



Artificial neural network based predictive negative hydrogen ion helicon plasma source for fusion grade large sized ion source

Vipin Shukla¹ · Mainak Bandyopadhyay² · Vivek Pandya¹ · A. Pandey² · A. Maulik¹

Received: 3 January 2020 / Accepted: 16 May 2020
© Springer-Verlag London Ltd., part of Springer Nature 2020

Abstract

Ion source operation and control is a challenging task, and the manual search of the optimized input parameter by the operator becomes a time consuming and cumbersome process. HELicon Experiment for Negative Ion Source (HELEN-I) with single driver developed at the Institute for Plasma Research (IPR) for the production negative hydrogen ion involves a multitude of interacting systems (such as Radio Frequency (RF) power, gas feed pressure, magnetic fields). This paper presents an Artificial Intelligence (AI) driven multiple-input and multiple-output model for HELEN-I to get optimized values of ion saturation current (mA) and high plasma density of order $10^{18}/\text{m}^3$. The developed Artificial Neural Network (ANN) based model predicts the outputs of the HELEN-I and is further utilizes a heuristic global optimization Particle Swarm Optimization (PSO) algorithm to select optimal source parameters to predict the desired outputs without conducting real-time experiments. The experimental studies further validate the results obtained by the proposed approach.

Keywords HELEN-I · Negative hydrogen ion · Helicon plasma · Artificial neural network (ANN) · Ion source · Particle swarm optimization (PSO)

1 Introduction

The Institute for Plasma Research has longstanding research and development program in the area of magnetically confined hot plasmas and plasma technologies for industrial applications. HELEN-I is a permanent ring magnet-based helicon plasma source developed at Institute Plasma Research (IPR) [1]. Helicon plasma sources have very high ionizing efficiency and produce very high plasma densities of the order of $\sim 10^{14}$ – 10^{18} per m^3 by utilizing low RF power (\sim few KW) [2]. The large sized fusion grade ion source is intending to use the module unit of HELEN-I [3]. The primary purpose of HELEN-I with single driver is to produce negative hydrogen ions. The HELEN-I module uses hydrogen and argon gases to produce negative ions. The

experimental setup of the module has many interactive systems, and the operator should adequately tune the ion source to achieve the optimized value of ion saturation current (mA) which in-turn produces high-density plasma. Ion source tuning is a highly complicated and time-consuming process. Operators should cautiously search the source (input) parameters manually by conducting real-time experiments to optimize the ion saturation current and plasma density.

In light of recent advances in the field of AI, the introduction of Machine Learning techniques is a promising avenue to provide a satisfactory solution. The scientific community, working in the area of fusion related research, have attempted for the automatic tuning and control of ion sources using AI techniques. While a few papers are available on the application of ANN for ion source modelling and control, Edelen et al. [4] discuss the neural-network-based control scheme for the resonance control of an RF electron gun used in a particle accelerator. Scheinker et al. [5] demonstrate a model-independent feedback control method combined with a trained neural network (NN) to achieve a relatively fast, automatic and precise electron beam phase space control. The approach is also applicable to an n -dimensional dynamic system. Huang et al. [6] presented an algorithm for online optimization of performance parameters of

✉ Vipin Shukla
Vipin.Shukla@pdpu.ac.in
Mainak Bandyopadhyay
mainak@ipr.res.in

¹ Pandit Deendayal Petroleum University, Gandhinagar, Gujarat, India

² Institute for Plasma Research, HBNI, Gujarat, Gandhinagar 382428, India

accelerator using *Robust Conjugate Direction Search (RCDS)* but did not utilize the neural network technique for modelling. Brown et al. [7] illustrated the development of an ANN-based automated controller for the optimization of operating point and control of a small angle ion source. Kong et al. [8] developed the ion source model for finding the optimized parameters for RFT-30 cyclotron using ANN. None except [8] utilizes both neural network technique and heuristic optimization technique for developing the model and optimizing the input parameters. Moreover, in all the papers [4–8] authors considered multiple inputs and single output systems and also ignored the function approximation of respective ion sources. The contributions of the previous papers are summarized in Table 1.

This paper uses neural network-based techniques for modelling the HELEN-I. First, the proposed approach builds a shallow ANN model and trains it using the experimental data obtained from the ion source operation. In this paper, various iterative optimization techniques are utilized for training the ANN model of ion source and tests and compares the performance of different algorithms. Based on the performance, the optimized ANN model is selected for further investigation. The proposed ANN-based ion source model predicts the outputs for the different input settings and validates it through the experimental outputs for the same input settings. The optimized ANN ion source model enables the operator to predict the optimized value of ion saturation current and plasma density for different input settings without manually operating the ion source. This proposed approach also allows the user to approximate the input–output function for the specific ion source, which is further utilized as an objective function to optimize the input parameters. The findings in the paper show that the model proposed in the paper can efficiently predict the outputs of ion source without manual operation and is also capable of optimizing the input source parameters.

2 The ion source HELEN-I

The schematic of HELEN-I showing the functional description and the photograph of the experimental setup is shown in Fig. 1. The HELicon Experiment for Negative ion source-I (HELEN-I) setup has a vacuum chamber which consists of a source chamber and an expansion chamber. The cylindrical source chamber is made of borosilicate glass having 70 mm length. The inner and outer diameters of the cylinder are 50 mm and 60 mm, respectively. The source chamber is placed on a cylindrical expansion chamber of length 300 mm, inner diameter 100 mm. The expansion chamber is made of stainless steel. The upper end of the cylindrical source chamber has a top conducting plate made up of stainless steel as one extreme vacuum boundary. A Nagoya-III antenna is wrapped from outside on the cylindrical source chamber wall to produce helicon plasma using hydrogen gas. A 13.56 MHz radio frequency (RF) power source with auto impedance matching unit is connected to the antenna to create plasma inside the glass chamber.

The diffusion pump is used to create the desired vacuum. Two Pirani gauges are installed along the pumping line, and a hot cathode penning gauge is connected to the expansion chamber for vacuum monitoring. There are 8 ports on the expansion chamber, these ports are meant for gas feed supply, vacuum gauges and diagnostic probes. HELEN-1 setup uses permanent ring magnets for creating a required axial magnetic field for helicon wave launching and permanent bar magnets for plasma confinement in the expansion chamber. Further details of the experimental setup can be found in references [1, 3].

The operator manipulates source operational parameters (e.g. RF power, pressure and magnetic field) to get enhanced plasma density (10^{18} m^{-3}) by monitoring Langmuir probe characteristics, particularly the Ion Saturation Current (ISC). Ion saturation current values lie in the order of few mA.

Table 1 Summary of literature review for ion source tuning and optimization

Reference no.	Type of device	Contributions	Limitation
Edelen et al. [4]	RF electron gun	Control of Resonance using neural networks and model predictive control	Multiple inputs and single output model Function approximation neglected
Scheinker et al. [5]	Particle Accelerator	A neural network is combined with model-independent feedback for automatic particle accelerator control and optimization. Total 6 parameters which suffer from time drift are tuned. The optimization is based on model-independent extremum.	Multiple inputs and single output model Function approximation neglected
Huang et al. [6]	Particle Accelerator	Online optimization of vertical emittance and injection optics	Application of neural network neglected
Brown et al. [7]	Small angle ion source	optimization of operating point and adaptive control scheme	Multiple inputs and single output model Function approximation neglected
Kong et al. [8]	Filament ion source	ANN-based modelling and optimization of parameters using the simulated annealing algorithm	Multiple inputs and single output model Function approximation neglected

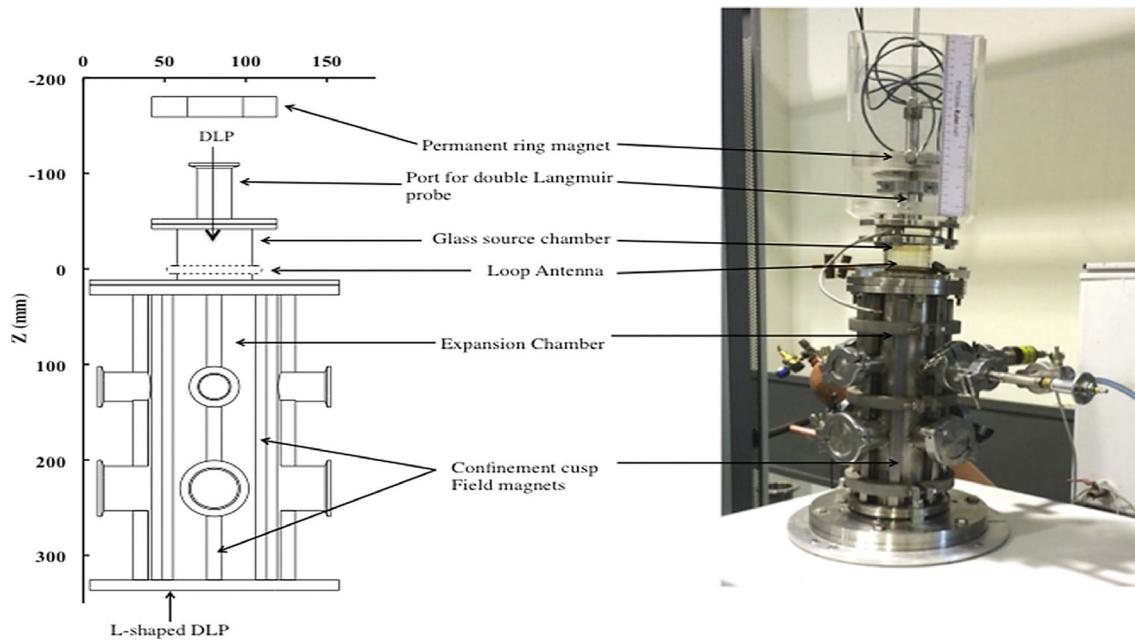


Fig. 1 Schematic and Photograph of the HELEN-I device Reproduced from [Citation: Review of Scientific Instruments **88**, 103,509 (2017) <https://doi.org/10.1063/1.4994058>], with the permission of AIP Publishing

H_2 gas is inserted in the vacuum vessel through a gas dosing valve, and the operator varies the source parameters to get the optimized plasma density and ion saturation current. The power generated by a 13.56 MHz RF power source is transferred to the antenna through an auto-tuning impedance matching network. RF power from 400 to 950 W is adjusted. Plasma is produced at three different pressure 6 mTorr; 7.5 mTorr and 30 mTorr. Further, the plasma density is enhanced by varying the magnetic fields (B field: 40 Gauss, 55 Gauss and 86 Gauss).

