

RESEARCH

Open Access



# A hybrid model based on general regression neural network and fruit fly optimization algorithm for forecasting and optimizing paclitaxel biosynthesis in *Corylus avellana* cell culture

Mina Salehi<sup>1\*</sup> , Siamak Farhadi<sup>1</sup>, Ahmad Moieni<sup>1</sup>, Naser Safaie<sup>2</sup> and Mohsen Hesami<sup>3</sup>

## Abstract

**Background:** Paclitaxel is a well-known chemotherapeutic agent widely applied as a therapy for various types of cancers. In vitro culture of *Corylus avellana* has been named as a promising and low-cost strategy for paclitaxel production. Fungal elicitors have been reported as an impressive strategy for improving paclitaxel biosynthesis in cell suspension culture (CSC) of *C. avellana*. The objectives of this research were to forecast and optimize growth and paclitaxel biosynthesis based on four input variables including cell extract (CE) and culture filtrate (CF) concentration levels, elicitor adding day and CSC harvesting time in *C. avellana* cell culture, as a case study, using general regression neural network-fruit fly optimization algorithm (GRNN-FOA) via data mining approach for the first time.

**Results:** GRNN-FOA models (0.88–0.97) showed the superior prediction performances as compared to regression models (0.57–0.86). Comparative analysis of multilayer perceptron-genetic algorithm (MLP-GA) and GRNN-FOA showed very slight difference between two models for dry weight (DW), intracellular and extracellular paclitaxel in testing subset, the unseen data. However, MLP-GA was slightly more accurate as compared to GRNN-FOA for total paclitaxel and extracellular paclitaxel portion in testing subset. The slight difference was observed in maximum growth and paclitaxel biosynthesis optimized by FOA and GA. The optimization analysis using FOA on developed GRNN-FOA models showed that optimal CE [4.29% (v/v)] and CF [5.38% (v/v)] concentration levels, elicitor adding day (17) and harvesting time (88 h and 19 min) can lead to highest paclitaxel biosynthesis (372.89  $\mu\text{g l}^{-1}$ ).

**Conclusions:** Great accordance between the predicted and observed values of DW, intracellular, extracellular and total yield of paclitaxel, and also extracellular paclitaxel portion support excellent performance of developed GRNN-FOA models. Overall, GRNN-FOA as new mathematical tool may pave the way for forecasting and optimizing secondary metabolite production in plant in vitro culture.

**Keywords:** Anticancer, In vitro culture, Secondary metabolite, Optimization problem, Artificial intelligence, Mathematical modeling

## Background

Paclitaxel as a microtubule-stabilizing agent is widely used for the treatment of a vast range of cancers [1]. This natural source diterpene alkaloid, paclitaxel, is the most prosperous anticancer drug owing to its unique action

\*Correspondence: mina.salehi@modares.ac.ir; salehi.minasm@gmail.com

<sup>1</sup> Department of Plant Genetics and Breeding, Faculty of Agriculture, Tarbiat Modares University, P.O. Box 14115-336, Tehran, Iran  
Full list of author information is available at the end of the article



© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

mechanism [2]. Paclitaxel arrests the disassembly of the microtubule, and in this unique way inhibits mitosis and proliferation of cancerous cells [3, 4].

In vitro culture of hazel (*Corylus avellana*) has been named as a promising and low-cost strategy for paclitaxel production [5–13]. The advantages of paclitaxel production through *C. avellana* cell culture are that the establishment of its in vitro culture is more straightforward than that of *Taxus* [6–12], and also the response of hazel to genetic manipulation through *Agrobacterium* is likely more hopeful as compared to that of *Taxus* since *C. avellana* is a dicotyledonous plant [14]. Obtaining high-producing cell cultures is essential for producing secondary metabolites by way of plant in vitro culture [15]. Biosynthesizing bioactive compounds in plants are influenced by various factors [6–8, 16–19]. Fungal elicitors including cell extract (CE) and culture filtrate (CF) have been described as an impressive strategy for improving paclitaxel biosynthesis in cell suspension culture (CSC) of *C. avellana* [6, 7, 10–13]. Fungal elicitor type, concentration level and adding time as well as exposure time of cell culture to it (harvesting time) should be optimized to achieve the highest biosynthesis of paclitaxel in *C. avellana* CSC [6, 7, 10–13]. Precise analysis of the effects of these factors and their optimal selection would be a step forward to commercialize the bioprocess of *C. avellana* cells for paclitaxel mass production. Paclitaxel biosynthesis and its elicitation are the complex biological processes because they are influenced by multiple factors and their nonlinear interactions. Optimizing these mentioned factors by performing experiment is laborious, costly and time-consuming. Robust nonlinear computational methods can effectively predict the optimized conditions for multifactorial process [20, 21] such as paclitaxel biosynthesis.

Traditional modeling and forecasting methods including regression models display insignificant non-linear predictive and fitting ability [7, 12, 13]. Artificial intelligence (AI) is applied to address matters that cannot be clarified by traditional computational methods. Artificial neural networks (ANNs) are one of the main parts of AI discovering complex nonlinear relationships amongst input and output data [7, 13, 24–30]. Indeed, ANNs are brain-inspired systems that emulate human brain capability of sensing and thinking, in a simplified way, to processes information and identify patterns [31]. ANNs obtain their intelligence by discovering the relationships and patterns in data, and learn using experience [31].

General regression neural network (GRNN) developed by Specht [32] is a kind of radial basis function (RBF) networks, and one of the most popular neural networks. GRNN as a powerful regression method with a dynamic network structure can successfully solve problems with

extremely difficult and unknown solution in various fields [33–39]. GRNN displays strong non-linear mapping capability, high fault tolerance, high robustness in the solution of complex problems, very fast network training speed, ease of implementation and simplicity of network structure [32, 40]. It is highly regretful that GRNN has not been used to model secondary metabolite biosynthesis in plant in vitro culture.

