surface: Experiments under nucleate pool boiling regime. Int J Heat Mass Transf 154:119775. https://doi.org/10.1016/j.ijheatmasstransfer.2020.119775
3. Tran N, Sajjad U, Lin R, Wang CC (2020) Effects of surface inclination and type of surface roughness on the nucleate boiling heat transfer performance of HFE-7200 dielectric fluid. Int J Heat Mass Transf 147:119015. https://doi.org/10.1016/j.ijheatmasstransfer.2019.119015
4. McHale JP, Garimella SV (2010) Bubble nucleation characteristics in pool boiling of a wetting liquid on smooth and rough surfaces. Int J Multiph Flow 36:249–260. https://doi.org/10.1016/j.ijmultiphaseflow.2009.12.004
5. Lamas MI, Jabardo JMS, Arce A, Fariñas P (2012) Numerical analysis of the bubble detachment diameter in nucleate boiling. J Phys Conf Ser. Institute of Physics Publishing. https://doi.org/10.1088/1742-6596/395/1/012174
6. Bovard S, Asadinia H, Hosseini G, Alavi Fazel SA (2017) Investigation and experimental analysis of the bubble departure diameter in pure liquids on horizontal cylindrical heater. Heat Mass Transf Und Stoffuebertragung 53:1199–1210. https://doi.org/10.1007/s00231-016-1885-3
7. Hamzekhani S, Maniavi Falahieh M, Akbari A (2014) Bubble departure diameter in nucleate pool boiling at saturation: Pure liquids and binary mixtures. Int J Refrig 46:50–58. https://doi.org/10.1016/j.ijrefrig.2014.07.003
8. Hazi G, Markus A (2009) On the bubble departure diameter and release frequency based on numerical simulation results. Int J Heat Mass Transf 52:1472–1480. https://doi.org/10.1016/j.ijheatmasstransfer.2008.09.003
9. Zeng LZ, Klausner JF, Mei R (1993) A unified model for the prediction of bubble detachment diameters in boiling systems- I. Pool boiling. Int J Heat Mass Transf 36:2261–2270. https://doi.org/10.1016/S0017-9310(05)80111-5
10. Lee HC, Do Oh B, Bae SW, Kim MH (2003) Single bubble growth in saturated pool boiling on a constant wall temperature surface. Int J Multiph Flow 29:1857–1874. https://doi.org/10.1016/j.ijmultiphaseflow.2003.09.003

11. Sakashita H (2011) Bubble growth rates and nucleation site densities in saturated pool boiling of water at high pressures. J Nucl Sci Technol 48:734–743. https://doi.org/10.1080/18811248.2011.9711756
12. Suszko A, El-Genk MS (2015) Saturation boiling of PF-5060 on rough Cu surfaces: Bubbles transient growth, departure diameter and detachment frequency. Int J Heat Mass Transf 91:363–373. https://doi.org/10.1016/j.ijheatmasstransfer.2015.07.083
13. Van Stralen SJD, Sluyter WM (1969) Local temperature fluctuations in saturated pool boiling of pure liquids and binary mixtures. Int J Heat Mass Transf 12. https://doi.org/10.1016/0017-9310(69)90061-1
14. Hutter C, Kenning DBR, Sefiane K, Karayiannis TG, Lin H, Cummins G, Walton AJ (2010) Experimental pool boiling investigations of FC-72 on silicon with artificial cavities and integrated temperature microsensors. Exp Therm Fluid Sci 34:422–433. https://doi.org/10.1016/j.expthermflusci.2009.03.010
15. Gorenflo D, Knabe V, Bieling V (1986) Bubble density on surfaces with nucleate boiling - its influence on heat transfer and burnout heat flux at elevated saturation pressures. Heat Transf Proc Int Heat Transf Conf Hemisphere Publ Corp, pp 1995–2000. https://doi.org/10.1615/ihtc8.3920
16. Hamzekhani S, Falahieh MM, Kamalizadeh MR, Nazari Z (2015) Experimental study on bubble departure frequency for pool boiling of water/NaCl solutions. Heat Mass Transf Und Stoffuebertragung 51:1313–1320. https://doi.org/10.1007/s00231-015-1502-x
17. Yang C, Wu Y, Yuan X, Ma C (2000) Study on bubble dynamics for pool nucleate boiling. Int J Heat Mass Transf 43:203–208. https://doi.org/10.1016/s0017-9310(99)00132-5
18. Li X, Tang J, Sun L, Li J, Bao J, Liu H (2020) Enhancement of subcooled boiling in confined space using ultrasonic waves. Chem Eng Sci 223:115751. https://doi.org/10.1016/j.ces.2020.115751