The ion source tuning is an arduous task because of its complex and nonlinear nature. Absence of input & output relationship between the source parameters and the outputs makes it challenging to predict the source parameters for a given process output. Moreover, there are several other interfering parameters such as azimuthal current, and axial voltage etc. affects the performance of ion source. The challenges, as mentioned earlier associated with the tuning of the ion source makes it cumbersome and tedious job for the researchers. The optimum source parameters are found through conducting the chain of real experiments by repeated and varied attempts continued until success. The limitation in predicting the optimized parameters correctly summons to have a faithful process model of an ion source. HELEN-I is used as a test setup to benchmark the developed algorithm since HELEN-I is a plasma source which has already been characterized in [9] manually with hydrogen plasma. Those experimental data, ~ 70% of the whole data set are used to teach the machine learning algorithms and to

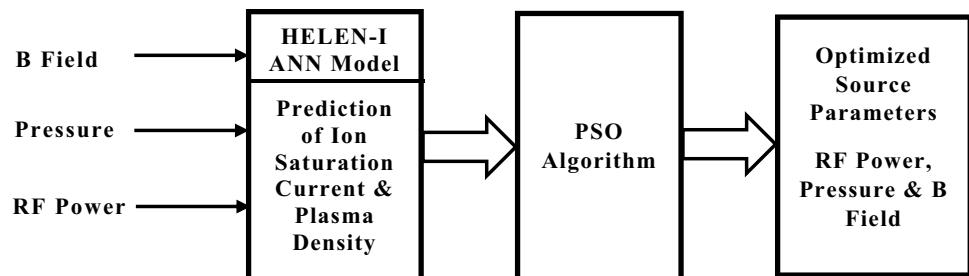
remaining ~ 30% data are reserved for testing the learning level. The ion saturation current and plasma density measurement values are considered as two independent experimental observed data sets while applying our algorithm. Though in reality, probe ion saturation current values are correlated with plasma density through plasma (mainly electron) temperature. However, in this work plasma temperature information is not taken into account and therefore ion saturation current and plasma density measurement values are treated as two independent experimental output data sets.

3 Outline for developing HELEN-I process model

The proposed approach, for intelligent process model development of HELEN-I ion source using ANN and Particle Swarm Optimization Technique (PSO), is shown in Fig. 2. The approach presented in the paper mainly consists of two steps:

- Data generation by conducting a series of experiments on HELEN-I.
- Development of ANN-based model of HELEN-I based on the experiments conducted in the laboratory.
- Optimization of source parameters PSO based optimization technique. PSO uses the ANN model as its objective function.

Fig. 2 Proposed process model of HELEN-I



This integrated approach allows the operator to accurately find the optimized source parameter for the desired output values without conducting additional experiments. Hence, saving lots of human efforts and time. The detailed description of the previous steps for developing the process model of HELEN is covered in the next sections.

4 ANN-based model for HELEN-I

In line with the prime objective of the present work outlined in the previous section, ANN techniques were studied for ion source modelling [6]. ANN mimics human brain information processing capabilities by gathering knowledge, relationships in the data to be processed and train itself through experience. ANNs have the capability of function approximation and considered a universal function estimators [10]. ANN can produce a model with an ability to accumulate experiential knowledge, which can be utilized for prediction. It gathers information and knowledge through learning process inspired by its environment [11]. The acquired knowledge is stored in the interneuron connection strengths, also called as synaptic weights. Neurons are the basic building block of ANNs [12], [13]. The ability of fast learning and function approximation from the inexplicit data make ANNs more worthy candidate for the modelling of a nonlinear and complex HELEN-I ion source [14], [15].

Hence, an encyclopedic process model of HELEN-I ion source having 3 input & 2 output is developed using ANN for fast and precise prediction of ion source outputs. Based on the previous studies conducted in [8] RF power, B Field, pressure, plasma density and ion saturation current are considered as three inputs and two outputs of the ion source. Experiments were conducted for generating the dataset having 189 samples for building the ANN-based ion source process model. Table 2 shows the ranges and setting of inputs selected for the experimental scheme for generating the data for ANN training based on studies carried out in [1]:

HELEN-1 ion source plasma experiment is to optimize the operational parameters within the range, suitable for an efficient fusion grade negative hydrogen ion source. The above consideration decides the pressure range. The range of the magnetic field is decided based on analytical modelling

Table 2 Experimental scheme for data generation for training the ANN model

Input	Ranges	Inputs settings
Pressure	6–30 mTorr	6 mTorr
		7.5 mTorr
		30 mTorr
B Field	40–86 Gauss	40 Gauss
		55 Gauss
		86 Gauss
RF power	200–950 Watts	400–900 Watts in steps of 25 Watts

considering simplified plasma description [3]. The available RF generator performance limits the RF power range.

Conducting real-time experiments is a time-consuming task. Each data point is an average over ten or even more data collected over time. With time the plasma-wall interaction and thermal effect come into play. Each experiment conducted for dataset generation took significant time and human efforts because tuning of input parameters and recording measurement subsequently is challenging due to the nonlinear nature of HELEN-I ion source hence dataset comprising 189 ($21 \times 3 \times 3$) data points is chosen for developing ANN model. The present NN model can quickly predict the optimum operational parameter space; otherwise, it usually takes a very long time to measure and to optimize the complicated operational parameter space through experiment only. Therefore, it can be concluded that the NN-based model is advantageous in this regard.

Although NN-based modelling techniques are prevalent, yet their robustness towards the drift and experimental error is a matter of investigation [16]. Hence, the experimental data set from HELEN-I should comprise these behaviours. Arun et al. [1] presented that the equilibrium plasma parameters, i.e. density, temperature, floating potential and plasma potential of HELEN-I only fluctuate within the error bars provided that the neutral gas pressure, RF power forwarded to the antenna and magnetic fields did not change. The forwarded RF power may change due to heating of the transmission lines for continuous-wave operations. But for

pulsed operation, this effect was not observed. The neutral pressure may show fluctuations within 10% from the set nominal value. But the effect due to this is much lower than the contribution due to measurement errors. The maximum variation in measured temperature could be $\pm 10\%$ whereas, the measured density variation could be as high as 40%, but it would still remain in the same order. Hence, drift and measurement error is not an issue for the HELEN-I setup.

There are several ANN architectures and configuration available such as radial basis networks, feed-forward backpropagation network, Hopfield recurrent network etc. Considering the nonlinear and complex nature of HELEN-I ion source, two-layered feed-forward backpropagation network (BPNN) (3-N1-N2-2) is tried and tested for the prediction performance evaluation. The neural network achieves greater freedom through multilayered feed-forward architectures and is capable of imitating reasonably complex function with a two-layered network [17] [18]. However, feed-forward networks cannot perform temporal computation [19]. On the other hand, Backpropagation, which is created by generalizing the Widrow-Hoff learning rule to multilayered network configurations and nonlinear differentiable transfer functions [20]. The computations are performed in the backward direction through the network. The network is trained using the input vectors and the corresponding target vectors until it approximates a function [21]. With a primary objective of creating an accurate and precise ANN model for predicting the desired outputs, an ANN-BPNN model having 3 inputs and 2 output is created. The 3-N1-N2-2 BPNN model was trained with different training techniques such as the Levenberg–Marquardt algorithm and Scaled conjugate gradient algorithm [22], [23]. The results obtained through the trained BPNN model were validated by comparing the errors with the experimental results.

The detailed description of the backpropagation algorithm and the comprehensive BPNN model developed for HELEN-I is described below.

4.1 Development of the BPNN model for HELEN-I

Figure 3 shows the schematic diagram of a BPNN configuration for HELEN-I. The BPNN network has two input layers, one input and one output layer. The two-layered sigmoid/linear network is utilized to approximate the relationship between inputs and outputs. Extensive numerical experiments were carried out to decide upon the number of neurons in each hidden layers. A BPNN having 18 neurons in the first hidden layer and 2 neurons in the second hidden layer is selected. Though, there are many variations of the backpropagation algorithm (such as Levenberg–Marquardt (LM), Resilient Backpropagation, Scaled Conjugate Gradient (SCG), One-Step Secant). However, for faster training,

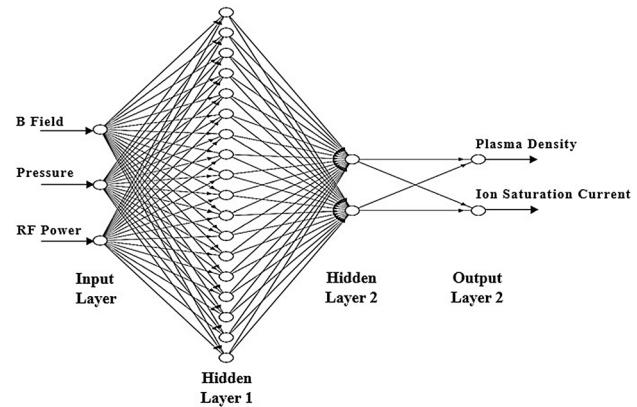


Fig. 3 Schematic diagram of a 3-18-2-2 BPNN model of HELEN-I

several algorithms are tested for training the 3-18-2-2 BPNN architecture. The detailed description of training the BPNN is given in the section to follow.

4.2 Training of HELEN-I BPNN model

There exist many algorithms for training a neural network model. Every algorithm has different characteristics in terms of speed, precision and memory requirements [24]. Gradient descent, Newton's method, Conjugate method, Quasi newton, Levenberg–Marquardt (LM) algorithm are the frequently used algorithms used for training the neural networks [25]. The BPNN model developed in the previous section utilizing the experimental data needs faster training for the prediction of outputs. Based on the previous studies, LM, Bayesian regularization, scaled conjugate gradient, BFGS quasi-newton algorithms are opted for training the BPNN model [26–29]. The performance of aforesaid training algorithms is measured using the Regression (R) value and the mean squared error (MSE) values [30]. R -value can be determined using the formula given by $R = \sqrt{1 - \frac{\text{SSE}}{\text{SS}_{yy}}}$

where $\text{SSE} = \sum (y - \hat{y})^2$; $\text{SS}_{yy} = \sum (y - \bar{y})^2$; y is the experimental outputs (targets) which are the experimental output vector–matrix having ion saturation current and plasma density values. \hat{y} is predicted output matrix (outputs) having the same dimension as y and \bar{y} is the mean of experimental output vector matrix.