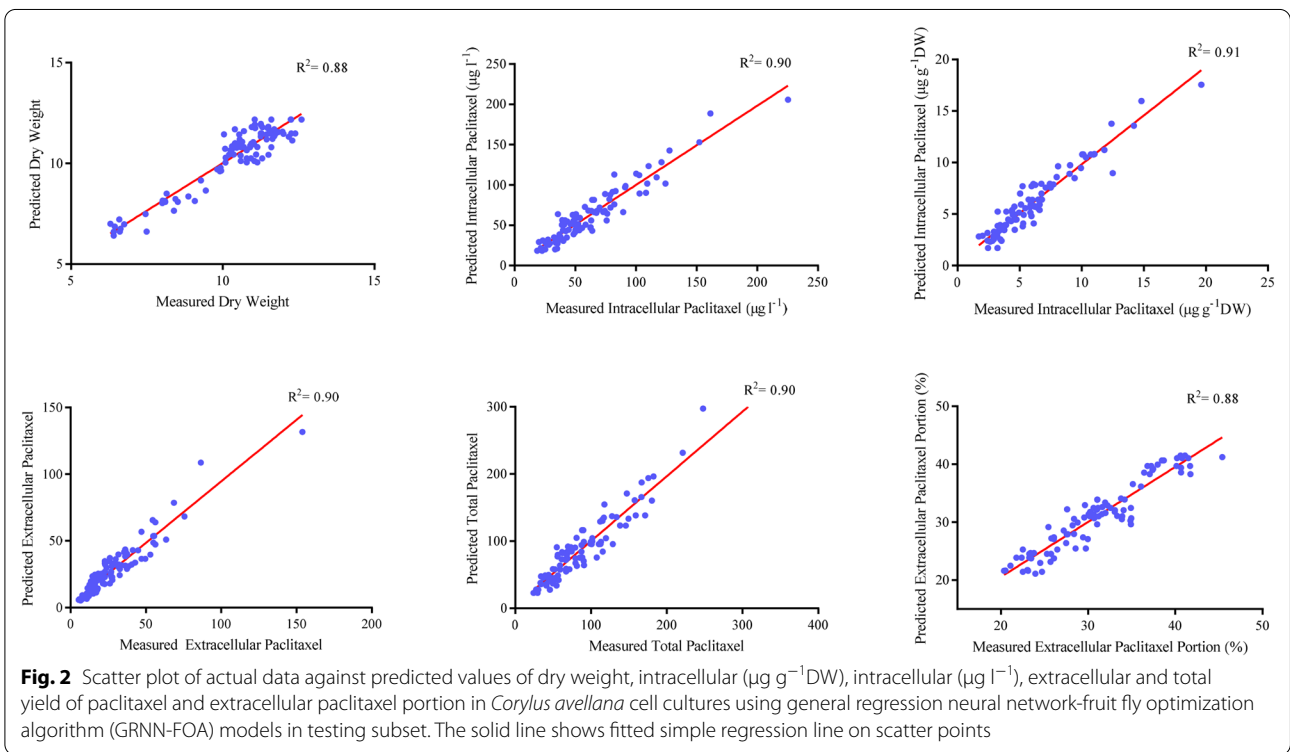
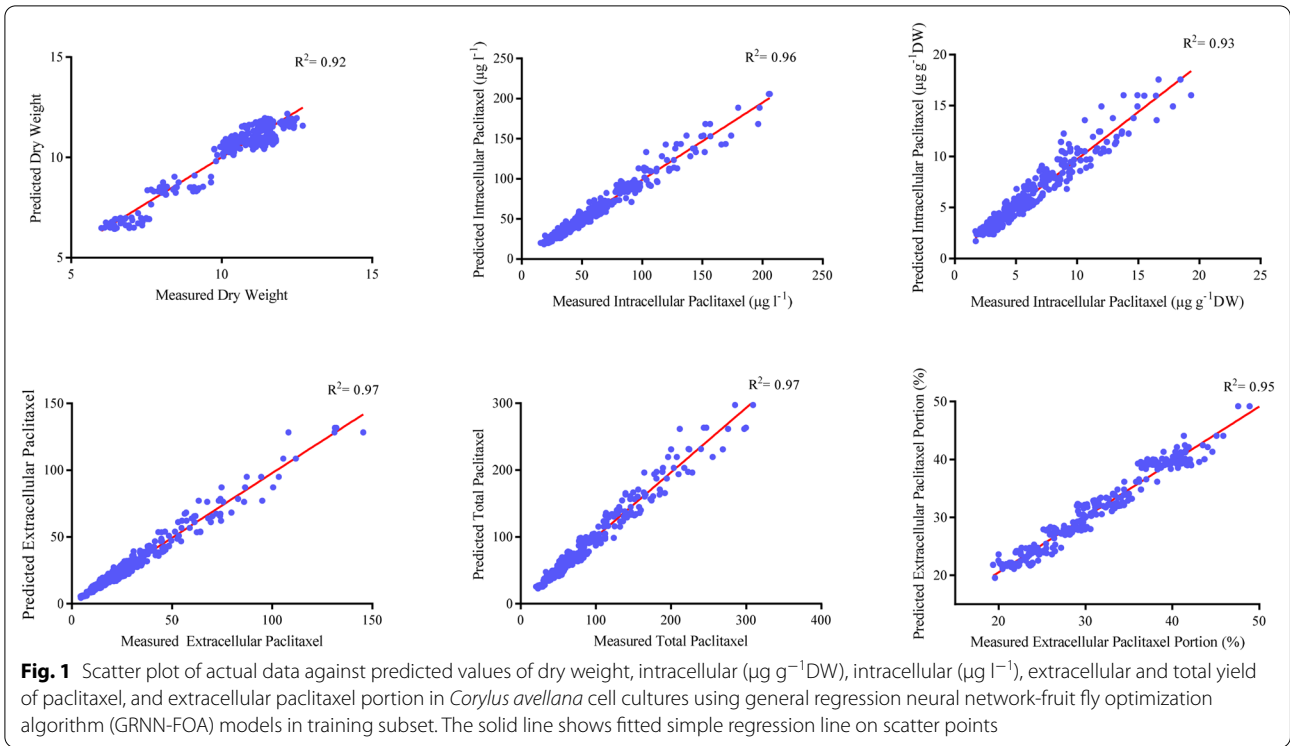
Smoothing (spread) parameter ( $\sigma$ ) in GRNN architecture has an important effect on predicting performance [41]. Indeed, the generalization capability of GRNN model depends on smoothing parameter. Intelligent optimization algorithms including fruit fly optimization algorithm (FOA) [42] was applied to determine parameters for predicting models.

Fruit Fly optimization algorithm or fly optimization algorithm (FOA) presented by Pan [43] is a new evolutionary optimization algorithm inspired from food finding behavior of fruit fly. The advantages of FOA are easy computational process, relatively simple and short program code and ease of understanding. So, this research attempted to apply FOA to automatically determine smoothing factor value of GRNN for enhancing predicting accuracy, and also optimize factors “CE and CF concentration levels, adding day of fungal elicitor and CSC harvesting time” for maximum paclitaxel biosynthesis and secretion in *C. avellana* cell culture treated with fungal elicitors.

## Results

### General regression neural network-fruit fly optimization analysis

Firstly, CE and CF concentration levels, elicitor adding day and CSC harvesting day were considered as input variables, and dry weight (DW), intracellular ( $\mu\text{g g}^{-1}$  DW), intracellular ( $\mu\text{g l}^{-1}$ ), extracellular and total yield of paclitaxel, and also extracellular paclitaxel portion as output variables. Afterwards, output variables were forecasted using developed GRNN-FOA models. The performance of developed GRNN-FOA models were evaluated by plotting the predicted values against the observed values of training (Fig. 1) and testing (Fig. 2) subsets. Great accordance between the predicted and observed values of DW, intracellular ( $\mu\text{g g}^{-1}$  DW), intracellular ( $\mu\text{g l}^{-1}$ ), extracellular and total yield of paclitaxel, and also extracellular paclitaxel portion was observed for both training and testing subset (Figs. 1, 2). Goodness-of-fit of developed GRNN-FOA models showed that they could accurately ( $R^2=0.88, 0.90, 0.91, 0.90, 90$  and  $0.88$ ) (Table 1) forecast DW, intracellular ( $\mu\text{g g}^{-1}$  DW), intracellular ( $\mu\text{g l}^{-1}$ ), extracellular and total yield of paclitaxel as well as extracellular paclitaxel portion of testing subset, respectively, not used during training processes (Fig. 2).



**Sensitivity analysis of models**

To rank input variables based on their relative importance in the model, variable sensitivity ratios (VSRs) were

estimated using entire data lines (training and testing subsets). VSRs were obtained for each output variables (DW, intracellular ( $\mu\text{g g}^{-1}\text{DW}$ ), intracellular ( $\mu\text{g l}^{-1}$ ),

**Table 1 Statistics and information on general regression neural network-fruit fly optimization algorithm (GRNN-FOA) models for growth, paclitaxel biosynthesis and secretion in *Corylus avellana* cell culture**

Measured factors	Training subsets			Test subsets		
	R <sup>2</sup>	RMSE	MBE	R <sup>2</sup>	RMSE	MBE
Dry weight (g l <sup>-1</sup> )	0.92	0.48	- 0.26 × 10 <sup>-15</sup>	0.88	0.81	0.03
Intracellular paclitaxel (μg l <sup>-1</sup> )	0.96	7.38	- 2.22 × 10 <sup>-15</sup>	0.90	10.99	0.65
Intracellular paclitaxel (μg g <sup>-1</sup> DW)	0.93	0.93	- 0.19 × 10 <sup>-15</sup>	0.91	1.37	0.01
Extracellular paclitaxel (μg l <sup>-1</sup> )	0.97	4.43	- 0.47 × 10 <sup>-15</sup>	0.90	6.93	- 0.09
Total yield of paclitaxel (μg l <sup>-1</sup> )	0.97	11.31	- 3.84 × 10 <sup>-15</sup>	0.90	16.93	0.55
Extracellular paclitaxel portion (%)	0.95	1.46	- 0.76 × 10 <sup>-15</sup>	0.88	2.88	- 0.23

R<sup>2</sup> coefficient of determination, RMSE root mean square error, MBE mean bias error

extracellular and total yield of paclitaxel, and also extracellular paclitaxel portion) regarding CE and CF concentration levels, elicitor adding day and CSC harvesting

time (Table 2). Analysis of DW model indicated that DW of *C. avellana* cells was more sensitive to CSC harvesting time (VSR = 1.002), followed by elicitor adding day

**Table 2 Importance (according to the sensitivity analysis) and optimal levels of the different factors including cell extract (CE), culture filtrate (CF) concentration levels [% (v/v)], fungal elicitor adding day and harvesting time (day) for achieving maximum growth, paclitaxel biosynthesis and its secretion in *Corylus avellana* cell culture by the optimization analysis using FOA and genetic algorithm (GA) on developed general regression neural network-fruit fly optimization (GRNN-FOA) models**

