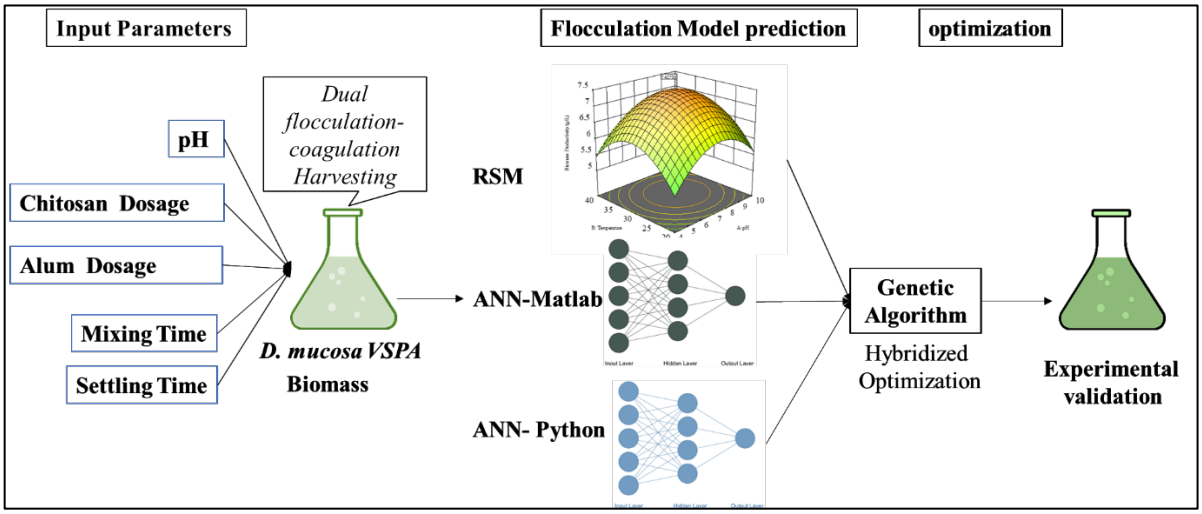
Chapter 7

An artificial neural network-genetic algorithm hybrid approach to optimize the dual flocculation-coagulation harvesting of *Diplosphaera mucosa VSPA*



An artificial neural network-genetic algorithm hybrid approach to optimize the dual flocculation-coagulation harvesting of *Diplosphaera mucosa VSPA*

7.1. Introduction

The main goal of this study is to improve key input parameters to find the best way to harvest Diplosphaera mucosa VSPA using a dual flocculation-coagulation (DFC) mechanism. Dual flocculation-coagulation (DFC) tests were conducted, simultaneously applying an inorganic salt (alum) and organic polymer (chitosan), aiming to enhance efficiency and reduce flocculant dosage. The parameters considered for optimization encompass pH, alum dosage, chitosan dosage, mixing time, and settling time, which are crucial for augmenting flocculation efficiency. Combining Response Surface Methodology (RSM) and Artificial Neural Networks (ANN) models with Genetic Algorithms (GA), optimal input parameter values were determined using predictive modelling. The ultimate objective was to discover parameter combinations that would maximise flocculation efficiency using a comprehensive optimization process that used the predictive powers of RSM and ANN, as well as the optimization acumen of GA. The outcomes of this research are anticipated to yield valuable insights for the optimization of microalgae harvesting, specifically focusing on the DFC mechanism. Moreover, these findings are poised to make significant contributions to the ongoing advancements in this domain.