The algorithms as mentioned above, use the 70%, 20% and 10% experimental dataset for training, testing and validation, respectively. Figure 4 shows the regression plots and error histograms for each training algorithm utilized for training the ANN model for the trained HELEN-I ANN model. The regression plots show the network's output or predicted output plotted in terms of corresponding experimental output [31]. The dashed line shown in Fig. 4a–d represents the perfect results when the network output exactly

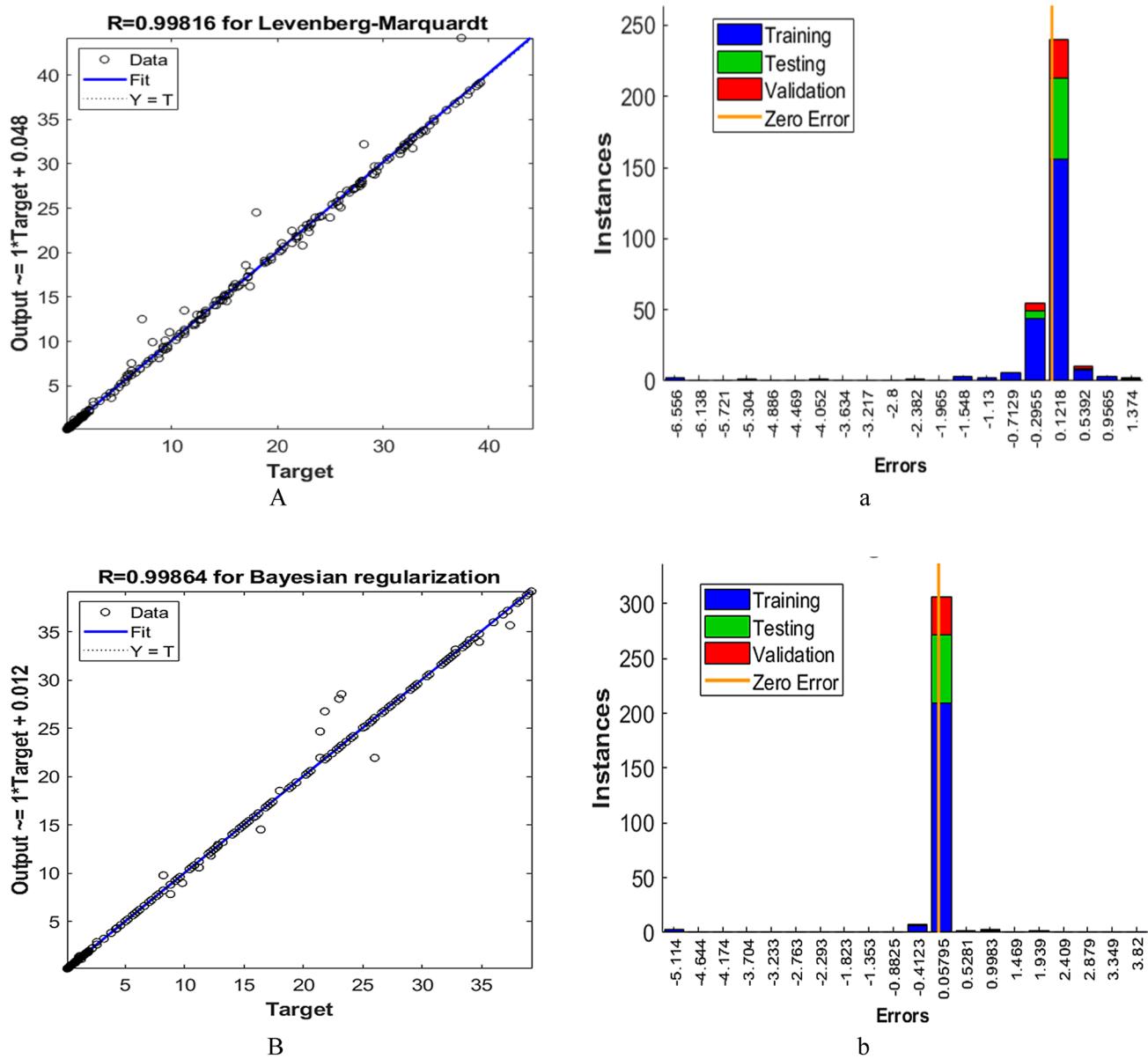
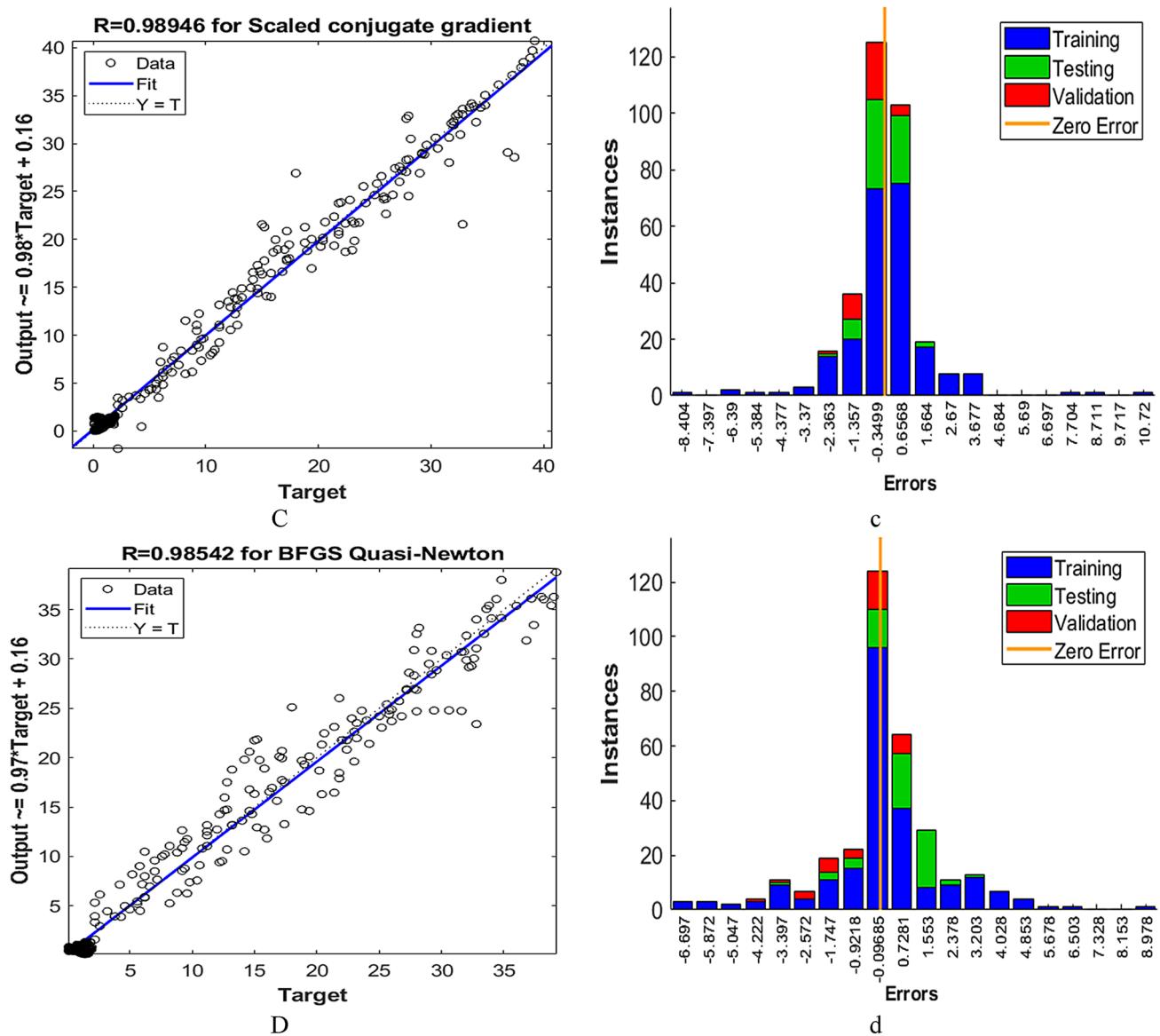


Fig. 4 Regression plot and associated error ($\text{Error } r = T - Y$) histograms for different training algorithms used for training the HELEN-I ANN model **A**, **a** Levenberg–marquardt (LM) **B**, **b** Bayesian regularization (BR) **C**, **c** Scaled conjugate gradient (SCG) **D**, **d** BFGS quasi-newton

matches with target (T) values (experimental output), i.e. ion saturation current and plasma density. In other words, the network output (Y) and target (T) should fall along a 45-degree dashed line ideally. The solid line shows the maximum linearity obtained between the outputs and targets for particular training algorithm. The scatter plot shows that some predicted outputs (the ion saturation and plasma density values) have poor fits with respect to associated target values. The large values of R (preferably close to 1) indicates the linear relationship between the targets and outputs hence the better ability of prediction.

Figure 4 a-d shows the error histograms for different training algorithms, which is an additional measure for the verification of network performance. The histograms shown give a clear indication of outliers, where the fit is significantly worse than the majority of data. It may be noted that the combined error for predicted output (i.e. network output in vector form) is very small for LM and BR algorithms. Table 3 summarizes the performance of the training algorithm.

After carefully analyzing the regression values, MSE and number of iterations for the training algorithms listed in Table 3, the superiority of training the BPNN network

**Fig. 4** (continued)**Table 3** Performance comparison of different training algorithms

Training algorithm	Regression value				Number of iterations	Mean squared error
	Training	Testing	Validation	Overall		
Levenberg–marquardt	0.99985	0.99445	0.99329	0.99816	50	0.5079
Bayesian regularization	1	0.99126	N.A	0.99864	1000	0.3673
Scaled conjugate gradient	0.99337	0.97423	0.98071	0.98946	67	2.7929
BFGS quasi-Newton	0.98762	0.96973	0.98465	0.98542	36	3.8657

model using LM over other algorithms can be established. Though the regression and MSE value for the Bayesian regularization is better than the LM, it has taken several

iterations in training the ANN model. Moreover, Bayesian regularization does not allow the user to fix the number of

iterations. The results obtained using LM are accurate and satisfies the performance goal criteria.

4.3 Function approximation and comparison of HELEN-I BPNN model performance with experimental data

The ANNs are known for their function approximation capabilities, and the 3-18-2-2 BPNN network efficiently approximates the function for HELEN-I. The approximated function is given by (1).

$$f_{\text{ANN}} = \left[\omega_2 \times \left(\frac{2}{1 + e^{-2[\omega_1] \times [I] + [b1]}} - 1 \right) + b_2 \right] \quad (1)$$

where,

$$I = \begin{bmatrix} i_{11} & i_{12} & \dots & \dots & \dots & i_{1,18} \\ i_{21} & i_{22} & \dots & \dots & \dots & i_{2,18} \\ i_{31} & i_{32} & \dots & \dots & \dots & i_{3,18} \end{bmatrix}_{3 \times 18} . \quad \text{is the input vector-matrix;}$$

$$\omega_1 = \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \\ \vdots & \vdots & \vdots \\ w_{18,1} & w_{18,2} & w_{18,3} \end{bmatrix}_{18 \times 3} \quad \text{is the weight matrix of hidden layer-1;}$$

$$\omega_2 = \begin{bmatrix} w_{o11} & w_{o12} & \dots & \dots & \dots & w_{o1,18} \\ w_{o21} & w_{o22} & \dots & \dots & \dots & w_{o2,18} \end{bmatrix}_{2 \times 18} ; \quad \text{is the weight matrix of hidden layer-2.}$$

$$b_1 = \begin{bmatrix} b_{11} \\ \vdots \\ b_{18} \end{bmatrix}_{18 \times 1} ; \quad b_2 = \begin{bmatrix} b_{i11} \\ b_{i12} \end{bmatrix}_{2 \times 1} ; \quad \text{are the biases of hidden layer-1 \& 2 respectively.}$$

After training BPNN, the performance of predictive HELEN-I model developed using the ANN technique is analyzed and validated using the experimental data. The HELEN-I function approximated by the Eq. (1) is utilized for predicting the outputs IS current, and Plasma Density for the experiment-1, 2 and 3 as per the experimental scheme suggested in Table 1. HELEN-I ANN model predicts the IS current and plasma density values at 30 mTorr constant pressure and three values of B -field (40 Gauss, 50 Gauss, 86 Gauss). The RF power is varied in 25 Watts steps starting from 400–900 Watts. Figure 5, 6, 7A and Fig. 5, 6, 7B show the comparison of predicted results versus actual measured output for IS current and plasma density for the input RFpower as explained formerly. As shown in figures, the predicted outputs are very much close to the actual experimental values of IS current (Fig. 5, 6, 7A). However, Figure 5, 6, 7B showing deviation from the experimental output in plasma density plots. The deviation can further be reduced by having a larger experimental data set for training. At the same time, it may also be noted that plasma density has very high order (i.e. 10^{18} per m^3) so such deviation may not significantly affect the performance of HELEN-I ANN Model.