Criteria	Variable	Importance value (according to VSR <sup>a</sup> )	Optimal level		Output optimal	
			FOA	GA	FOA	GA
Dry weight (g l <sup>-1</sup> )	CE concentration level	0.0068	4.26	5.34	12.57	12.18
	CF concentration level	0.0053	0.54	0.71		
	Adding day	0.0237	16.33	15.62		
	Harvest time	1.0022	20.58	20.86		
Intracellular paclitaxel (μg g <sup>-1</sup> DW)	CE concentration level	0.8904	4.12	3.45	19.26	18.53
	CF concentration level	0.6751	5.84	5.68		
	Adding day	0.4257	15.72	16.97		
	Harvest time	0.2443	20.34	20.41		
Intracellular paclitaxel (μg l <sup>-1</sup> )	CE concentration level	0.7458	4.43	5.07	224.78	213.78
	CF concentration level	0.4406	5.69	5.46		
	Adding day	0.4083	16.09	16.79		
	Harvest time	0.3961	20.47	21.09		
Extracellular paclitaxel (μg l <sup>-1</sup> )	CE concentration level	0.2862	4.58	4.73	152.15	141.11
	CF concentration level	0.7519	4.71	5.06		
	Adding day	0.1893	15.91	16.19		
	Harvest time	0.9477	22.06	22.86		
Total yield of paclitaxel (μg l <sup>-1</sup> )	CE concentration level	0.6891	4.29	4.97	372.89	369.04
	CF concentration level	0.6043	5.38	5.01		
	Adding day	0.0943	17.00	16.53		
	Harvest time	0.2018	20.68	20.16		
Extracellular paclitaxel portion (%)	CE concentration level	0.1003	4.62	5.06	50.36	49.63
	CF concentration level	0.0622	4.91	4.97		
	Adding day	0.1409	16.59	17.03		
	Harvest time	0.4224	22.66	21.98		

<sup>a</sup> Relative indication of the ratio between the variable sensitivity error and the error of the model when all variables are available

(VSR=0.024), CE concentration level (VSR=0.007) and CF concentration level (VSR=0.005). Intracellular paclitaxel ( $\mu\text{g g}^{-1}$  DW) displayed more sensitivity to CE concentration level (VSR=0.890), followed by CF concentration level (VSR=0.675), elicitor adding day (VSR=0.426) and CSC harvesting time (VSR=0.244). Intracellular paclitaxel ( $\mu\text{g l}^{-1}$ ) exhibited more sensitivity to CE concentration level (VSR=0.746), followed by CF concentration level (VSR=0.441), elicitor adding day (VSR=0.408) and CSC harvesting time (VSR=0.396). Extracellular paclitaxel showed more sensitivity to CSC harvesting day (VSR=0.948), followed by CF concentration level (VSR=0.752), CE concentration level (VSR=0.286) and elicitor adding day (VSR=0.189). Accordingly, total yield of paclitaxel exhibited more sensitivity to CE concentration level (VSR=0.689), followed by CF concentration level (VSR=0.604), CSC harvesting time (VSR=0.202) and elicitor adding day (VSR=0.094). Also, extracellular paclitaxel portion displayed more sensitivity to CSC harvesting time (VSR=0.422), followed by elicitor adding day (VSR=0.141), CE concentration level (VSR=0.100) and CF concentration level (VSR=0.062) (Table 2).

### Model optimization

The optimization analysis on developed GRNN-FOA models was performed using fruit fly optimization algorithm to determine optimal levels of input variables for achieving maximum growth, paclitaxel biosynthesis and its secretion in *C. avellana* CSCs (Table 2). The optimization results showed that adding 4.8% (v/v) of CE:CF (89:11) containing 4.26% (v/v) CE and 0.54% (v/v) CF on 16th day, and harvesting CSC 102 h after elicitation could result in the maximum DW ( $12.57 \text{ g l}^{-1}$ ) (Table 2). The highest content of intracellular paclitaxel ( $19.26 \mu\text{g g}^{-1}$  DW) may produce by adding 9.96% (V/V) of CE:CF (41:59) containing 4.12% (v/v) CE and 5.84% (v/v) CF on 16th day, and harvesting CSC 110 h and 53 min after elicitation (Table 2). *C. avellana* cell culture exposed with 10.12% (v/v) of CE:CF (44:56) containing 4.43% (v/v) CE and 5.69% (v/v) CF on 16th day, and harvesting it 105 h and 7 min after elicitation may obtain the highest intracellular paclitaxel ( $224.78 \mu\text{g l}^{-1}$ ). Also, the results showed that the highest extracellular paclitaxel ( $152.15 \mu\text{g l}^{-1}$ ) can be produced by adding 9.29% (v/v) of CE:CF (49:51) containing 4.58% (v/v) CE and 4.71% (v/v) CF on 16th day, and harvesting CSC 147 h and 36 min after elicitation (Table 2). Additionally, CSC exposed with 9.67% (v/v) of CE:CF (44:56) containing 4.29% (v/v) CE and 5.38% (v/v) CF on 17th day, and harvesting it 88 h and 19 min after elicitation may obtain the highest total yield of paclitaxel ( $372.89 \mu\text{g l}^{-1}$ ) (Table 2). The results of GRNN-FOA model optimization displayed that adding

9.53% (v/v) of CE:CF (48:52) containing 4.62% (v/v) CE and 4.91% (v/v) CF on 17th day, and harvesting CSC 145 h and 41 min after elicitation may lead to the highest extracellular paclitaxel portion (50.36) (Table 2).

GRNN-FOA was also linked to genetic algorithm (GA) to determine the optimal level of input variables for achieving maximum growth, paclitaxel biosynthesis and its secretion in *C. avellana* CSCs (Table 2). The optimization results of paclitaxel biosynthesis in GRNN-FOA model using GA showed that adding 6.05% (v/v) of CE:CF (88:12) containing 5.34% (v/v) CE and 0.71% (v/v) CF on 16th day, and harvesting CSC 125 h and 46 min after elicitation could result in the maximum DW ( $12.18 \text{ g l}^{-1}$ ) (Table 2). Also, optimization results indicated that intracellular paclitaxel ( $18.53 \mu\text{g g}^{-1}$  DW) may produce by adding 9.13% (V/V) of CE:CF (38:62) containing 3.45% (v/v) CE and 5.68% (v/v) CF on 17th day, and harvesting CSC 82 h and 34 min after elicitation. *C. avellana* cell culture exposed with 10.53% (v/v) of CE:CF (48:52) containing 5.07% (v/v) CE and 5.46% (v/v) CF on 17th day, and harvesting it 103 h and 12 min after elicitation may obtain the highest total intracellular paclitaxel ( $213.78 \mu\text{g l}^{-1}$ ). Additionally, the results showed that the highest extracellular paclitaxel ( $141.11 \mu\text{g l}^{-1}$ ) can be produced by adding 9.79% (v/v) of CE:CF (48:52) containing 4.73% (v/v) CE and 5.06% (v/v) CF on 16th day, and harvesting CSC 160 h and 6 min after elicitation (Table 2). Also, cell culture exposed with 9.98% (v/v) of CE:CF (50:50) containing 4.97% (v/v) CE and 5.01% (v/v) CF on 17th day, and harvesting it 87 h and 7 min after elicitation may obtain the highest total yield of paclitaxel ( $369.04 \mu\text{g l}^{-1}$ ) (Table 2). The results of optimizing GRNN-FOA model using GA showed that adding 10.03% (v/v) of CE:CF (50:50) containing 5.06% (v/v) CE and 4.97% (v/v) CF on 17th day, and harvesting CSC 118 h and 48 min after elicitation may lead to the highest extracellular paclitaxel portion (49.63) (Table 2).

### Validation experiment

*C. avellana* cell culture exposed to 4.29% (v/v) CE and 5.38% (v/v) CF on 17th day, and harvesting it 88 h after elicitation (optimized input variables in GRNN-FOA model using FOA) produced  $348.65 \pm 36.8 \mu\text{g l}^{-1}$  paclitaxel.

### Discussion

Paclitaxel biosynthesis in *C. avellana* CSC treated with fungal elicitors is affected by the type, concentration level and adding day of fungal elicitors and also CSC harvesting time [6, 7, 10–13]. Forecasting the optimized value of these mentioned factors is highly promising and essential for paclitaxel biosynthesis improvement. However, the optimization of these factors by experimental studies is

laborious, time-consuming, and costly. Paclitaxel biosynthesis is considered as complex biological process since it is affected by multiple factors in nonlinear ways [7, 13]. Therefore, the conventional computational methods are inefficient for modeling paclitaxel biosynthesis [7, 12, 13]. Some machine learning algorithms such as multi-layer perceptron [13], genetic algorithm [7, 13], adaptive neuro-fuzzy inference system [13] have been successfully used for forecasting and optimizing paclitaxel biosynthesis. This is the first study for forecasting the optimal conditions for maximum paclitaxel biosynthesis in *C. avellana* CSC exposed to fungal elicitors using GRNN-FOA model. To accurately forecast the optimized values of effective factors (CE and CF concentration levels, elicitor adding day and CSC harvesting time) on paclitaxel biosynthesis in *C. avellana* CSC, using a trustworthy modeling system is essential.