7.2. Materials and Methods

7.2.1. Strain Isolation and Characterization

The *D. mucosa VSPA* strain used in this investigation was obtained from the sewage treatment plant input in Bhagwanpur, Varanasi, India, which is exactly positioned at 25° 16' 21" North and 83° 0' 16.92" East. The isolation, sequencing, and characterization methods have been described in the past[286]. The BLAST tool was used to analyse the

sequencing of the strain, which revealed a high degree of similarity (89.60%) with *D. mucosa* and was therefore identified as *D. mucosa VSPA*. Textile wastewater was collected from local textile firms in Varanasi, India (25° 18' 0" N 82° 55' 48" E) and utilised to create an isolated strain. They were kept in a well-lit environment (photoperiod 16:8 light: dark) with 1 N HCl and 1 N NaOH to keep their pH stable. Various wastewater sample parameters, such as nitrogen, phosphorus, COD, and BOD, have already been measured and documented in the literature using standardised and globally accepted procedures [399]. This is the first research to optimise input parameters for the harvesting of *D. mucosa VSPA* in wastewater utilising the DFC harvesting mechanism, as far as the author is aware.

7.2.2. Preparation of the chitosan and Alum solution

Aluminium potassium sulphate hydrate was supplied by Labo-gens (with a purity level of 99.5 %). (India). Stock solutions with a concentration of 2 g/L were produced in 200 mL of distilled water and agitated at 100 rpm for one hour. Sigma-Aldrich provided chitosan produced from crab chitin shells (India). Because chitosan is insoluble in water, 0.2 g of chitosan was dissolved in 10 mL of a diluted acetaldehyde (0.1 %) solution, then diluted with 90 mL of distilled water to achieve the necessary 2 g/L stock concentration. The mixture was agitated at 100 rpm for 30 to 60 minutes. These stock solutions were stored at room temperature upon preparation and were utilized within three days to ensure their stability and efficacy [52].

7.2.3. Experimental set-up

Microalgal harvesting experiments were conducted using mid-stationary growth phase suspensions, chosen for their maximum biomass production. This phase occurs after rapid population growth, characterized by reduced cell division due to high cell density. Factors such as nutrient scarcity, light, pH changes, and carbon dioxide depletion contribute to this

slowdown. Consequently, harvesting microalgae during this phase is a well-accepted strategy to optimize the process[52]. To identify the stationary phase, the optical density (OD) of samples obtained from flasks at regular intervals had to be monitored. Microalgal suspension samples of 100 mL were placed in 250 mL beakers. According to the design specified in Table 2, the input parameters pH, alum dosage (mg/L of microalgal suspension), chitosan dosage (mg/L of microalgal suspension), settling time (min), and mixing time (min) were systematically changed. The pH was varied with 1 N HCl and 1 N NaOH and measured using a digital pH metre (Eutech pH Tutor, US). The mixing rate was maintained at 150 rpm until the mixing duration, and the temperature was maintained at 30°C. To evaluate flocculation efficiency, a 15 mL sample of the suspension's supernatant was extracted from the middle one-third to two-thirds of the settled volume. All studies were run in triplicate, and the average data were utilised to fine-tune the results. The optimal process parameter values found through optimization were empirically validated using the same experimental setup.

7.2.4. Harvesting efficiency analysis

Using a UV-VIS spectrophotometer with an optical density of 680 nm, the initial concentration of microalga biomass in quartz cuvettes was calculated (OD680). A 15 mL supernatant sample was pipetted from the solution between one-third and two-thirds up from the bottom to assess flocculation effectiveness. The optical density of the supernatant was measured at a height of 2/3 of the pellucid culture at the end of the flocculation phase. The absence of the component in the culture was employed as a control. The harvesting productivity was measured by [251]:

$$Harvesting\ efficiency\ (\%) = \left(1 - \frac{OD(final)}{OD(initial)} \times 100\%\right) \quad (7.1)$$

Before filtering via glass fibre filter paper, a 100 mL sample of microalgae cell culture was weighed. Microalgal biomass content was determined by drying filter paper holding the sample overnight at a constant weight of 60 °C. The dry microalgal biomass was measured by the weight of the filter paper following drying. The dosage of flocculant was established using:

$$Flocculent\ dosage = \frac{weight\ of\ flocculent\ (in\ mg)}{wieght\ of\ dry\ biomass\ (g)} \quad (7.2)$$

All experiments were replicated to ensure the authenticity of data.