The differences observed between predicted density values and the experimental data in Fig. 5, 6, 7B lie within the experimental uncertainty ~40%. Further, the performance of HELEN-I ANN model can be quantized by analyzing the prediction error for the individual case. Fig. 5, 6, 7a and Fig. 5, 6, 7b present the spectrum of prediction error for the network predicted IS current and plasma density. Careful observation of the histogram suggests that predicted IS current and plasma density values are very close to the measured values. Hence, most of the instances are having minimal errors or falling around the zero error point.

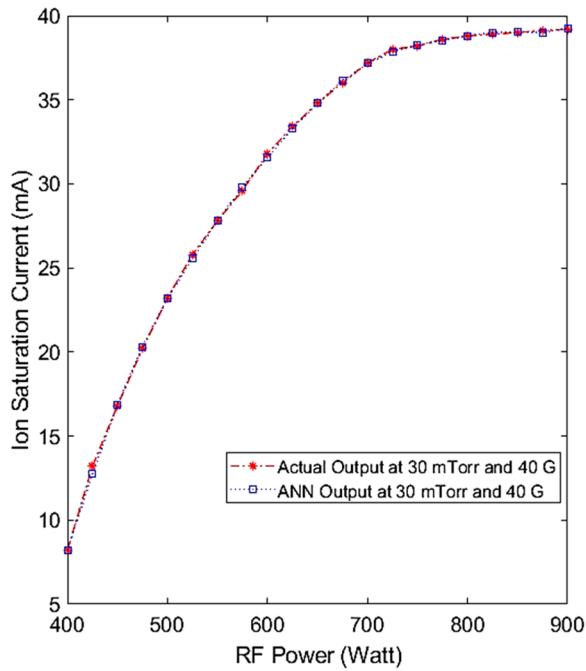
4.4 Retraining the ANN-Model

As discussed in Sect. 4, HELEN-I experimental setup has minimal effect of the drift and experimental error. However, small drift and measurement noise in the measured output for the previous input settings can adversely affect the trained ANN-model performance, so the model should be able to accommodate the drift and experimental error by transfer learning [32]. So the ANN-model trained in Sect. 4.3 is retrained using the sparse data set having experimental noise in the bounds mentioned in [1]. The ANN model was retrained using smaller data set while retaining the previously optimized weights of hidden layers of the trained ANN model.

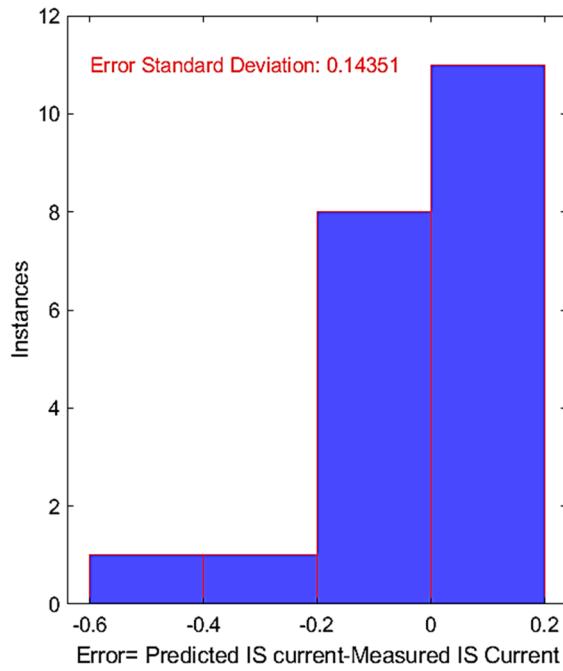
In our case, a dataset of 30 pair of inputs and outputs collected as per the experimental scheme mentioned in Sect. 4 was used to retrain the ANN. Figure 8a–d shows the results of retraining performance the ANN through the new dataset collected for each experimental conditions (40G, 55G, 86G). The pattern and standard deviation of error for predicted IS current, and the plasma density turns out to be very close to the patterns and standard deviation of prediction error estimated in Sect. 4.3 for the trained ANN. Hence, the developed ANN mimics the previously mentioned time-invariant nature of HELEN-I experimental setup. However, retraining of ANN with sparse data gives a fair idea about the reproducibility and robustness of elaborated ANN model and adds more features to the hidden layers for enhanced learning.

The detailed demonstration of results obtained by the training and retraining of ANN gives a clear understanding of the accuracy and robustness of the proposed model. Figure 9 shows the accuracy of the ANN model in terms of percentage error for predicted IS current and plasma density obtained through transfer learning via retraining.

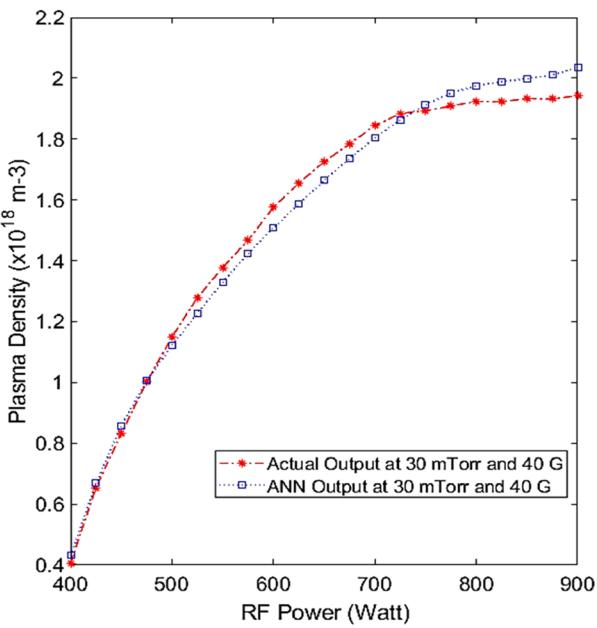
The HELEN-I ANN model presented in the paper predicts the IS current and plasma density with 99% and 95% accuracy, respectively. Hence, high accuracy and robust ANN model of HELEN-I device can be deployed as an objective function for the further optimization of the input parameters.



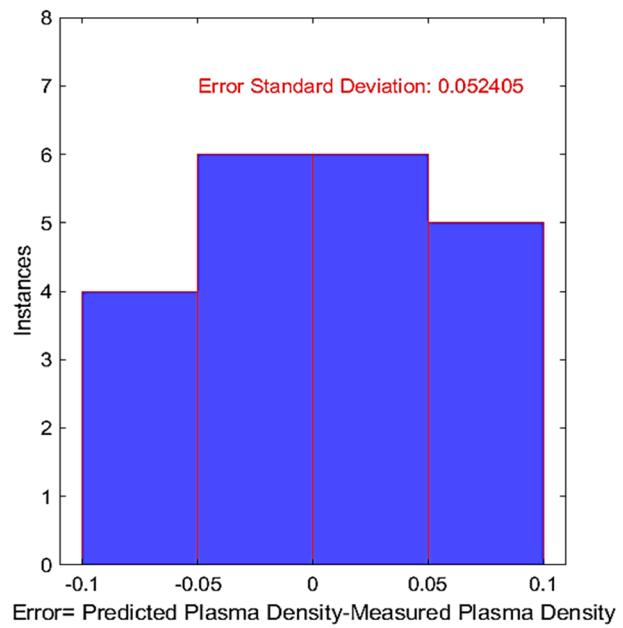
A



a



B



b

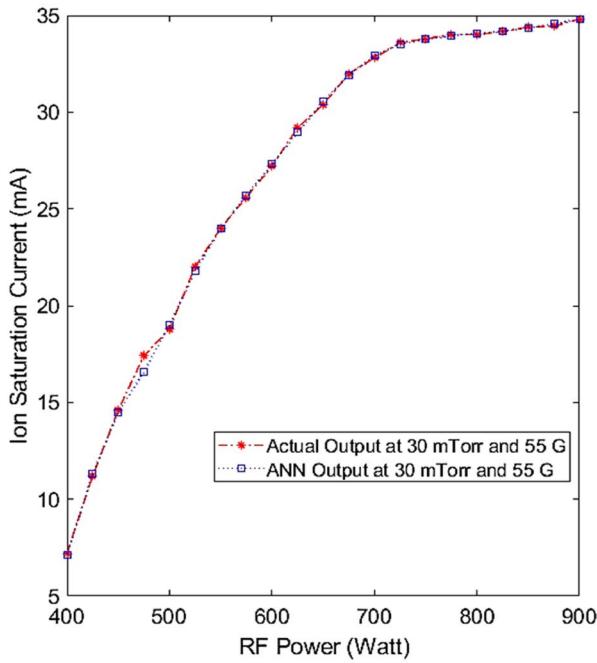
Fig. 5 **A** shows the Ion Saturation (IS) Current (mA) Versus RF Power (Watts) at 30 mTorr & 40 G **a** shows the corresponding error histogram for the predicted error and the associated standard deviation

B shows the Plasma Density Vresus RF Power (Watts) at 30 mTorr & 40 G **b** shows the corresponding histogram for the error for predicted plasma density and the associated error standard deviation

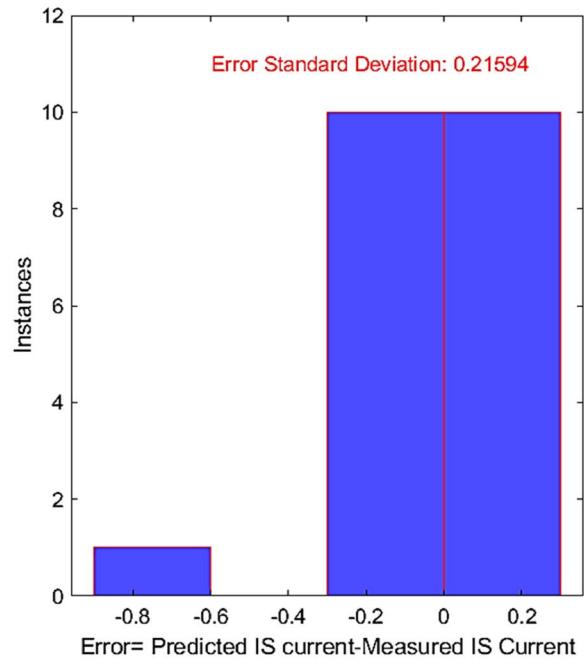
5 Optimization of source parameters of HELEN-I ion source

Currently, optimized source parameter selection of HELEN-I is based on the experimental methods, which