19. Zhao Z, Ma X, Li S, Yang S, Huang L (2020) Visualization-based nucleate pool boiling heat transfer enhancement on different sizes of square micropillar array surfaces. Exp Therm Fluid Sci 119:110212. https://doi.org/10.1016/j.expthermflusci.2020.110212
20. Wong KK, Leong KC (2019) Nucleate flow boiling enhancement on engineered three-dimensional porous metallic structures in FC-72. Appl Therm Eng 159:113846. https://doi.org/10.1016/j.applthermaleng.2019.113846
21. Lay KK, Ong JS, Yong KY, Tan MK, Hung YM (2019) Nucleate pool boiling enhancement by ultrafast water permeation in graphene-nanostructure. Int Commun Heat Mass Transf 101:26–34. https://doi.org/10.1016/j.icheatmasstransfer.2018.12.015
22. Karayiannis TG, Xu Y (1998) Electric field effect in boiling heat transfer. Part B: Electrode geometry. J Enhanc Heat Transf 5:231–247. https://doi.org/10.1615/jenhheattransf.v5.i4.20
23. Kweon YC, Kim MH (2000) Experimental study on nucleate boiling enhancement and bubble dynamic behavior in saturated pool boiling using a nonuniform de electric field. Int J Multiph Flow 26:1351–1368. https://doi.org/10.1016/S0301-9322(99)00090-7
24. Sefiane K (2001) A new approach in the modeling of the critical heat flux and enhancement techniques. AIChE J 47:2402–2412. https://doi.org/10.1002/aic.690471104
25. Chen F, Peng Y, Song YZ, Chen M (2007) EHD behavior of nitrogen bubbles in DC electric fields. Exp Therm Fluid Sci 32:174–181. https://doi.org/10.1016/j.expthermflusci.2007.03.006
26. Dong W, Li RY, Yu HL, Yan YY (2006) An investigation of behaviours of a single bubble in a uniform electric field. Exp Therm Fluid Sci 30:579–586. https://doi.org/10.1016/j.expthermflusci.2005.12.003
27. Ogata J, Yabe A (1993) Augmentation of boiling heat transfer by utilizing the EHD effect-EHD behaviour of boiling bubbles and heat transfer characteristics. Int J Heat Mass Transf 36:783–791. https://doi.org/10.1016/0017-9310(93)80054-X
28. Xu Y, Karayiannis TG (1998) Electric field effect in boiling heat transfer. Part B: electrode geometry. J Enhanc Heat Transf 5:231–247. https://doi.org/10.1615/jenhheattransf.v5.i4.20
29. Ahmad SW, Karayiannis TG, Kenning DBR, Luke A (2011) Compound effect of EHD and surface roughness in pool boiling and CHF with R-123. Appl Therm Eng 31:1994–2003. https://doi.org/10.1016/j.applthermaleng.2011.03.005
30. Zu YQ, Yan YY (2009) A numerical investigation of electrohydrodynamic (EHD) effects on bubble deformation under pseudo-nucleate boiling conditions. Int J Heat Fluid Flow 30:761–767. https://doi.org/10.1016/j.ijheatfluidflow.2009.03.008
31. Pascual CC, Jeter SM, Abdel-Khalik SI (2001) A statistical analysis of EHD-enhanced nucleate boiling along a heated wire. Int J Heat Mass Transf 44:1201–1212. https://doi.org/10.1016/S0017-9310(00)00149-6
32. Madadnia J, Koosha H (2003) Electrohydrodynamic effects on characteristic of isolated bubbles in the nucleate pool boiling regime. Exp Therm Fluid Sci 27:145–150. https://doi.org/10.1016/S0894-1777(02)00258-3
33. Siedel S, Cioulachtjian S, Robinson AJ, Bonjour J (2011) Electric field effects during nucleate boiling from an artificial nucleation site. Exp Therm Fluid Sci 35:762–771. https://doi.org/10.1016/j.expthermflusci.2010.06.006
34. Chen F, Liu D, Song Y (2012) Visualization of a single boiling bubble in a DC electric field. Am Soc Mech Eng Fluids Eng Div FEDSM. American Society of Mechanical Engineers Digital Collection, pp 245–252. https://doi.org/10.1115/FEDSM2012-72493
35. Kim J (2009) Review of nucleate pool boiling bubble heat transfer mechanisms. Int J Multiph Flow 35:1067–1076. https://doi.org/10.1016/j.ijmultiphaseflow.2009.07.008
36. Warrier GR, Dhir VK (2006) Heat transfer and wall heat flux partitioning during subcooled flow nucleate boiling - A review. J Heat Transfer 128:1243–1256. https://doi.org/10.1115/1.2349510