7.2.5. Design of Experiment (DoE)

7.2.5.1 Response Surface Methodology (RSM) modeling

The experimental design for this study used a mix of Central Composite Design (CCD) and Response Surface Methodology (RSM), which were carried out using Design Expert software (Version 13). The experimental design for this study used a mix of Central Composite Design (CCD) and Response Surface Methodology (RSM), which were carried out using Design Expert software (Version 13). The objective was to determine the relationship between harvesting efficiency (%) and five input parameters: Chitosan dosage (mg/L of microalgal suspension), Alum dosage (mg/L of microalgal suspension), pH, mixing time (min), and settling time (min). Based on previous study, the range of the input parameters was identified and separated into three coded levels for each parameter. As shown in **Table 7.1**, a total of 50 experimental runs were completed utilising the CCD design.

Table 7.1. Experimental range and coded levels of five input microalgae biomass harvesting parameters.

Factor	Input Parameter	Coded Levels		
		Low (-1)	Mean (0)	High (+1)

A	Chitosan dosage (mg/L)	10	20	30
B	Alum dosage (mg/L)	10	20	30
C	pH	6	8	10
D	Mixing time (min.)	6	16	24
E	Settling time (min.)	10	25	40

The effect of input parameters on flocculation efficiency was then studied, and their ideal levels for obtaining high flocculation efficiency were established using RSM. RSM is a statistical and mathematical tool for researching and modelling process parameter relationships. For the analysis, a full quadratic model was employed, as presented in Equation ():

$$Y_0 = b_0 + \sum_i b_i X_i + \sum_{ii} b_{ii} X_i^2 + \sum_{ij} X_i X_j \quad (7.3)$$

where Y_0 is the output response (harvesting efficiency), b_0 is the offset term, and b_i, b_{ii} and b_{ij} are the ith linear coefficient, ith quadratic coefficient and ijth interaction coefficient, and X_iX_j are the independent variables respectively.

7.2.5.2 Artificial Neural Network (ANN) model development

In this study, the process parameters were optimised using the feed-forward Artificial Neural Network (FANN) modelling technique. MATLAB, a popular programming language, was used to construct the ANN models. Using the NN toolkit in MATLAB, the ANN modelling was executed. The feed-forward ANNs are a form of neural network made up of linked artificial neuronal layers. The results of RSM-mediated experimental runs, which contained input parameters and harvesting efficiency data, were used in ANN modelling. As indicated in **Figure 7.1**, the ANN structure comprised of three main layers: the input layer, the hidden layer, and the output layer [303]. Nodes or neurons connected these levels, enabling information to travel between them. The input layer received the five

input parameters to begin the modelling process. They were subsequently sent to the hidden layer for processing. Based on the complexity of the problem, several hidden layers were deployed, and the number of hidden layers was established by trial and error, where different numbers were examined until the error was close to zero.

The following equation describes the connection between layers [304]:

$$Z_{j=}f\sum_{i=1}^{N}|xw+b|_i \tag{7.4}$$

where Z and x are the i^{th} and j^{th} inputs and outputs, f is the operating function, w is the weight, and b is the bias. After testing 1 to 15 secret levels, we picked 10 hidden layers for this inquiry. Based on the input, the hidden layers predicted the output, which was then transferred to the output layer for comparison with the experimental data. Finally, the inaccuracy was expected.