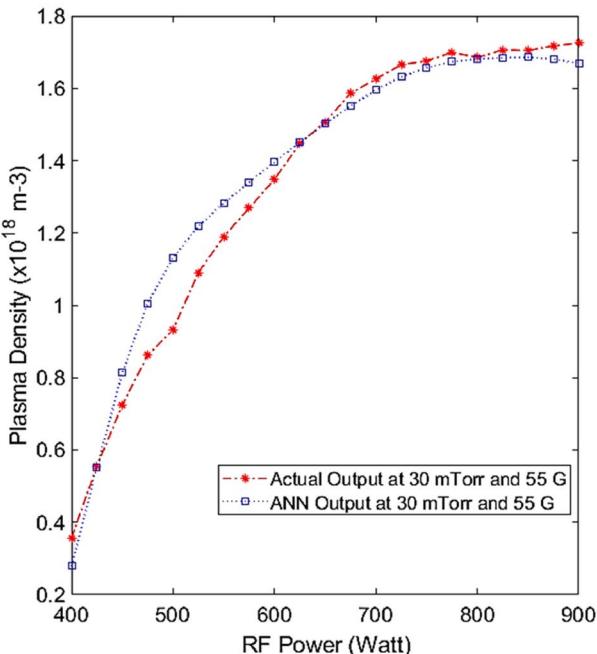
often takes much time and human efforts [1, 32, 33]. The input parameters and output performance parameters of an ion source have complex and nonlinear relationships, so it becomes very challenging to select the optimum source parameters for desired IS current and plasma density of



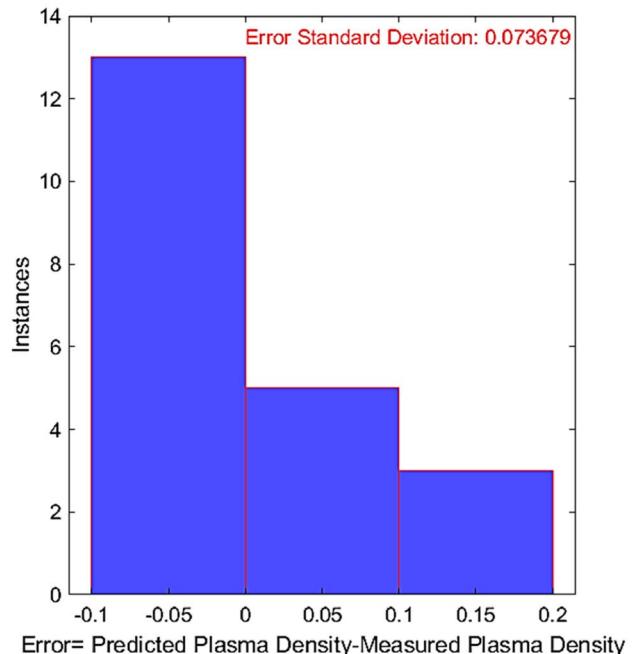
A



a



B



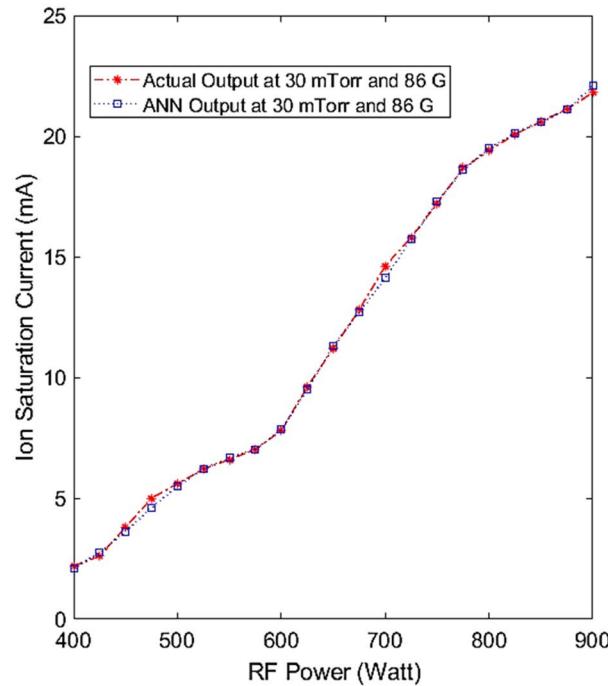
b

Fig. 6 **A** shows the Ion Saturation (IS) Current (mA) Versus RF Power (Watts) at 30 mTorr & 55 G **a** shows the corresponding error histogram for the predicted error and the associated standard deviation

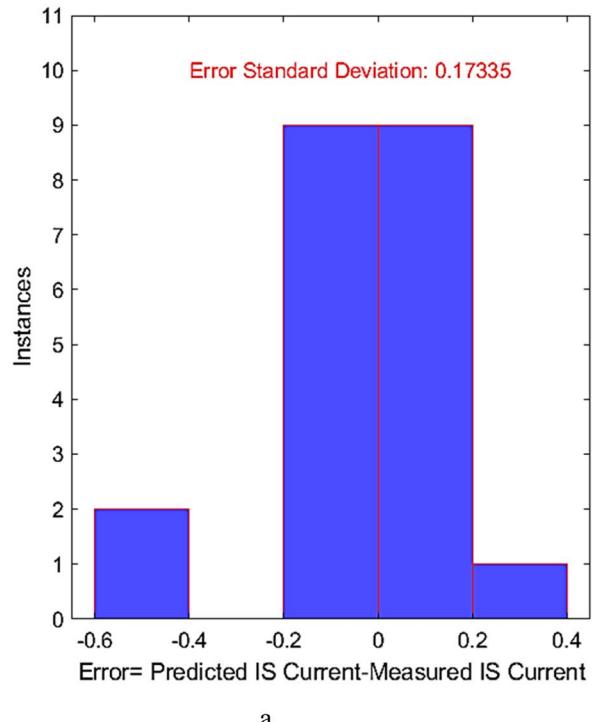
B shows the Plasma Density Versus RF Power (Watts) at 30 mTorr & 55 G **b** shows the corresponding error histogram of the predicted plasma density and the associated error standard deviation

ion source. Literature suggests Particle Swarm Optimization (PSO) is a better optimization technique among other available technique due to its fast convergence, robustness, more straightforward implementation and have

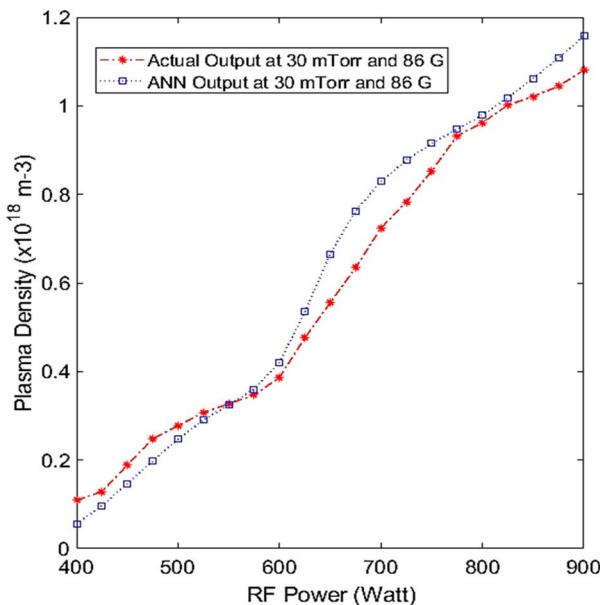
few parameters to adjust [34], [35]. In the present work, Fig. 10 shows the ANN-PSO approach for finding out the optimized source parameters of HELEN-I ion source as a test setup.



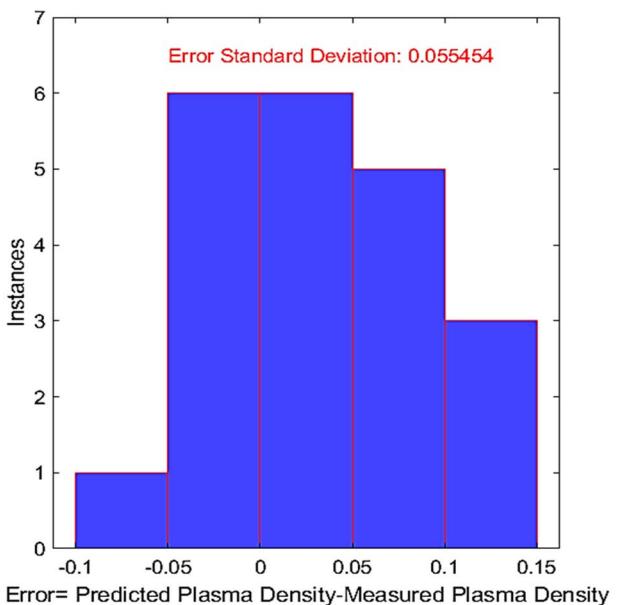
A



a



B



b

Fig. 7 **A** shows the Ion Saturation (IS) Current (mA) Versus RF Power (Watts) at 30 mTorr & 55 G. **a** shows the corresponding error histogram for the predicted IS current and the associated standard

deviation **B** shows the Plasma Density Versus RF Power (Watts) at 30 mTorr & 86 G. **b** shows the corresponding error histogram of the predicted plasma density and the associated error standard deviation

PSO is a heuristic global optimization algorithm. PSO technique searches several points in the iteration simultaneously. The swarm intelligence inspires the PSO technique. It computes the best position of the individual from the total population size as well as that of the total population

in terms of the objective function. p_{best} and g_{best} denote the best objective function value of the individual as well as of the particular group, respectively [36], [37]. PSO technique has better optimization capability, simple principle,