In this study, GRNN-FOA modeling was used to evaluate the relationships among four studied factors “CE and CF concentration levels, elicitor adding time and CSC harvesting time” and the parameters “DW, intracellular, extracellular and total yield of paclitaxel and extracellular paclitaxel portion”, and also the possibility of forecasting of paclitaxel biosynthesis by the determined factors. Such mathematical predictions using GRNN-FOA model have not been described in this area.

Our results suggested that GRNN-FOA models could accurately forecast DW, intracellular paclitaxel ( $\mu\text{g g}^{-1}$  DW), intracellular paclitaxel ( $\mu\text{g l}^{-1}$ ), extracellular paclitaxel, total yield of paclitaxel and extracellular paclitaxel portion ( $R^2=0.88, 0.90, 0.91, 0.90, 0.90$  and  $88$ , respectively) in testing subset (Fig. 1), not used in training process. Small bias values (Table 2) showed the high potential of GRNN-FOA models in forecasting output variables.

It is noteworthy that our group was previously used multivariate statistical methods including “stepwise regression, ordinary least squares regression, principal component regression and partial least squares regression [12]. Goodness-of-fit showed no difference regarding the accuracy of different regression models for all output variables, 0.67, 0.57, 0.62, 0.60 and 0.86 for DW, intracellular paclitaxel, extracellular paclitaxel, total yield of paclitaxel and extracellular paclitaxel portion, respectively for training subset [12]. The fit of regression models was presented by  $R^2$  for testing subset, suggesting the best-mentioned regression models can explain 67, 62, 68, 65 and 86% of the variability in DW, intracellular paclitaxel, extracellular paclitaxel, total yield of paclitaxel and paclitaxel extracellular portion, respectively, when they faced unseen data [12]. As shown in Table 1, the statistical values for GRNN-FOA models displayed higher prediction accuracy than

regression models in previous study [12]. This finding was in line with the previous studies [7, 13] showing AI technology had the superior performances as compared to conventional modeling methods for forecasting growth and paclitaxel biosynthesis in *C. avellana* cell culture.

Additionally, multilayer perceptron-genetic algorithm (MLP-GA) was used to forecast growth and paclitaxel biosynthesis in *C. avellana* CSC treated with fungal elicitors [13]. Comparative analysis of MLP-GA [13] and GRNN-FOA (Table 1) showed very slight difference between two models for DW, intracellular and extracellular paclitaxel in testing subset, the unseen data. However, MLP-GA was slightly more accurate as compared to GRNN-FOA for total paclitaxel and extracellular paclitaxel portion in testing subset.  $R^2$  for GRNN-FOA (Table 1) vs. MLP-GA [13] were; DW = 0.89 vs. 0.90, intracellular paclitaxel = 0.90 vs. 0.89, extracellular paclitaxel = 0.90 vs. 0.92, total yield of paclitaxel = 0.90 vs. 0.95, and extracellular paclitaxel portion = 0.88 vs. 0.91.

As shown in Fig. 3, residual plots for all the developed GRNN-FOA models displayed a high density of points close to the origin and a low density of points away from the origin, and symmetric shape about the origin. Indeed, the residuals appear to behave randomly (normal distribution), it suggests that developed GRNN-FOA models for forecasting DW, intracellular paclitaxel ( $\mu\text{g g}^{-1}$  DW), intracellular paclitaxel ( $\mu\text{g l}^{-1}$ ), extracellular paclitaxel, total yield of paclitaxel and extracellular paclitaxel portion fit the data well.

The results of optimization analysis using “GA” and “FOA” on developed GRNN-FOA models displayed the slight difference in maximum growth and paclitaxel biosynthesis optimized by these optimization algorithms.

As previously mentioned, sensitivity analysis displayed that CE and CF concentration levels are the most important variables affecting total yield of paclitaxel (Table 2). Accordingly, CSC harvesting time and CF concentration level had the greatest effect on extracellular paclitaxel content (Table 2). The increment of paclitaxel secretion from the cells to culture medium decrease toxicity and feedback inhibition of paclitaxel [6, 13]. Paclitaxel secretion to culture medium undoubtedly makes easy extraction and the purification of it which is required for steady production of paclitaxel at the commercial level. Extracellular paclitaxel content is important for paclitaxel biosynthesis in continuous system. Sensitivity analysis displayed that CSC harvesting time is the most important factors affecting extracellular paclitaxel (Table 2). Paclitaxel biosynthesis is the complex biological processes which require the accurate techniques for modeling and optimization. GRNN-FOA has been efficiently used to

solve problems with extremely difficult and unknown solution in various fields [40, 44–47].

Based on high forecasting accuracy of training and testing subsets (Figs. 1, 2) and also residual analysis (Fig. 3), it can be concluded that developed GRNN-FOA could precisely forecast DW, paclitaxel biosynthesis and secretion in *C. avellana* CSC. Additionally, the validation experiment revealed that GRNN-FOA hybrid method is an efficient method for forecasting and optimizing paclitaxel biosynthesis in *C. avellana* cell culture responding fungal elicitors.

In conclusion, this research applied GRNN-FOA for forecasting and optimizing paclitaxel biosynthesis in *C. avellana* cell culture treated with fungal elicitors for the first time. Great accordance between the predicted and observed values of DW, intracellular, extracellular and total yield of paclitaxel, and also extracellular paclitaxel portion support excellent performance of developed GRNN-FOA models. This research introduced GRNN-FOA as a new mathematical tool for forecasting and optimizing the complex systems including secondary metabolite biosynthesis in plant in vitro culture, paclitaxel biosynthesis in *C. avellana* CSC responding to fungal elicitors as a case study. Overall, GRNN-FOA could be useful as a strong method for forecasting and optimizing in various fields of plant systems.

## Methods

### Cell suspension culture

*C. avellana* CSC was established as described by Salehi et al. [8–11].

### Preparation of elicitors and elicitation experiment

Endophytic fungus applied in this research was a strain of *Camarosporomyces flavigenus*, HEF<sub>17</sub>, isolated from the leaf of *C. avellana* grown in Iran [13]. CE and CF were prepared as described previously [10]. For elicitation, 1.5 ± 0.1 g of *C. avellana* cells (fresh mass) was cultured in 100 ml flasks containing 30 ml of Murashige and Skoog (MS) medium supplemented with 2 mg l<sup>-1</sup> 2,4-D and 0.2 mg l<sup>-1</sup> BAP.

Three concentrations [2.5, 5 and 10% (v/v)] of fungal elicitors “CE:CF (100:0, 75:25, 50:50, 25:75, 0:100 v/v)”, and also mid (day 13) and late (day 17) log phase of *C. avellana* cell cultures were selected for adding fungal elicitors. Control received an equal volume of water (for CE)/potato dextrose broth (PDB) (for CF).

### Cell growth measurement

Cell dry weight (DW) was measured as cell growth [6–13].

### Quantification of paclitaxel

The extraction of intracellular and extracellular paclitaxel, and also HPLC analysis were performed with a procedure described by Salehi et al. [8–11].

### Experimental design

The experiment was planned based on completely randomized design (CRD) with factorial arrangement, three factors containing fungal elicitor type with 10 levels [(CE:CF (100:0, 75:25, 50:50, 25:75, 0:100 v/v) and water:PDB (100:0, 75:25, 50:50, 25:75, 0:100 v/v), elicitor concentration with three levels (2.5, 5, and 10% (v/v)), elicitor adding day with two levels (days 13 and 17), and three replicates. The cultures were harvested in two-day intervals after elicitation until 23rd day.

### Model development

Before testing machine learning algorithm, Box-Cox transformation [48] was used for normalizing the datasets. Also, principal component analysis (PCA) was applied to detect outliers; however, no outlier was detected in this case.

Five-fold cross-validation method with ten repetitions were used to calculate the prediction accuracy of all the tested models. Thus, we found the model with the best prediction on unknown data from the entire data set. The advantages of K-fold cross-validation are low computation time, low bias, every data dataset is used for both training (k – 1) and testing (1) subset.