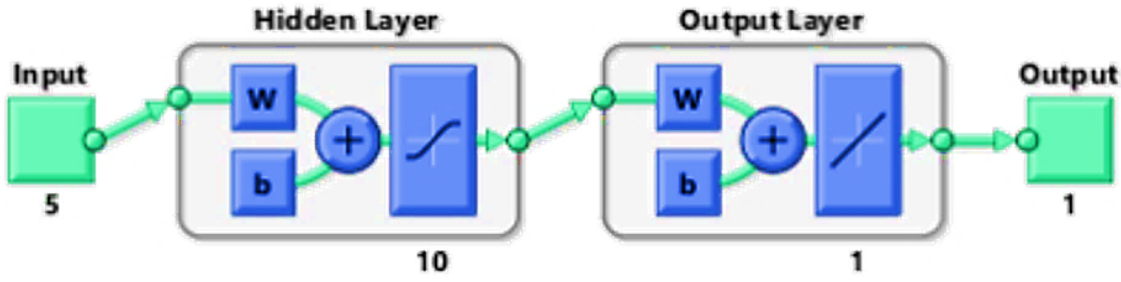


Figure 7.1. An artificial neural network's design, which includes 5 input parameters, a hidden layer, and an output layer.

For the training of the ANN model, the dataset was divided into three subsets using the "dividerand" function: training, validation, and testing. The default configuration was utilised, with 70% of the data going to training (34 data points), 15% going to validation (8 data points), and 15% going to testing (8 data points). This was done to evaluate the ANN models' generalisation capacity, or their ability to make predictions on fresh, unseen data that was not used during the training phase [400]. Throughout the training phase, input and output data from both training sets were utilised to train the network. To alter the model's

parameters, the Levenberg-Marquardt technique, which belongs to the Quasi-Newton class of optimization algorithms, was utilised. To recognise patterns in the data and make accurate predictions, the ANN model was trained across many cycles/iterations. Each repetition is known as an epoch. The validation set was used to fine-tune the model by evaluating its performance after each epoch and adjusting the model's hyperparameters to increase performance. Following the training method, the model's ultimate performance was evaluated using the testing subset. Only input data was utilised to make predictions during this phase, and the trained model was used to anticipate the output.

7.2.5.3 Multi-input optimization by Genetic Algorithm (GA)

Genetic Algorithms (GAs) are highly effective optimization approaches inspired by natural selection and genetics. They've been used successfully in a variety of industries, including chemistry and environmental engineering. The GA approach was utilised in this work to optimise both RSM and ANN models in order to find the ideal conditions for the five input parameters that maximise flocculation efficiency (RSM-GA and ANN-GA). GA is a heuristic search technique inspired by Darwin's theory of evolution that is often employed to address optimisation challenges. The GA algorithm utilised in this study produced a population of 50 individuals, each representing a chromosome and carrying five genes that corresponded to the input parameters. The fitness value of each chromosome was calculated using the optimised RSM and ANN models. The fittest individuals were picked, and genetic operators such as selection, crossover, and mutation were utilised to generate the children of the next generation. The mutation rate was determined.

7.2.6. Statistical Analysis

To ensure accuracy, each batch experiment was repeated three times, and the average results were used for further optimization. The performance of RSM-GA and ANN-GA

was evaluated using three performance parameters: coefficient of regression (R^2), root mean square error (RMSE), and fitness function [401]. The fitness function is a mathematical function that evaluates the solution of the optimisation problem generated by GA. The results of the RSM modelling were also evaluated based on ANOVA results.

7.3. Results and Discussion

7.3.1. The Influence of Input Parameters on Harvesting Efficiency

7.3.1.1 Effect of chitosan and Alum dosage

Chitosan, a biobased polymer with no toxic properties, holds promise for microalgae harvesting through flocculation. The mechanism of flocculation of chitosan depends on the pH [402]. At low pH (< 7) flocculation occurs through bridging, adsorption and charge neutralization mechanisms as because amino group of chitosan is protonated at low pH. Chitosan's high cationic charge density allows for strong charge-neutralization ability through fast bridging and adsorption. The positively charged chitosan attaches to negatively charged microalgae cells, resulting in neutralization of charges and reduction in interparticle repulsion. This process weakens electrostatic repulsion, leading to flocculation. High pH (>7) causes the deprotonation of amino group, which leads the flocculation through the sweep flocculation mechanism [403]–[405]. The current study discovered that a low chitosan concentration is necessary to obtain the best %age of microalgae cell clearance. The clearance % increased with increasing chitosan dose at these low concentrations. A minimal dose of 20mg/L removed 70.3±0.7 % of the cells, which is a considerable removal %age. The proportion of microalgae cells destroyed reduced considerably when greater concentrations of chitosan (more than 30 mg/L) were used. These observations are explained by the charge density process. The ideal chitosan concentration was discovered to be ~20mg of chitosan per litre of microalgal suspension, since greater doses resulted in cell restabilization and decreased removal efficiency. At