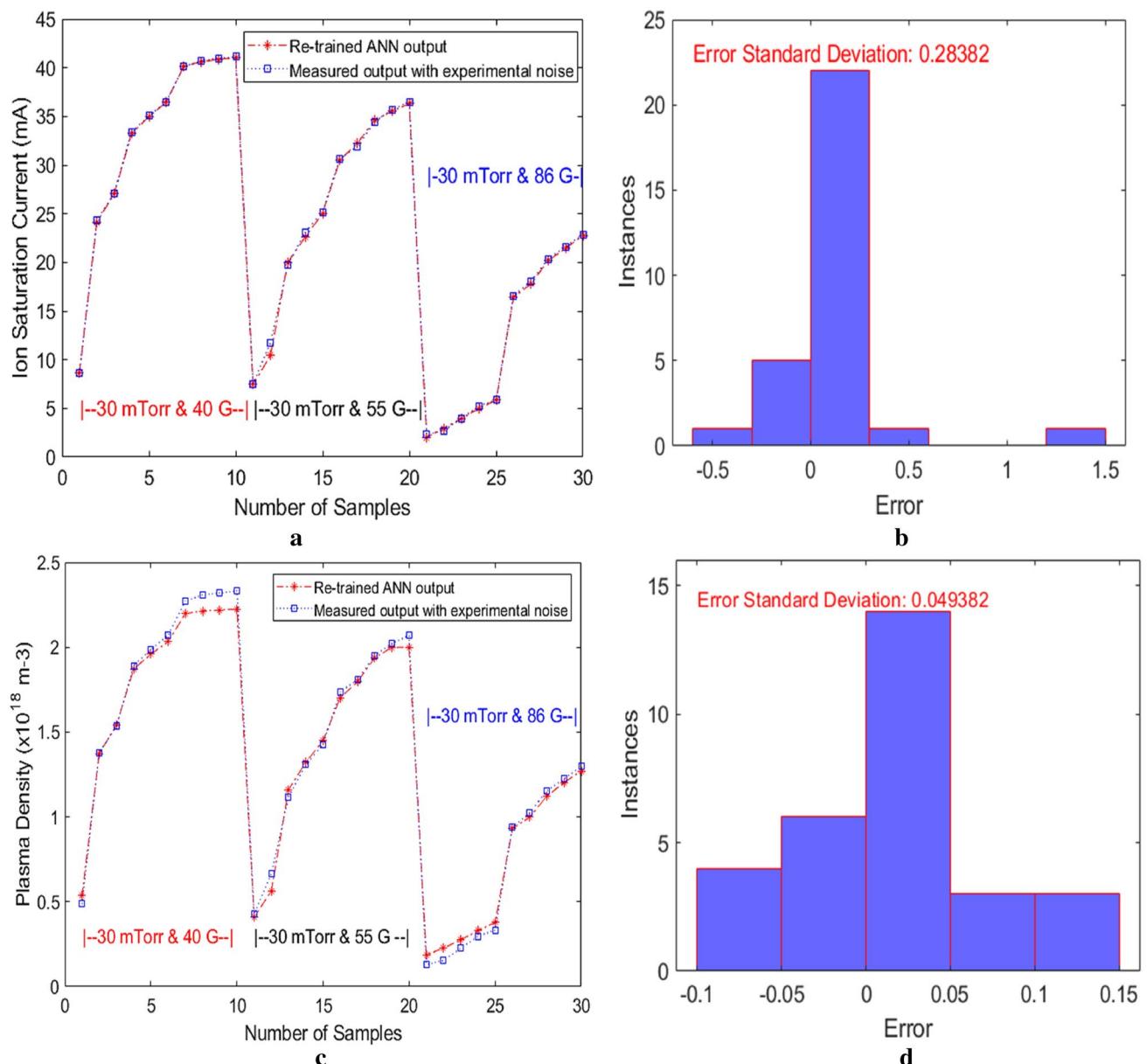


Fig. 8 **a** shows the number of samples (data taken randomly from the whole set of each experimental conditions (40G, 55G, 86G) Versus Measured IS current at having approximately 10% experimental noise and predicted IS current obtained through retraining the ANN network without manipulating the previously optimized weights of the hidden layer **b** shows the error [(predicted—measured) IS current] pattern, and its standard deviation associated with predicted IS current **c** shows the number of samples (data taken randomly from the

whole set of each experimental conditions (40G, 55G, 86G) Versus measured plasma density having approximately 40% experimental noise and predicted IS current obtained through retraining the ANN network without manipulating the previously optimized weights of the hidden layer **d** shows the error [(predicted—measured) plasma density] pattern and its standard deviation associated with predicted plasma density

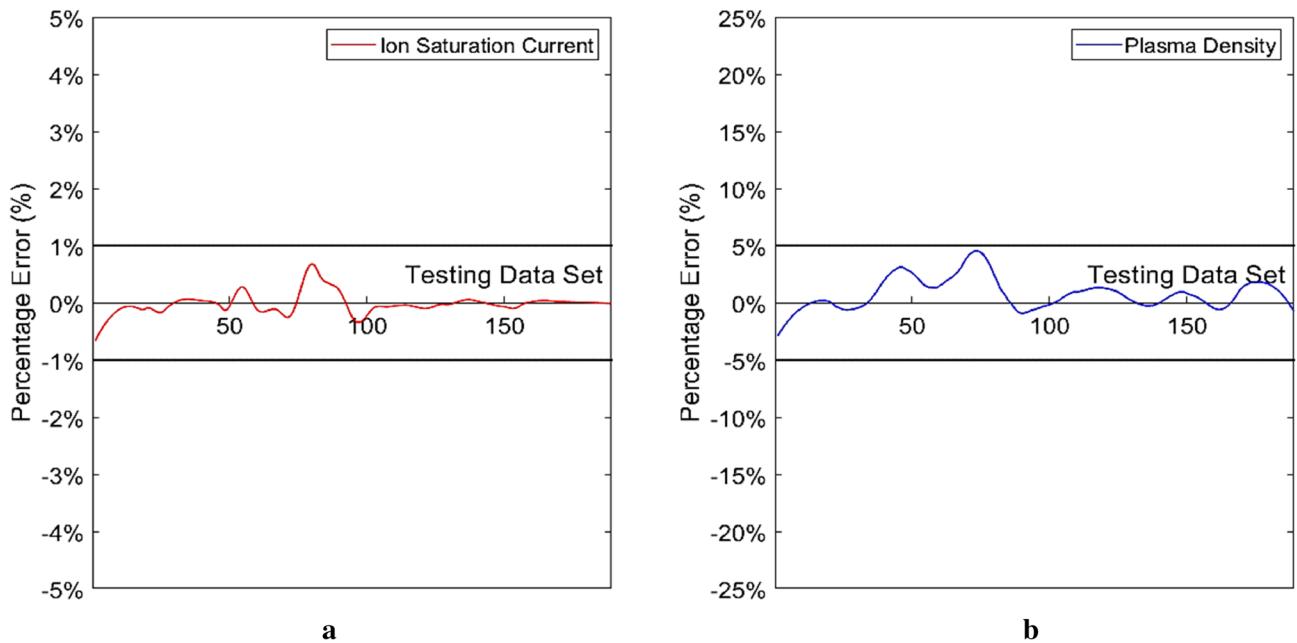


Fig. 9 Prediction accuracy of HELEN-I ANN Model. **a** Percentage error for predicted IS current **b** Percentage error for predicted plasma density

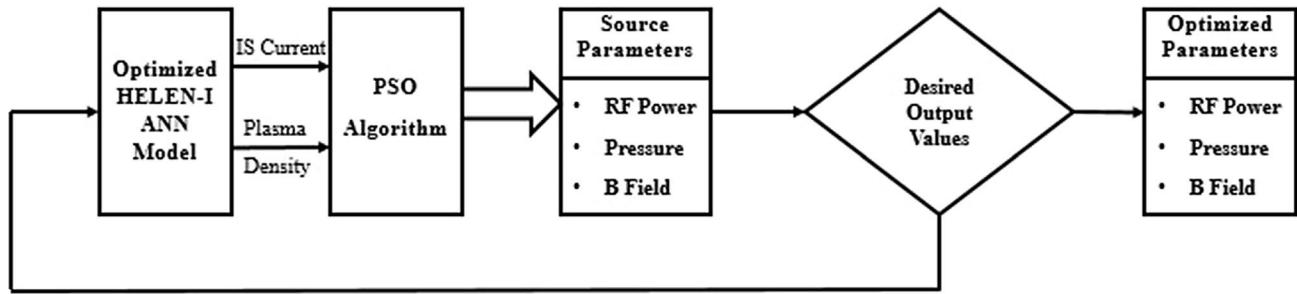


Fig. 10 Integrated ANN-PSO technique for *source* parameter optimization of HELEN-I

fast convergence characteristics, low computational cost and very few tunable parameters. Hence they have been extensively used to solve various optimization problems.

5.1 Following are the steps involved in the PSO algorithm

Step 1 Define tuning parameters c_1, c_2 , population size, number of iterations, X_{\min} (Lower bound), X_{\max} (Upper Bound), weights, number of inputs and outputs and fix the desired value of outputs (IS current and Plasma Density).

Step 2 Initialize starting position using Eq. (2) and initialize the initialize starting velocity $V_{\text{pso}} = 0$

$$X_{\text{pso}} = X_{\min} + \Gamma(0, 1) \otimes (X_{\max} - X_{\min}) \quad (2)$$

Step 3 For each member of the initial population compute the desired values of outputs using the HELEN-I approximated function Eq. (3)

Step 4 Define the fitness function as

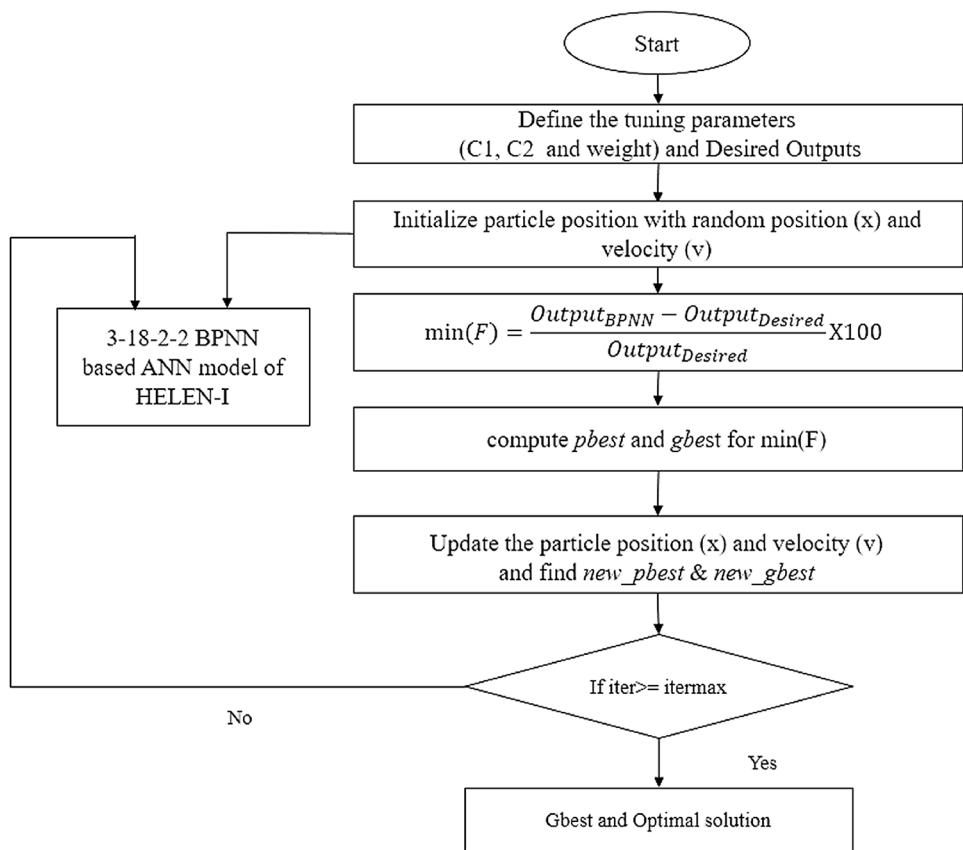
$$\text{fit} = 100 \times \frac{f_{\text{out}}^{\text{pso}} - f_{\text{out}}^{\text{desired}}}{f_{\text{out}}^{\text{desired}}} \quad (3)$$

Step 5 The p_{best} and g_{best} in the population are determined by the comparison of the fitness values the particles.