### General regression neural network (GRNN) model

GRNN modeling was used to define the influences of CE and CF concentration levels, elicitor adding day and harvesting day on DW, paclitaxel biosynthesis (intracellular, extracellular and total) and extracellular paclitaxel portion.

GRNN is established on a standard statistical method named Gaussian kernel regression [49]. As shown in Fig. 4, GRNN is made up of four layers including input, pattern, summation and output layers. Input layer (distribution unit) stores information as an input vector  $X$ , and is totally connected to pattern layer. The neurons of input layer, input neurons, feed input variables to all neurons on second layer (pattern unit). Pattern layer applies a non-linear transformation from input space to pattern one. Pattern neurons, the neurons in pattern layer, memorize the relation among input neuron and the proper response of pattern layer. Pattern Gaussian function “pi” given in Eq. (1) is applied to compute an output  $\pi_i$  by a pattern neuron  $i$ .

$$p_i = \exp\left[-\frac{(x - x_i)^T(x - x_i)}{2\sigma^2}\right] (i = 1, 2, \dots, n), \quad (1)$$

where  $X$  denotes input variable,  $X_i$  is a specific training vector of pattern neuron  $i$ , and  $\sigma$  signifies smoothing parameter.

Summation neurons, the neurons in summation layer, pass on the outputs of pattern unit to third layer, summation unit. Third layer has two summations including simple summation ( $S_s$ ) and weighted summation ( $S_w$ ) while  $S_s$  (Eq. 2) computes the summation of all pattern layer outputs.  $S_w$  (Eq. 3) computes weighted sum of pattern layer outputs, where  $w_i$  is interconnection weight of pattern neuron  $i$  to summation layer.

$$S_s = \sum_{i=1} p_i. \quad (2)$$

$$S_w = \sum_{i=1} w_i p_i. \quad (3)$$

Then, summation layer feed both  $S_s$  (numerator) and  $S_w$  (denominator) to output layer. Output layer computes output  $Y$  of GRNN model by dividing summation layer outputs (Eq. 4).

$$Y = S_s/S_w. \quad (4)$$

Smoothing parameter “ $\sigma$ ” is only parameter that needs to be defined in GRNN model. This research applied fruit fly optimization algorithm (FOA) to automatically determine appropriate smoothing parameter value in GRNN model.

### Fruit fly optimization algorithm (FOA)

FOA was used (1) to determine appropriate value of smooth parameter ( $\sigma$ ), and (2) to optimize the values of input variables (CE and CF concentration, elicitor elicitor adding day and CSC harvesting day) in developed GRNN-FOA models for maximum paclitaxel biosynthesis and its secretion.

FOA is a new intelligence method inspired from food searching behavior of fruit fly which can find global optimal solution [43]. Food searching process of fruit fly includes two steps: (1) fruit fly detects the food location using osphresis organ and flies towards it, (2) when fruit fly gets close to the food source, the sensitive vision is likewise applied for detecting source and fruit flies flocking location, and fly towards that direction. Food finding iterative behavior of fruit fly group is presented in Fig. 5.

The procedure of FOA for detecting the optimal values is described as follows.

Step 1. Randomly initialize FOA parameters including population size (sizepop), maximum iteration number (maxgen), location coordinate (LC) ( $X, Y$ ) of fruit fly group, and flight distance range (FDR).

Step 2. Give the random distance and direction (Eq. 5) to an individual fruit fly such that they can detect the food by osphresis organ.

$$X_i = X + \text{Random Value}, \quad (5)$$

$$Y_i = Y + \text{Random Value}.$$

Step 3. Compute the distance of food location to the origin ( $Dist$ ) (Eq. 6), smell concentration judgment value ( $S_i$ ) (Eq. 7), and smell concentration ( $Smell_i$ ) of individual fruit fly location by putting smell concentration judgment value ( $S_i$ ) into the smell concentration judgment function (fitness function) (Eq. 8). At last, determine the fruit fly with highest smell concentration (highest  $Smell_i$  value) (Eq. 9) among the fruit fly group:

$$Dist_i = \sqrt{(X_i^2 + Y_i^2)}. \quad (6)$$

$$S_i = 1/Dist_i. \quad (7)$$

$$Smell_i = \text{Function}(S_i). \quad (8)$$

$$[bestSmell \ bestIndex] = \max(Smell_i). \quad (9)$$

Step 4. Keep the highest smell concentration value (Eq. 10), and find fly location coordinate with highest smell concentration value (Eq. 11), and at this point, fruit fly group flies towards that location using vision. Enter iterative optimization until (1) current iteration numbers is less than maxgen (2) highest smell concentration is superior as compared to previous iterative one.

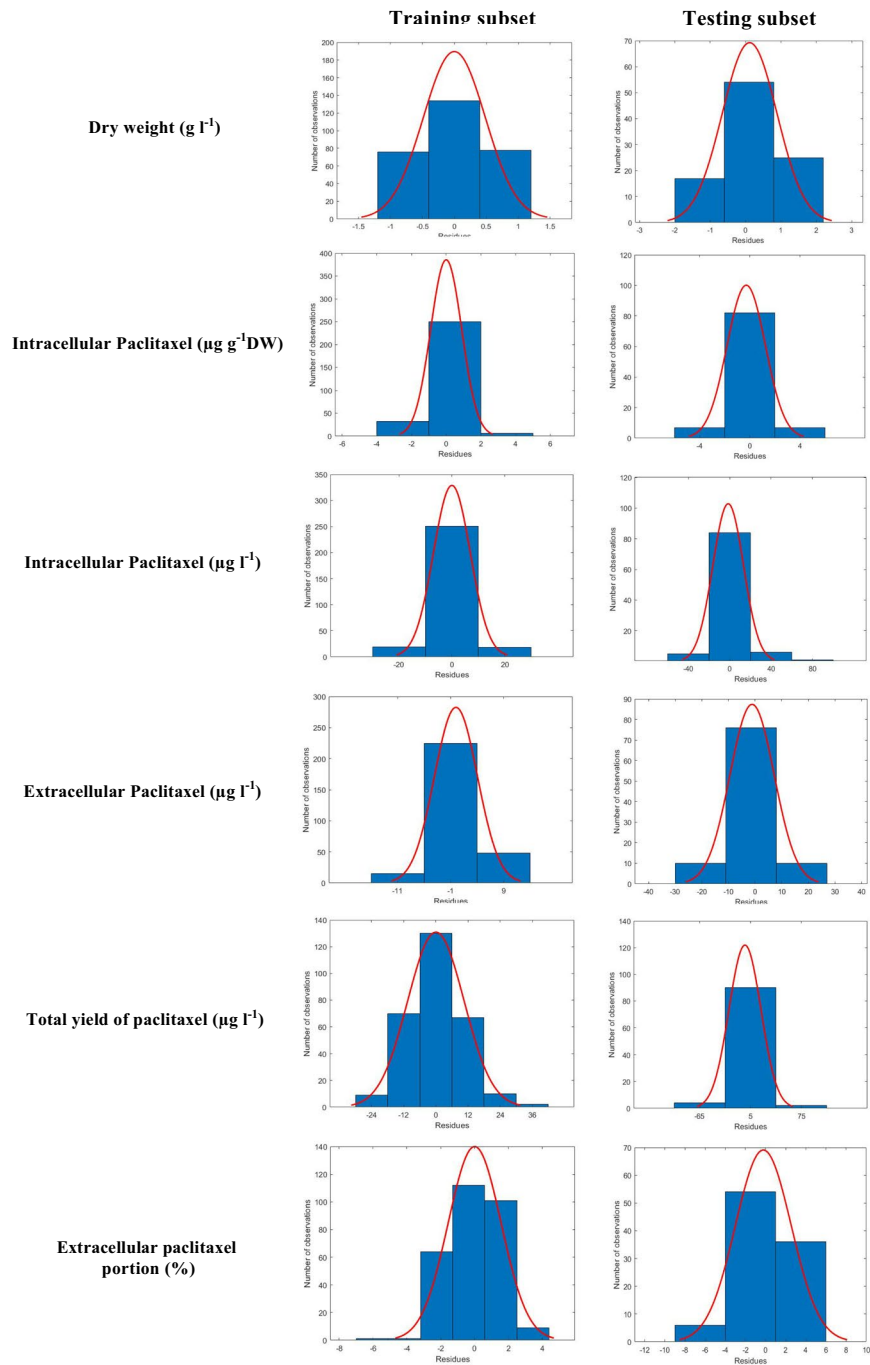
$$Smell_{best} = bestSmell. \quad (10)$$

$$X = X(bestIndex). \quad (11)$$

$$Y = Y(bestIndex).$$

The optimization procedure for searching appropriate value of smoothing parameter in GRNN model, and also optimal input variables for maximum paclitaxel biosynthesis through FOA in GRNN-FOA model is presented in Fig. 6. Maxgen of 100, sizepop of 10, LC of [0, 1] and FDR





**Fig. 3** Histogram of residuals for general regression neural network-fruit fly optimization algorithm (GRNN-FOA) models developed for forecasting dry weight, intracellular ( $\mu g g^{-1}DW$ ), intracellular ( $\mu g l^{-1}$ ), extracellular and total yield of paclitaxel and extracellular paclitaxel portion in *Corylus avellana* cell cultures treated with fungal elicitors

of  $[-10, 10]$  [40] were set to establish the fittest GRNN structure, and also optimize input variables for maximum paclitaxel biosynthesis in GRNN-FOA model.