higher concentrations (> 30mg/L of chitosan per litre of microalgal suspension), chitosan-adsorbed particles resisted each other, preventing floc formation and lowering removal %. The high repulsion between chitosan's polycations and monolayer adsorption has been observed to cause restabilization. In the past, findings that were comparable to these were reported in the research literature. Fareza et al. [81] discovered that increasing the dosage of chitosan increased microalgae collection, with collection efficiency reaching 98 % at 30 mg/L [406]. While Farid et al. were able to harvest 98% of their cells with 60 mg/L of chitosan, Dong et al. were only able to harvest 95% of their cells with 2 mg/L [407], [408].

The principal mechanism of coagulation using inorganic flocculants is charge neutralization [409], [410]. Suspension stability of small microalgae cells is due to the repulsion stemming from their negatively charged surface. Thus, to counteract this electrostatic stabilization and induce coagulation, positively charged alum ions are necessary for charge neutralization [411]. The present study revealed that harvesting efficiency reaching 65±3% for alum at ~22±1 alum per litre of microalgal suspension. Hence, to achieve optimal harvesting efficiency, it requires notably high dosages of alum.

7.3.1.2 Synergetic effect of Chitosan and alum

Combining inorganic flocculant (alum) with organic flocculant (chitosan) improved flocculation significantly [412], [413]. Compared to individual methods, dual flocculation coagulation (DFC) showed much better harvesting efficiency. For instance, using alum and chitosan together achieved 95±3 % removal at ~20 mg chitosan per liter (equivalent to 10.52 mg chitosan/g dry biomass) and at ~22 mg alum per liter (equivalent to 11.57 mg alum/g dry biomass). This demonstrated 30±2 % improvement in harvesting efficiency compared to single flocculation methods (Section 3.1.1). The synergy in dual flocculation, involving alum and chitosan, significantly boosted efficiency, up to 1.3 to 1.6 times. This is most likely the result of the cooperation of multiple flocculation mechanisms, such as

charge neutralisation, adsorption, bridging, and sweeping, which were all working together to create the floc. Chitosan and inorganic flocculants were involved in this process. When alum is added, it causes cell collisions, which then results in the production of tiny flocs. Chitosan then begins the process of particle trapping and bridging after it has been applied [414]. Chitosan chains bind to existing microalgal-alum flocs, causing them to clump together (macroflocs). These findings imply that combining modest dosages of alum and chitosan might result in improved microalgae biomass harvesting, less contamination, and cost savings. Vu et al. found that mixing chitosan with inorganic flocculants ($FeCl_3$ and $Al_2(SO_4)_3$) resulted in better harvesting of microalgae than using the flocculants alone. For $FeCl_3$/chitosan and $Al_2(SO_4)_3$/chitosan, the harvesting capacity was increased by 57% and 24%, respectively [251]. Additionally, Loganathan et al. showed that the dual harvesting system of chitosan and $Al_2(SO_4)_3$ produced a synergistic impact on the harvesting of microalgae, which means that the combined effect of the two flocculants was larger than the sum of their separate effects [415]. This observation is consistent with existing literature, where enhanced flocculation performance (up to 90%) with inorganic flocculants like ferric chloride and aluminum sulfate demands substantial doses (as illustrated in **Table 7.2**). Discrepancies in microalgal culture and growth conditions could contribute to the variations in optimal dosages observed across these studies.