37. Richenderfer A, Kossolapov A, Seong JH, Saccone G, Bucci M, Buongiorno J (2017) Direct measurement of heat flux partitioning in boiling heat transfer (ASME). Am Soc Mech Eng Fluids Eng Div FEDSM. American Society of Mechanical Engineers. https://doi.org/10.1115/FEDSM2017-69347
38. Scalabrin G, Condosta M, Marchi P (2006) Modeling flow boiling heat transfer of pure fluids through artificial neural networks. Int J Therm Sci 45:643–663. https://doi.org/10.1016/J.IJTHERMALSCI.2005.09.009
39. Scalabrin G, Condosta M, Marchi P (2006) Flow boiling of pure fluids: local heat transfer and flow pattern modeling through artificial neural networks. Int J Therm Sci 45:739–751. https://doi.org/10.1016/J.IJTHERMALSCI.2005.09.010
40. Wen T, Zhu G, Lu L (2021) Experimental and artificial neural network based study on the heat transfer and flow performance of ZnO-EG/water nanofluid in a mini-channel with serrated fins. Int J Therm Sci 170:107149. https://doi.org/10.1016/J.IJTHERMALSCI.2021.107149
41. Cho E, Lee H, Kang M, Jung D, Lee G, Lee S, Kharangate CR, Ha H, Huh S, Lee H (2022) A neural network model for free-falling condensation heat transfer in the presence of non-condensable gases. Int J Therm Sci 171:107202. https://doi.org/10.1016/J.IJTHERMALSCI.2021.107202
42. Scalabrin G, Condosta M, Marchi P (2006) Mixtures flow boiling: modeling heat transfer through artificial neural networks. Int J Therm Sci 45:664–680. https://doi.org/10.1016/J.IJTHERMALSCI.2005.09.011
43. Peng Y, Ghahnaviyeh MB, Ahamd MN, Abdollahi A, Bagherzadeh SA, Azimy H, Mosavi A, Karimipour A (2021) Analysis of the effect of roughness and concentration of Fe3O4/water nanofluid on the boiling heat transfer using the artificial neural network: An experimental and numerical study. Int J Therm Sci 163:106863. https://doi.org/10.1016/J.IJTHERMALSCI.2021.106863

44. Diaz G, Campo A (2009) Artificial neural networks to correlate in-tube turbulent forced convection of binary gas mixtures. Int J Therm Sci 48:1392–1397. https://doi.org/10.1016/J.IJTHERMALSCI.2008.12.001
45. Mohanraj M, Jayaraj S, Muraleedharan C (2015) Applications of artificial neural networks for thermal analysis of heat exchangers – A review. Int J Therm Sci 90:150–172. https://doi.org/10.1016/J.IJTHERMALSCI.2014.11.030
46. Nafey AS (2009) Neural network based correlation for critical heat flux in steam-water flows in pipes. Int J Therm Sci 48:2264–2270. https://doi.org/10.1016/J.IJTHERMALSCI.2009.04.010
47. Kurt H, Atik K, Özkaymak M, Recebli Z (2008) Thermal performance parameters estimation of hot box type solar cooker by using artificial neural network. Int J Therm Sci 47:192–200. https://doi.org/10.1016/J.IJTHERMALSCI.2007.02.007
48. Sayahi T, Tatar A, Bahrami M (2016) A RBF model for predicting the pool boiling behavior of nanofluids over a horizontal rod heater. Int J Therm Sci 99:180–194. https://doi.org/10.1016/J.IJTHERMALSCI.2015.08.010
49. Wei H, Su GH, Tian WX, Qiu SZ, Yang XB (2010) Study on the onset of nucleate boiling in narrow annular channel by genetic neural network. Int Commun Heat Mass Transf 37:596–599. https://doi.org/10.1016/j.icheatmasstransfer.2009.11.017
50. Zendehboudi A, Tatar A (2017) Utilization of the RBF network to model the nucleate pool boiling heat transfer properties of refrigerant-oil mixtures with nanoparticles. J Mol Liq 247:304–312. https://doi.org/10.1016/j.molliq.2017.09.105
51. Balcilar M, Dalkilic AS, Suriyawong A, Yiamsawas T, Wongwises S (2012) Investigation of pool boiling of nanofluids using artificial neural networks and correlation development techniques. Int Commun Heat Mass Transf 39:424–431. https://doi.org/10.1016/j.icheatmasstransfer.2012.01.008
52. Pare A, Ghosh SK (2021) Surface qualitative analysis and ANN modelling for pool boiling heat transfer using Al2O3-water based nanofluids. Colloids Surfaces A Physicochem Eng Asp 610:125926. https://doi.org/10.1016/j.colsurfa.2020.125926