The performance of GRNN-FOA models is determined by three statistical criteria including root mean square error (RMSE) (Eq. 12), mean bias error (MBE) (Eq. 13) and coefficient of determination ( $R^2$ ) (Eq. 14).

$$RMSE = \sqrt{\left(\sum_{i=1}^n (y_{est} - y_{act})^2\right) / n}, \tag{12}$$

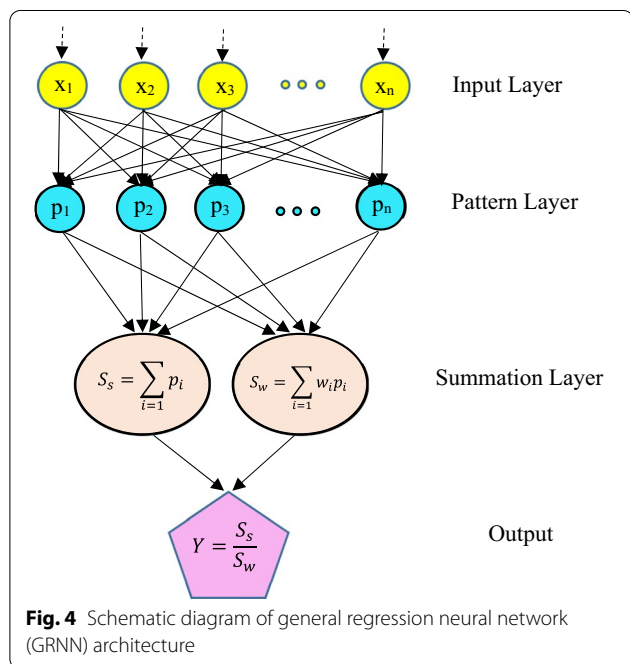
$$MBE = 1/n \sum_{i=1}^n (y_{est} - y_{act}), \tag{13}$$

$$R^2 = 1 - \left(\sum_{i=1}^n (y_{est} - y_{act})^2 / \sum_{i=1}^n (y_{act} - \bar{y})^2\right), \tag{14}$$

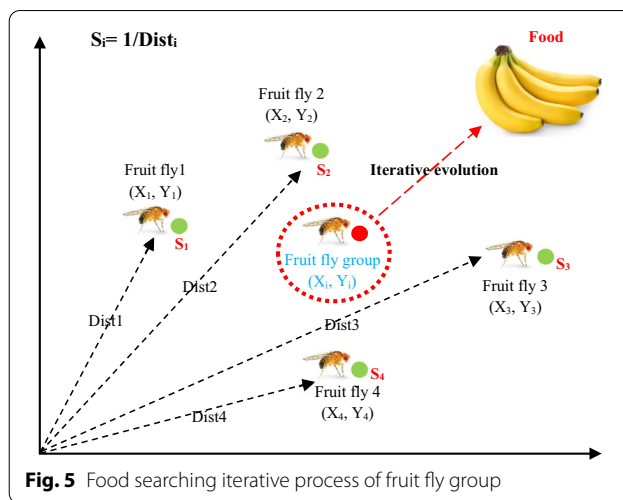
where “ $y_{act}$ ” are the actual values, “ $y_{est}$ ” are the predicted values, and “ $n$ ” is the number of data.

**Sensitivity analysis of the models**

Sensitivity analysis was done on GRNN-FOA models to determine the importance degree of the factors (CE and CF concentration levels, elicitor adding day and



**Fig. 4** Schematic diagram of general regression neural network (GRNN) architecture



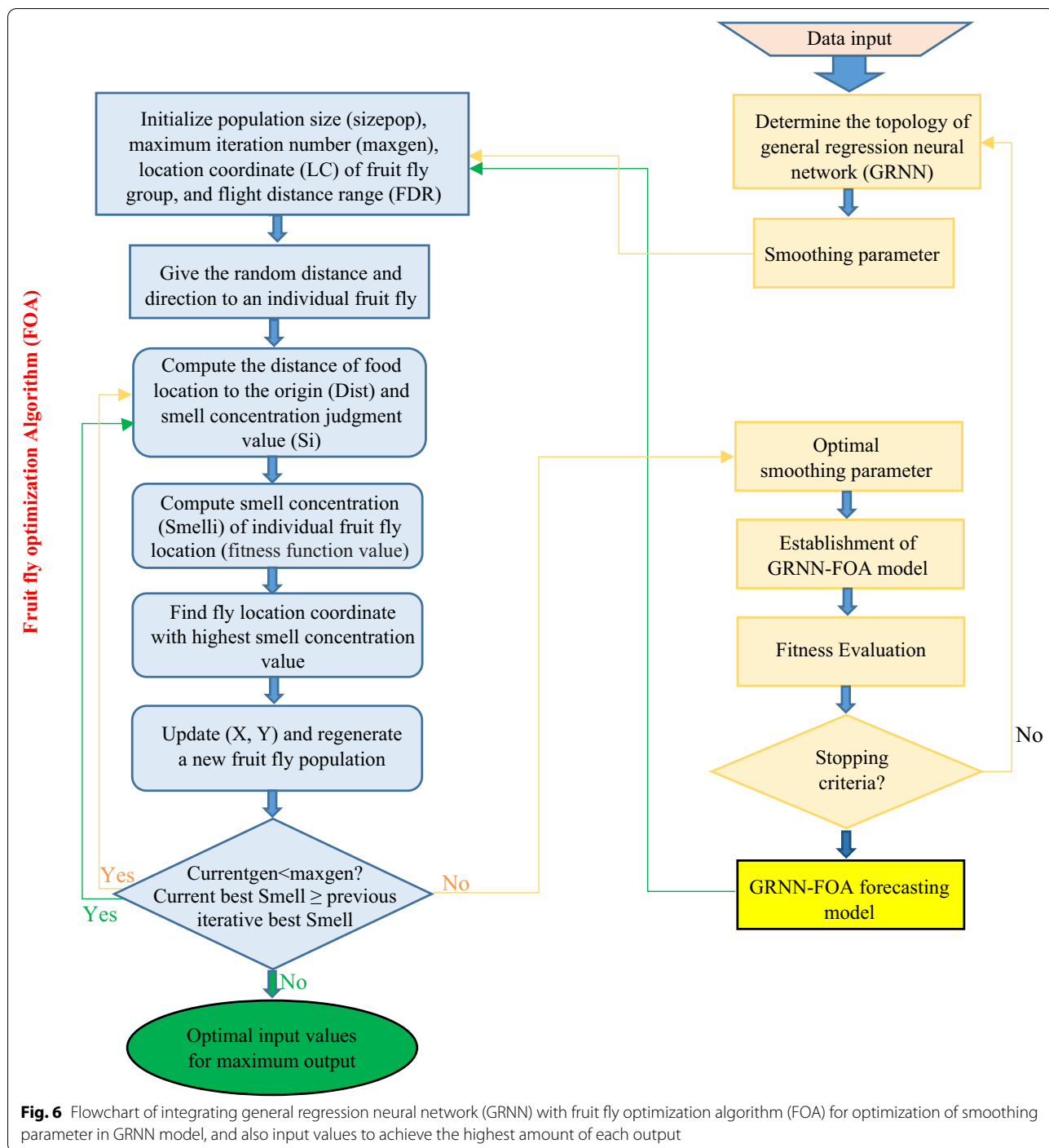
**Fig. 5** Food searching iterative process of fruit fly group

harvesting time) on the model parameters (DW, paclitaxel biosynthesis and its secretion). The sensitivity of DW, paclitaxel biosynthesis (intracellular, extracellular and total yield) and extracellular paclitaxel portion was determined by the criteria including variable sensitivity error (VSE) value displaying the performance (RMSE) of GRNN-FOA model when that particular input variable is unavailable from the model. Variable sensitivity ratio (VSR) value was calculated as ratio of VSE and GRNN-FOA model error (RMSE value) when all input variables are available. The input variable with higher VSR was considered as higher important variable in model [7, 13, 50–52]. Finally, calculated VSR values were rescaled within range [0, 1] to make them more easily comparable.