Step 6 The velocity and the position are updated using Eqs.(4) and (5), respectively. p_{best} and g_{best} are the minimum value after evaluating the objective function.

$$V_{\text{pso}} = \omega V_{\text{pso}} + c_1 \Gamma(0, 1)(p_{\text{best}} - X_{\text{pso}}) + c_2 \Gamma(0, 1)(g_{\text{best}} - X_{\text{pso}}) \quad (4)$$

Fig. 11 Flowchart of ANN-PSO approach for optimizing the source parameters of HELEN-I



$$X_{pso} = X_{pso} + V_{pso} \quad (5)$$

Step 7 Steps 2–6 are repeated for the maximum number of iterations defined in step 1.

Step 8 After the termination, the *g_{best}* solution becomes the optimal solution.

The ANN approximated BPNN model of HELEN-I is used as an objective function for finding the tuning parameters for the desired IS current and plasma density using PSO algorithm. The flowchart of ANN-PSO approach is shown in Fig. 11.

Before finding the optimal source parameters for tuning the ion source, a function space is constructed for or the trained BPNN model based on three experimental configurations, ie. *B* Fields 40 G, 55G, 86G & RF power from 400–900 W. The Fig. 12a and Fig. 12b show the function surface of BPNN output IS current and plasma density for the *B* Field and RF power, respectively.

As shown in Fig. 11a & 11b, the BPNN model outputs are current, and plasma density are produced for three *B* Field settings (40 G, 55G, 86 G), and RF power is varied in all three cases. The trained BPNN model outputs do not reflect the changes in *B* Field parameters. Therefore, the PSO algorithm searches the optimized RF power parameter. In the simulations, source parameters (RF power) for a total

21 data point of desired IS current, and plasma density from the experiment are optimized. The population size, upper bounds and lower bounds of the RF power input parameters are kept as 60, 950 Watts and 200 Watts, respectively. Figure 13 shows the convergence patterns for PSO algorithm implemented for optimizing the input parameter (RF power). Figure 13a shows the plot of the fitness of the global best particle with the iteration number. Since the control variables are generated randomly between the lower and upper allowable limits, some of the control variables generated randomly lie close to the optimum value. This makes the fitness value of the global best particle start close to the final optimal fitness value, and the algorithm converges very fast and takes approximately 20 iterations before saturation. However, all the randomly generated control variables do not start from the vicinity of the optimal value. A plot of the fitness value of the particle lying furthest away from the optimal solution is shown in Fig. 13b to demonstrate this phenomenon which takes approximately 50 iterations. It is seen that as iteration progresses, the entire population converges to the optimal value, and the worst fitness value in the current population approaches the fitness value of the global best particle.

The performance results obtained with the approach proposed in the present work are given in Table 4. As shown

Fig. 12 a b Function Space of the BPNN model. The desired IS current can be predicted with B Field and RF power Function Space of the BPNN model. The desired plasma density can be predicted with B Field and RF power

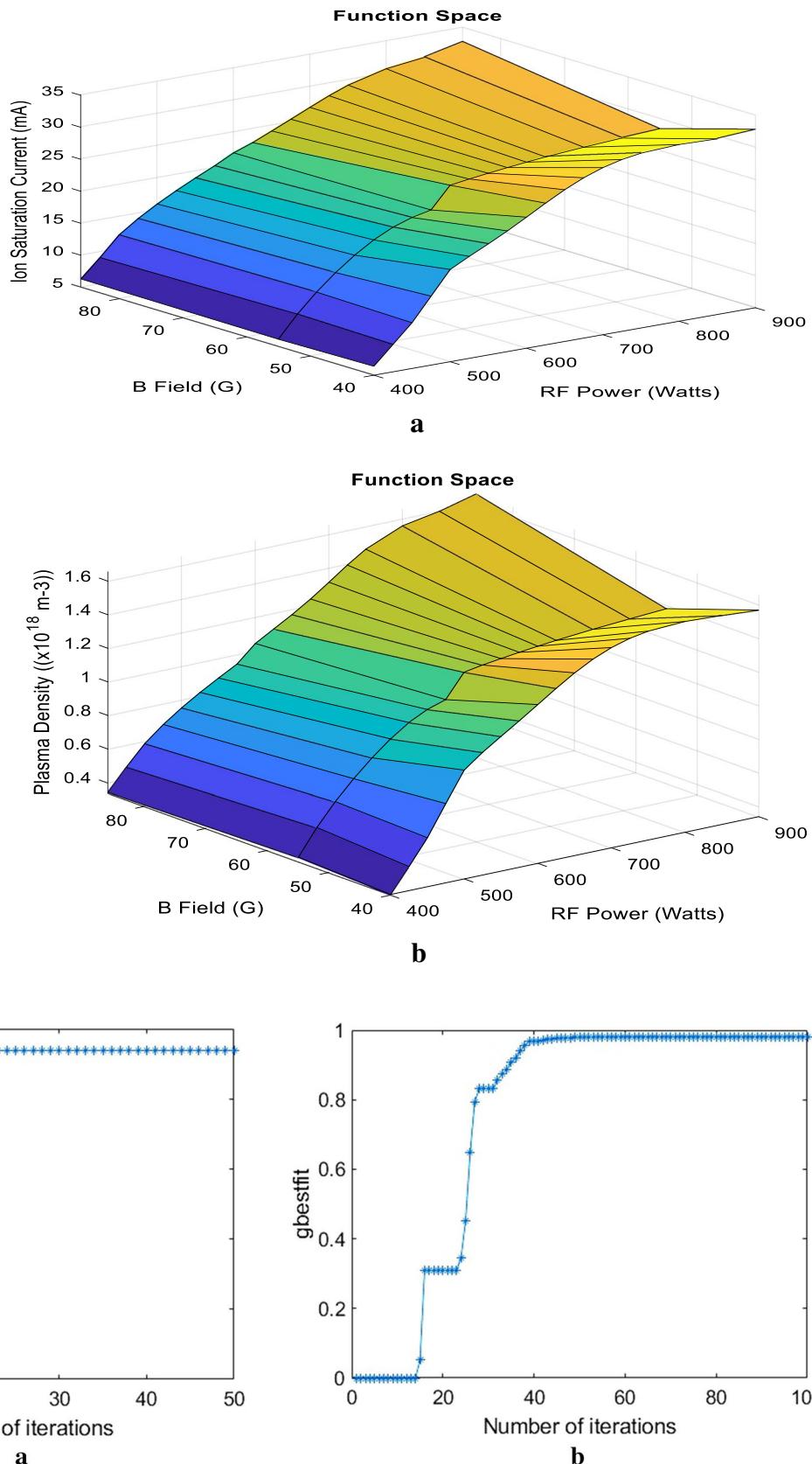


Fig. 13 gbestfit versus the number of iterations **a** when particle fitness value is close to the optimum fitness value. **b** when particle fitness value is furthest away from the optimum fitness value

Table 4 Performance result of the proposed approach. The predictive control parameter is close to the experimental value, and the predictive output is almost the same to the desired output

Desired results		Predicted parameter	Experimental	Values	ANN predicted outputs
Ion saturation current (mA)	Plasma density (10^{18} m^{-3})	RF power (Watt)	RF power (Watt)	Ion saturation current (mA)	Plasma density (10^{18} m^{-3})
8.2	0.4065	398.7219	400	8.0448	0.4142
13.2	0.6543	426.3426	425	13.2764	0.6505
16.8	0.8328	450.28	450	16.9759	0.8222
20.2	1.0013	476.36	475	20.2725	0.9977
23.2	1.15	499.9511	500	23.1729	1.1513
25.8	1.2789	524.4344	525	25.7041	1.2836
27.8	1.378	547.6344	550	27.6689	1.3840
29.6	1.4672	571.6344	575	29.4704	1.4736
31.8	1.5763	602.4142	600	31.6945	1.5815
33.4	1.6556	625.5814	625	33.3228	1.6594
34.8	1.725	647.7444	650	34.7517	1.7274
36	1.7845	670.7040	675	36.0008	1.7866
37.2	1.844	700.7040	700	37.2075	1.8436
38	1.8836	733.1645	725	38.0275	1.8822
38.2	1.8935	745.1718	750	38.2330	1.8919
38.5	1.9083	780.2518	775	38.5398	1.9509
38.8	1.9233	815.9034	800	38.8559	1.9205
38.89	1.9224	830.7531	825	38.9062	1.9290
39	1.9332	864.7860	850	39.0706	1.9297
39.1	1.9318	880.1267	875	39.0048	1.9300
39.2	1.9431	901.4927	900	39.2874	1.9388

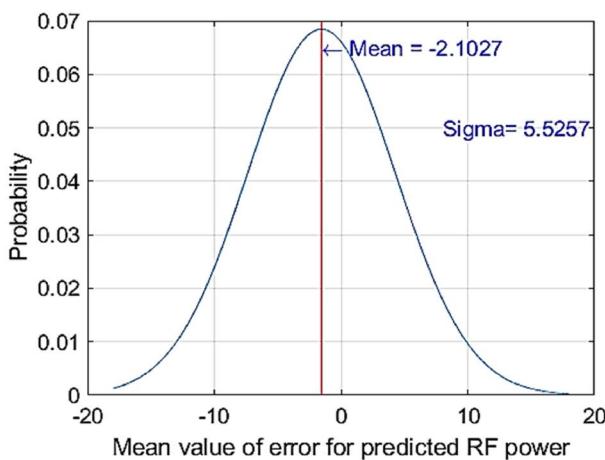
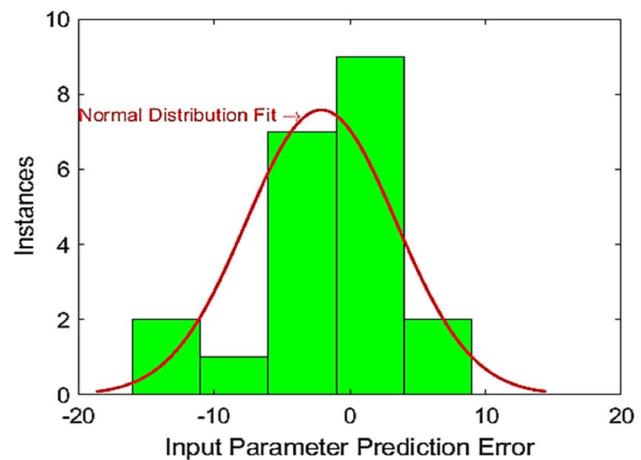
**a****b**

Fig. 14 **a** Shows the Probability Density Function (PDF) for prediction error for optimized RF power parameters **b** Prediction error histogram with the normal distribution fit

in Table 4, the predicted RF power source parameters are very close to the real RF power parameters, and the BPNN predicted IS current, and plasma density is almost the same as experimental outputs.

The simulation results show that the proposed ANN-based intelligent approach successfully searched the optimal source parameter for the desired IS current and the plasma density of HELEN-I ion source. The codes were

written using MATLAB 2016a on a laptop with Intel core i5 processor.

Figure 14 presents the probability density function (PDF) for the prediction error [38]. The predicted error falls in six sigma range having the mean error of -2.1207 for 21 optimized input parameter, i.e. RF power for the target values. The predicted values of RF power input parameter are very close to the experimental RF power input parameter and falls within the $\pm 2\%$ error bound for the reference target IS current and plasma density.