53. Liang X, Xie Y, Day R, Meng X, Wu H (2021) A data driven deep neural network model for predicting boiling heat transfer in helical coils under high gravity. Int J Heat Mass Transf 166:120743. https://doi.org/10.1016/j.ijheatmasstransfer.2020.120743
54. Dong ZG, Cheng DL, Li RY, Yu HL, Dong W (2011) ICMREE2011 - Proc 2011 Experimental study of influence on bubble in electric field. Int Conf Mater Renew Energy Environ. pp 1280–1283. https://doi.org/10.1109/ICMREE.2011.5930570
55. Gao M, Cheng P, Quan X (2013) An experimental investigation on effects of an electric field on bubble growth on a small heater in pool boiling. Int J Heat Mass Transf 67:984–991. https://doi.org/10.1016/j.ijheatmasstransfer.2013.08.098
56. Ghazvini M, Maddah H, Peymanfar R, Ahmadi MH, Kumar R (2020) Experimental evaluation and artificial neural network modeling of thermal conductivity of water based nanofluid containing magnetic copper nanoparticles. Phys A Stat Mech Its Appl 124127. https://doi.org/10.1016/j.physa.2019.124127
57. Ertuğrul ÖF (2018) A novel type of activation function in artificial neural networks: Trained activation function. Neural Netw 99:148–157. https://doi.org/10.1016/j.neunet.2018.01.007
58. Herzog S, Tetzlaff C, Wörgötter F (2020) Evolving artificial neural networks with feedback. Neural Netw 123:153–162. https://doi.org/10.1016/j.neunet.2019.12.004
59. Mirjalili S, Lewis A (2016) The whale optimization algorithm. Adv Eng Softw 95:51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008
60. Rao SS (2019) Engineering Optimization Theory and Practice. Wiley. https://doi.org/10.1002/9781119454816
61. Seixas Gomes de Almeida B, Coppo Leite V (2019) Particle Swarm optimization: a powerful technique for solving engineering problems. Swarm Intell - Recent Adv New Perspect Appl Intech Open. https://doi.org/10.5772/intechopen.89633
62. Sardashti A, Daniali HM, Varedi SM (2013) Optimal free-defect synthesis of four-bar linkage with joint clearance using PSO algorithm. Meccanica 48:1681–1693. https://doi.org/10.1007/s11012-013-9699-6
63. Praveen C, Duvigneau R (2009) Low cost PSO using metamodels and inexact pre-evaluation: Application to aerodynamic shape design. Comput Methods Appl Mech Eng 198:1087–1096. https://doi.org/10.1016/j.cma.2008.11.019
64. Karaboga D, Karaboga D (2005) An idea based on honey bee swarm for numerical optimization. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.714.4934. Accessed 28 July 2020
65. Karaboga D, Basturk B (2007) A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm 39:459–471. https://doi.org/10.1007/s10898-007-9149-x
66. Karaboga D, Basturk B (2007) Artificial Bee Colony (ABC) optimization algorithm for solving constrained optimization problems. Lect Notes Comput Sci (Including Subser Lect Notes Artif Intell Lect Notes Bioinformatics). Springer Verlag, pp 789–798. https://doi.org/10.1007/978-3-540-72950-1_77
67. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. Adv Eng Softw 69:46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007
68. Mech LD (1999) Alpha status, dominance, and division of labor in wolf packs. Can J Zool 77:1196–1203. https://doi.org/10.1139/z99-099
69. Abusnaina AA, Ahmad S, Jarrar R, Mafarja M (2018) Training neural networks using salp swarm algorithm for pattern classification. ACM Int Conf Proceeding Ser Association for Computing Machinery. New York, New York, USA, pp 1–6. https://doi.org/10.1145/3231053.3231070
70. Mirjalili S, Gandomi AH, Mirjalili SZ, Saremi S, Faris H, Mirjalili SM (2017) Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. Adv Eng Softw 114:163–191. https://doi.org/10.1016/j.advengsoft.2017.07.002
71. Sayed A, Sardeshmukh M, Limkar S (2014) Optimisation using Levenberg-Marquardt algorithm of neural networks for Iris. Adv Intell Syst Comput. Springer Verlag, pp 91–98. https://doi.org/10.1007/978-3-319-02931-3_12