The mathematical codes for the development and evaluation of GRNN-FOA and GRNN-FOA-GA models were written using MATLAB [53] software, and the graphs were made by GraphPad Prism 5 [54] software.

**Validation experiment**

CE and CF concentration levels, elicitor adding day, and harvesting time of CSC optimized by FOA were tested to evaluate the efficiency of GRNN-FOA model for forecasting and optimizing paclitaxel biosynthesis in *C. avellana* cell culture responding to fungal elicitors.



**Abbreviations**

CE: Cell extract; CF: Culture filtrate; CSC: Cell suspension culture; AI: Artificial intelligence; ANN: Artificial neural network; GRNN: General regression neural network; FOA: Fruit fly optimization algorithm; GA: Genetic algorithm; MLP: Multilayer perceptron; PDB: Potato dextrose broth; DW: Dry weight; VSR: Variable sensitivity ratio.

**Authors' contributions**

MS carried out all experiments and analyses. MS and SF interpreted the results and wrote the manuscript. AM and NS directed the research. MH performed data modeling. All authors read and approved the final manuscript.

**Funding**

No funding was received.

**Availability of data and materials**

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup> Department of Plant Genetics and Breeding, Faculty of Agriculture, Tarbiat Modares University, P.O. Box 14115-336, Tehran, Iran. <sup>2</sup> Department of Plant Pathology, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran. <sup>3</sup> Gosling Research Institute for Plant Preservation, Department of Plant Agriculture, University of Guelph, Guelph, ON, Canada.

Received: 28 May 2020 Accepted: 22 January 2021

Published online: 05 February 2021

**References**

- Wani MC, Taylor HL, Wall ME, Coggon P, McPhail AT. Plant antitumor agents. VI. Isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J Am Chem Soc*. 1971;93(9):2325–7. <https://doi.org/10.1021/ja00738a045>.
- Weaver BA. How Taxol/paclitaxel kills cancer cells. *Mol Biol Cell*. 2014;25(18):2677–81. <https://doi.org/10.1091/mbc.E14-04-0916>.
- Jordan MA, Wilson L. Microtubules as a target for anticancer drugs. *Nat Rev Cancer*. 2004;4(4):253–65. <https://doi.org/10.1038/nrc1317>.
- Schiff PB, Fant J, Horwitz SB. Promotion of microtubule assembly in vitro by taxol. *Nature*. 1979;277(5698):665–7. <https://doi.org/10.1038/277665a0>.
- Gallego A, Malik S, Yousefzadi M, Makhzoum A, Tremouillaux-Guiller J, Bonfill M. Taxol from *Corylus avellana*: paving the way for a new source of this anti-cancer drug. *Plant Cell Tissue Organ Cult*. 2017;129(1):1–16. <https://doi.org/10.1007/s11240-017-1175-x>.
- Farhadi S, Moieni A, Safaie N, Sabet MS, Salehi M. Fungal cell wall and methyl- $\beta$ -cyclodextrin synergistically enhance paclitaxel biosynthesis and secretion in *Corylus avellana* cell suspension culture. *Sci Rep*. 2020;10(1):1–10. <https://doi.org/10.1038/s41598-020-62196-4>.
- Farhadi S, Salehi M, Moieni A, Safaie N, Sabet MS. Modeling of paclitaxel biosynthesis elicitation in *Corylus avellana* cell culture using adaptive neuro-fuzzy inference system-genetic algorithm (ANFIS-GA) and multiple regression methods. *PLoS ONE*. 2020;15(8):e0237478. <https://doi.org/10.1371/journal.pone.0237478>.
- Salehi M, Moieni A, Safaie N. A novel medium for enhancing callus growth of hazel (*Corylus avellana* L.). *Sci Rep*. 2017;7(1):1–9. <https://doi.org/10.1038/s41598-017-15703-z>.
- Salehi M, Moieni A, Safaie N. Elicitors derived from hazel (*Corylus avellana* L.) cell suspension culture enhance growth and paclitaxel production of *Epicoccum nigrum*. *Sci Rep*. 2018;8(1):1–10. <https://doi.org/10.1038/s41598-018-29762-3>.
- Salehi M, Moieni A, Safaie N, Farhadi S. Elicitors derived from endophytic fungi *Chaetomium globosum* and *Paraconiothyrium brasiliense* enhance paclitaxel production in *Corylus avellana* cell suspension culture. *Plant Cell Tissue Organ Cult*. 2019;136(1):161–71. <https://doi.org/10.1007/s11240-018-1503-9>.
- Salehi M, Moieni A, Safaie N, Farhadi S. New synergistic co-culture of *Corylus avellana* cells and *Epicoccum nigrum* for paclitaxel production. *J Ind Microbiol Biotechnol*. 2019;46(5):613–23. <https://doi.org/10.1007/s10295-019-02148-8>.
- Salehi M, Moieni A, Safaie N, Farhadi S. Whole fungal elicitors boost paclitaxel biosynthesis induction in *Corylus avellana* cell culture. *PLoS ONE*. 2020;15(7):e0236191. <https://doi.org/10.1371/journal.pone.0236191>.
- Salehi M, Farhadi S, Moieni A, Safaie N, Ahmadi H. Mathematical modeling of growth and paclitaxel biosynthesis in *Corylus avellana* cell culture responding to fungal elicitors using multilayer perceptron-genetic algorithm. *Front Plant Sci*. 2020; 11. <https://doi.org/10.3389/fpls.2020.01148>.
- Miele M, Mumot AM, Zappa A, Romano P, Ottaggio L. Hazel and other sources of paclitaxel and related compounds. *Phytochem Rev*. 2012;11(2–3):211–25. <https://doi.org/10.1007/s11101-012-9234-8>.
- Smetanska I. Production of secondary metabolites using plant cell cultures. In: *Food biotechnology*. Springer, Berlin; 2008: 187–228. [https://doi.org/10.1007/10\\_2008\\_103](https://doi.org/10.1007/10_2008_103).
- Salehi M, Karimzadeh G, Naghavi MR. Synergistic effect of coronatine and sorbitol on artemisinin production in cell suspension culture of *Artemisia annua* L. cv, Anamed. *Plant Cell Tissue Organ Cult*. 2019;137(3):587–97. <https://doi.org/10.1007/s11240-019-01593-8>.
- Salehi M, Karimzadeh G, Naghavi MR, Badi HN, Monfared SR. Expression of artemisinin biosynthesis and trichome formation genes in five *Artemisia* species. *Ind Crop Prod*. 2018;112:130–40. <https://doi.org/10.1016/j.indcrop.2017.11.002>.
- Salehi M, Karimzadeh G, Naghavi MR, Badi HN, Monfared SR. Expression of key genes affecting artemisinin content in five *Artemisia* species. *Sci Rep*. 2018;8(1):1–11. <https://doi.org/10.1038/s41598-018-31079-0>.
- Salehi M, Naghavi MR, Bahmankar M. A review of *Ferula* species: biochemical characteristics, pharmaceutical and industrial applications, and suggestions for biotechnologists. *Ind Crop Prod*. 2019;139:111511. <https://doi.org/10.1016/j.indcrop.2019.111511>.
- Gallego PP, Gago J, Landín M. Artificial neural networks technology to model and predict plant biology process artificial neural networks-methodological advances and biomedical applications Rijeka, Croatia. *Intech Open Access Publ* 2011:197–217. <https://doi.org/10.5772/14945>.
- Struik PC, Yin X, de Visser P. Complex quality traits: now time to model. *Trends Plant Sci*. 2005;10(11):513–6. <https://doi.org/10.1016/j.tplants.2005.09.005>.
- Gago J, Martínez-Núñez L, Landín M, Gallego P. Artificial neural networks as an alternative to the traditional statistical methodology in plant research. *J Plant Physiol*. 2010;167(1):23–7. <https://doi.org/10.1016/j.jplph.2009.07.007>.
- Nezami-Alanagh E, Garoosi G-A, Landín M, Gallego PP. Combining DOE with neurofuzzy logic for healthy mineral nutrition of pistachio rootstocks in vitro culture. *Front Plant Sci*. 2018;9:1474. <https://doi.org/10.3389/fpls.2018.01474>.
- Patnaik P. Applications of neural networks to recovery of biological products. *Biotechnol Adv*. 1999;17(6):477–88. [https://doi.org/10.1016/S0734-9750\(99\)00013-0](https://doi.org/10.1016/S0734-9750(99)00013-0).
- Hesami M, Condori-Apfata JA, Valderrama Valencia M, Mohammadi M. Application of artificial neural network for modeling and studying in vitro genotype-independent shoot regeneration in wheat. *Appl Sci*. 2020;10(15):5370. <https://doi.org/10.3390/app10155370>.
- Hesami M, Naderi R, Tohidfar M, Yousefzadeh-Najafabadi M. Development of support vector machine-based model and comparative analysis with artificial neural network for modeling the plant tissue culture procedures: effect of plant growth regulators on somatic embryogenesis of chrysanthemum, as a case study. *Plant Methods*. 2020;16(1):112. <https://doi.org/10.1186/s13007-020-00655-9>.
- Hesami M, Alizadeh M, Naderi R, Tohidfar M. Forecasting and optimizing Agrobacterium-mediated genetic transformation via ensemble model-fruit fly optimization algorithm: a data mining approach using chrysanthemum databases. *PLoS ONE*. 2020;15(9):e0239901. <https://doi.org/10.1371/journal.pone.0239901>.
- Hesami M, Jones AMP. Application of artificial intelligence models and optimization algorithms in plant cell and tissue culture. *Appl Microbiol Biotechnol*. 2020;104:9449–85. <https://doi.org/10.1007/s00253-020-10888-2>.
- Hesami M, Naderi R, Tohidfar M. Introducing a hybrid artificial intelligence method for high-throughput modeling and optimizing plant tissue culture processes: the establishment of a new embryogenesis medium for chrysanthemum, as a case study. *Appl Microbiol Biotechnol*. 2020;104:10249–63. <https://doi.org/10.1007/s00253-020-10978-1>.
- Yousefzadeh Najafabadi M, Earl HJ, Tulpan D, Sulik J, Eskandari M. Application of machine learning algorithms in plant breeding: predicting yield from hyperspectral reflectance in soybean. *Front Plant Sci*. 2020;11:2169. <https://doi.org/10.3389/fpls.2020.624273>.
- Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research.