6 Conclusion

The ion source is a highly nonlinear and complex system; hence tuning the ion source is a challenging and time-consuming task. The manual operation of the system for optimizing the source parameter to get the desired plasma density and ion saturation current may increase cost and efforts. Any approach which can predict the desired outputs without the manual operation of ion source may significantly help the scientific community and researchers involved in research related to fusion energy. In-line with the same, this paper developed an ANN-based model of HELEN-I ion source for predicting the outputs and also combined the proposed model with a stochastic optimization technique for optimizing the source parameters with minimal intervention with the experimental setup. The work presented in the paper demonstrates the efficiency of various learning algorithms used for training the ANN model and compares their performance. A three input (RF power, B Field, Pressure) and two output (IS current, plasma density) hydrogen ion helicon plasma source is successfully modelled using the ANN technique. The multilayered BPNN is trained using the Levenberg–Marquardt algorithm. The algorithm utilises the experimental data averaged over multiple values which significantly reduces the drift effects. The optimised ANN model is further retrained using a sparse data set having experimental noise while the features of formerly optimised ANN model were preserved. This approach significantly adds more learning data points to the network and broaden the prediction horizon of the ANN model. Hence, exhibits reproducibility and robustness of the proposed model over-time.

The optimised ANN model is further utilized for the source parameter optimization using PSO algorithm. The experimental studies validate the predicted optimized source parameter values. The paper also approximates the function for ANN-based HELEN-I ion source and allows the operator to retrain the network with the new dataset. The easy to train and retrain capability of this model allows the operator to utilize it for modelling other ion sources with slight modification. However, initialization of hidden layer weights

is warranted prior to the training in case of the different experimental setup. The proposed ANN combined with PSO algorithm can significantly save human efforts and time. Finally, the study demonstrated an efficient, robust ANN-based model combined with a stochastic optimization technique for control parameter tuning. The future study aims at conducting an exhaustive study to explore other machines learning-based technique for modelling and control of ion source with the larger dataset and develop a machine learning-driven controller for automatic input parameter optimization and measurements.

References

- Pandey A, Bandyopadhyay M, Sudhir D, Chakraborty A (2017) Performance evaluation of a permanent ring magnet based helicon plasma source for negative ion source research. *Rev Sci Instrum.* <https://doi.org/10.1063/1.4994058>
- Boswell RW (1984) Very efficient plasma generation by whistler waves near the lower hybrid frequency. *Plasma Phys Control Fusion* 26:1147–1162. <https://doi.org/10.1088/0741-3335/26/10/001>
- Pandey A, Sudhir D, Bandyopadhyay M, Chakraborty A (2016) Conceptual design of a permanent ring magnet based helicon plasma source module intended to be used in a large size fusion grade ion source. *Fusion Eng Des* 103:1–7. <https://doi.org/10.1016/j.fusengdes.2015.11.025>
- Edelen AL, Biedron SG, Chase BE, Edstrom D, Milton SV, Stabile P (2016) Neural networks for modeling and control of particle accelerators. *IEEE Trans Nucl Sci* 63(2):878–897. <https://doi.org/10.1109/TNS.2016.2543203>
- Scheinker A, Edelen A, Bohler D, Emma C, Lutman A (2018) Demonstration of model-independent control of the longitudinal phase space of electron beams in the linac-coherent light source with femtosecond resolution. *Phys Rev Lett* 121:44801. <https://doi.org/10.1103/PhysRevLett.121.044801>
- Huang X, Corbett J, Safranek J, Wu J (2013) An algorithm for online optimization of accelerators. *Nucl Instruments Methods Phys Res Sect A Accel Spectrometers Detect Assoc Equip* 726:77–83. <https://doi.org/10.1016/j.nima.2013.05.046>
- Brown SK, Mead WC, Bowling PS, Jones RD, Barnes CW (1994) Optimization and control of a small angle ion source using an adaptive neural network controller (invited). *Rev Sci Instrum* 65:1411–1415. <https://doi.org/10.1063/1.1144978>
- Kong YB, Hur MG, Lee EJ, Park JH, Park YD, Yang SD (2016) Predictive ion source control using artificial neural network for RFT-30 cyclotron. *Nucl Instruments Methods Phys Res Sect A Accel Spectrometers Detect Assoc Equip* 806:55–60. <https://doi.org/10.1016/j.nima.2015.09.095>
- Scheinker DBA, Rees D, Garnett B, Milton S, Edelen AL (2018) Applying artificial intelligence to accelerators. IPAC 2018 Ninth Int Part Accel Conf. <https://doi.org/10.18429/JACOW-IPAC2016-TUOB02>
- Pandey A, Mukherjee D, Borah D, Bandyopadhyay M, Tyagi H, Yadav R, Chakraborty A (2019) Characterization of hydrogen plasma in a permanent ring magnet based helicon plasma source for negative ion source research. *Plasma Phys Control Fusion*. <https://doi.org/10.1088/1361-6587/ab0f09>
- Szegedy C, Zaremba W, Sutskever I, Bruna J, Erhan D, Goodfellow I, Fergus R (2014) Intriguing properties of neural networks.

- In: 2nd International Conference on Learning Representations ICLR 2014 - Conference Track Proceedings. Banff, pp 1–10
12. Shanmuganathan S, Samarasinghe S (2016) Artificial neural network modelling. <https://doi.org/10.1088/0031-9155/49/12/004>
 13. R. Murray-Smith, D. Neumerkel, D. Sbarbaro-Hofer (1992) Neural networks for modelling and control of a nonlinear dynamic system. IEEE Int Symp Intell Control—Proc <https://doi.org/10.1109/ISIC.1992.225125>
 14. Fernandez FG, Santos ISL, Redondo FL, Izquierdo S, Vega-carriño HR, Cervantes-viramontes JM, Martín V, Caocci G, Radi A, Luis J, Rosa G, Mareš J (2013) Artificial neural networks – architectures and applications. In: Tech. <https://doi.org/10.5772/3409>
 15. Hussain MA (1999) Review of the applications of neural networks in chemical process control—simulation and online implementation. *Artif Intell Eng* 13:55–68. [https://doi.org/10.1016/S0954-1810\(98\)00011-9](https://doi.org/10.1016/S0954-1810(98)00011-9)
 16. Shen YF, Pokharel R, Nizolek TJ, Kumar A, Lookman T (2019) Convolutional neural network-based method for real-time orientation indexing of measured electron backscatter diffraction patterns. *Acta Mater* 170:118–131. <https://doi.org/10.1016/j.actam.2019.03.026>
 17. Pieck M (2009) Artificial intelligence research in particle accelerator control systems for beam line tuning. In: Proceedings 24th LINAC 2008, Victoria, British Columbia, pp 314–316
 18. Lecun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444. <https://doi.org/10.1038/nature14539>
 19. Demuth H, Beale M (2005) MATLAB neural networks toolbox v4 user's guide. Mathworks Inc. <https://doi.org/10.1016/j.neuett.2005.10.002>
 20. Wilamowski B, Irwin J (eds) (2011) The industrial electronics handbook - five volume set. CRC Press, Boca Raton. <https://doi.org/10.1201/NOE1439802892>
 21. Wang W, Pedretti G, Milo V, Carboni R, Calderoni A, Ramaswamy N, Spinelli AS, Ielmini D (2019) Computing of temporal information in spiking neural networks with ReRAM synapses. *Faraday Discuss* 213:453–469. <https://doi.org/10.1039/c8fd0097b>
 22. Abdi H, Valentin D, Edelman B, O'Toole AJ (1996) A Widrow-Hoff learning rule for a generalization of the linear auto-associator. *J Math Psychol* 40:175–182. <https://doi.org/10.1006/jmps.1996.0017>
 23. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representations by error propagation (No. ICS-8506). Calif Univ San Diego La Jolla Inst Cogn Sci 1:318–362. <https://doi.org/10.1016/B978-1-4832-1446-7.50035-2>
 24. Hagan MT, Menhaj MB (1996) Brief Papers. *Brain Cogn* 32:273–344. <https://doi.org/10.1006/brcg.1996.0066>
 25. Møller MF (1993) A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw* 6:525–533. [https://doi.org/10.1016/S0893-6080\(05\)80056-5](https://doi.org/10.1016/S0893-6080(05)80056-5)
 26. Awolusi TF, Oke OL, Akinkurodere OO, Sojobi AO, Aluko OG (2019) Performance comparison of neural network training algorithms in the modeling properties of steel fiber reinforced concrete. *Heliyon* 5:e01115. <https://doi.org/10.1016/j.heliyon.2018.e01115>
 27. Mukherjee I, Routroy S (2012) Comparing the performance of neural networks developed by using Levenberg–Marquardt and Quasi-Newton with the gradient descent algorithm for modelling a multiple response grinding process. *Expert Syst Appl* 39:2397–2407. <https://doi.org/10.1016/j.eswa.2011.08.087>
 28. Andrei N (2007) Scaled conjugate gradient algorithms for unconstrained optimization. *Comput Optim Appl* 38:401–416. <https://doi.org/10.1007/s10589-007-9055-7>
 29. Lourakis MIA (2005) A brief description of the Levenberg–Marquardt algorithm implemented by levmar. *Matrix* 3:2. <https://doi.org/10.1016/j.ijinfomgt.2009.10.001>
 30. Botchkarev A (2018) Evaluating performance of regression machine learning models using multiple error metrics in azure machine learning studio. *SSRN Electron J.* <https://doi.org/10.2139/ssrn.3177507>
 31. Hagan MT, Demuth HB, De Jesús O (2002) An introduction to the use of neural networks in control systems. *Int J Robust Nonlinear Control* 12:959–985. <https://doi.org/10.1002/rnc.727>
 32. Ian G, Yoshua B, Aaron C (2016) Deep learning. https://www.deeplearningbook.org/front_matter.pdf. Accessed 08 May 2020
 33. Kang KU, An DH, Chang HS, Chai JS (2008) Performance optimization of H—multicusp ion source for KIRAMS-30 cyclotron. *Rev Sci Instrum* 79:2006–2009. <https://doi.org/10.1063/1.2819336>
 34. Abdoule Z, Gastli A, Ben-Brahim L, Haouari M, Al-Emadi NA (2017) Review of optimization techniques applied for the integration of distributed generation from renewable energy sources. *Renew Energy* 113:266–280. <https://doi.org/10.1016/j.renene.2017.05.087>
 35. Bai Q (2016) Analysis of particle swarm optimization algorithm. *Comput Inf Sci* 3:180–184
 36. Chunkai Zhang, Huihe Shao, Yu Li (2002) Particle swarm optimisation for evolving artificial neural network, pp 2487–2490. <https://doi.org/10.1109/icsmc.2000.884366>
 37. P.N. Suganthan (1999) Particle swarm optimiser with neighbourhood operator. In: Proc 1999 Congr Evol Comput CEC 1999. Vol 3, pp 1958–1962 10.1109/CEC.1999.785514
 38. Ribeiro MI (2004) Gaussian probability density functions: properties and error characterization. Institute for systems and robotics, Instituto superior tcnico. [Online]. Available: http://hans.fugal.net/comps/papers/ribeiro_2004.pdf Accessed 08 May 2020

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