- J Pharmaceut Biomed. 2000;22(5):717–27. [https://doi.org/10.1016/S0731-7085\(99\)00272-1](https://doi.org/10.1016/S0731-7085(99)00272-1).
32. Specht DF. A general regression neural network. *IEEE Trans Neural Netw.* 1991;2(6):568–76. <https://doi.org/10.1109/72.97934>.
  33. Kulkarni SG, Chaudhary AK, Nandi S, Tambe SS, Kulkarni BD. Modeling and monitoring of batch processes using principal component analysis (PCA) assisted generalized regression neural networks (GRNN). *BiochemEng J.* 2004;18(3):193–210. <https://doi.org/10.1016/j.bej.2003.08.009>.
  34. Chen T-C, Yu C-H. Motion control with deadzone estimation and compensation using GRNN for TWUSM drive system. *SystAppl.* 2009;36(8):10931–41. <https://doi.org/10.1016/j.eswa.2009.02.025>.
  35. Shahlaei M, Sabet R, Ziari MB, Moenifard B, Fassihi A, Karbaksh R. QSAR study of anthranilic acid sulfonamides as inhibitors of methionine aminopeptidase-2 using LS-SVM and GRNN based on principal components. *Eur J Medic Chem.* 2010;45(10):4499–508. <https://doi.org/10.1016/j.ejmech.2010.07.010>.
  36. Chelgani SC, Jorjani E. Microwave irradiation pretreatment and peroxyacetic acid desulfurization of coal and application of GRNN simultaneous predictor. *Fuel.* 2011;90(11):3156–63. <https://doi.org/10.1016/j.fuel.2011.06.045>.
  37. Chang P-C, Liu C-H, Fan C-Y. Data clustering and fuzzy neural network for sales forecasting: a case study in printed circuit board industry. *Knowl-Based Syst.* 2009;22(5):344–55. <https://doi.org/10.1016/j.knosys.2009.02.005>.
  38. Guo Z-h, Wu J, Lu H-y, Wang J-z. A case study on a hybrid wind speed forecasting method using BP neural network. *Knowl-Based Syst.* 2011;24(7):1048–56. <https://doi.org/10.1016/j.knosys.2011.04.019>.
  39. Leung MT, Chen A-S, Daouk H. Forecasting exchange rates using general regression neural networks. *ComputOper Res.* 2000;27(11–12):1093–110. [https://doi.org/10.1016/S0305-0548\(99\)00144-6](https://doi.org/10.1016/S0305-0548(99)00144-6).
  40. Li H-Z, Guo S, Li C-J, Sun J-Q. A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. *Knowl-Based Syst.* 2013;37:378–87. <https://doi.org/10.1016/j.knosys.2012.08.015>.
  41. Zhang Y, Niu J, Na S. A novel nonlinear function fitting model based on FOA and GRNN. *Math ProblEng.* 2019. <https://doi.org/10.1155/2019/2697317>.
  42. Wang L, Zheng X-L, Wang S-Y. A novel binary fruit fly optimization algorithm for solving the multidimensional knapsack problem. *Knowl-Based Syst.* 2013;48:17–23. <https://doi.org/10.1016/j.knosys.2013.04.003>.
  43. Pan W-T. A new fruit fly optimization algorithm: taking the financial distress model as an example. *Knowl-Based Syst.* 2012;26:69–74. <https://doi.org/10.1016/j.knosys.2011.07.001>.
  44. Pan W-T. Using modified fruit fly optimisation algorithm to perform the function test and case studies. *Connect Sci.* 2013;25(2–3):151–60. <https://doi.org/10.1080/09540091.2013.854735>.
  45. Pan W-T. Mixed modified fruit fly optimization algorithm with general regression neural network to build oil and gold prices forecasting model. *Kybernetes.* 2014. <https://doi.org/10.1108/K-02-2014-0024>.
  46. Kang L, Xiong X, Yi L, Guo Y. A study of cutting tool wear prediction utilizing generalized regression neural network with improved fruit fly optimization. In: 2018 Prognostics and system health management conference (PHM-Chongqing); 2018. IEEE. pp. 1–7. <https://doi.org/10.1109/PHM-Chongqing.2018.00008>.
  47. Niu D, Wang H, Chen H, Liang Y. The general regression neural network based on the fruit fly optimization algorithm and the data inconsistency rate for transmission line icing prediction. *Energies.* 2017;10(12):2066. <https://doi.org/10.3390/en10122066>.
  48. Box GE, Cox DR. An analysis of transformations. *J Roy Statist Soc B (Methodol)* 1964;26(2): 211–243. <https://www.jstor.org/stable/2984418>
  49. Celikoglu HB, Cigizoglu HK. Public transportation trip flow modelling with generalized regression neural networks. *AdvEngSoftw.* 2007;38(2):71–9. <https://doi.org/10.1016/j.advengsoft.2006.08.003>.
  50. Hesami M, Naderi R, Tohidfar M. Modeling and optimizing medium composition for shoot regeneration of chrysanthemum via radial basis function-non-dominated sorting genetic algorithm-II (RBF-NSGAI). *Sci Rep.* 2019;9(1):1–11. <https://doi.org/10.1038/s41598-019-54257-0>.
  51. Hesami M, Naderi R, Tohidfar M. Modeling and optimizing in vitro sterilization of chrysanthemum via multilayer perceptron-non-dominated sorting genetic algorithm-II (MLP-NSGAI). *Front Plant Sci.* 2019; 10. <https://doi.org/10.3389/fpls.2019.00282>.
  52. Hesami M, Naderi R, Tohidfar M, Yoosefzadeh-Najafabadi M. Application of adaptive neuro-fuzzy inference system-non-dominated sorting genetic Algorithm-II (ANFIS-NSGAI) for modeling and optimizing somatic embryogenesis of Chrysanthemum. *Front Plant Sci.* 2019;10:869. <https://doi.org/10.3389/fpls.2019.00869>.
  53. Matlab V: 7.10. 0 (R2010a). The MathWorks Inc, Natick, Massachusetts 2010.
  54. GraphPad Prism 5 (2005) GraphPad Prism 5. GraphPad Software Inc., San Diego.

## Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