Table 7.2. Summary of previous studies on flocculation using chitosan and aluminium sulphate in relation to the results of this study

Microalgae species	**Dry Microalgal biomass (g/L)**	**Flocculant type**	**Optimal Flocculant dose (mg/g dry biomass)**	**Harvesting Efficiency (%)**	**References**
D. mucosa VSPA	**1.9**	• Alum • Chitosan	11.52 12.67	>95	This Study
Chlorella vulgaris	0.36	• Aluminium sulphate/Ferric chloride	504/ 448	>89	[251]

		• Chitosan	224		
Chlorella vulgaris (freshwater)	1.0	• Aluminium sulphate • Ferric chloride	350 300	>95	[266]
Season algal cells	-	• $Al_2(SO_4)_3$ • chitosan	35-45 mg/L 0.4-0.7 mg/L	-	
Scenedesmus obliquus		• Chitosan • buoy-bead flotation (BBF)	20 mg/L 50 mg/L	>83.77	[416]
Chlorella vulgaris		• Chitosan • buoy-bead flotation (BBF)	30 mg/L 50 mg/L	>92.47	[416]
Chlorella vulgaris	1.2	• Aluminium sulphate • Chitosan	2083 208	>90	[264]
Chlorella sp.	0.12	• Aluminium sulphate • Ferric chloride	1266 1191	>90	[265]

7.3.1.3 Effect of pH

The efficiency of *D. mucosa VSPA* harvesting is influenced by several factors, including pH. The harvesting efficiency increases with pH, but seems to plateau at about pH 8. This is consistent with the findings of earlier research. This is in line with earlier research, which shows that variations in protonation and structural changes in the floc cause auto-flocculation to happen at high pH. In statistical analysis, it was determined that pH, which is 8 in this case, had less of an impact on flocculation efficiency and that flocculation efficiency increased as pH increased. Changes in pH are known to affect the structure of the flocculant. At acidic pH, chitosan exists in a stretched-out form and remains distributed due to the repulsion between adjacent NH_2 groups and positively charged NH^+ groups. This repulsion keeps microalgae cells dispersed and prevent the formation of floc [417]. On the flip side, when the environment turns alkaline, the positive charges on chitosan gradually get neutralized, making it less effective in cancelling out the charges on microalgae cells

[403]. The optimal flocculation of microalgae using chitosan happens within a specific pH range, roughly between 6 and 8 [404]. Microalgae auto-flocculation has been observed in several articles to boost flocculation efficacy at high pH [418], [419], [420] . The flocculants, on the other hand, turn negative when the pH is too high, causing repulsion and lowering flocculation effectiveness [421]. Additionally, Vandamme et al. (2012) found that the repulsive force might reduce particle interaction, which disturbs the flocculation process [422].

7.3.1.4 Effect of Mixing time

In this DFC study, it was found that flocculants (Alum+ Chitosan) require a mixing time of 16 minutes to coagulate 95.0 ± 2% of microalgae cells. Due to less contact or collision between flocculant particles and microalgal cells, flocculation rates were lower when mixing durations were shorter (8–10 minutes). Collisions between flocculants and microalgal cells increased as mixing duration was extended from 8 to 16 minutes, allowing for robust bridging and adsorption to occur. Longer mixing times can increase the surface area of oil-water interfaces by breaking up more oil droplets [423]. Longer mixing durations could result in more microalgae cells breaking up when chitosan, alum, and microalgae cells are combined. This discovery may also be applicable to this scenario. This could result in more chitosan being adsorbed to the microalgae cells, which might enhance the stability of the resultant emulsion. Further increasing the mixing time resulted decrease removal %age, this is may be due to dispersion of alum flocs by prolonged mixing [251].

7.3.1.5 Effect of Settling time

According to the research discussed above, *D. mucosa VSPA* harvesting efficiency is significantly influenced by the settling period (sedimentation). The optimum settling time is typically in the range of 30-120 minutes, but the exact optimum time may vary depending on the specific application and mechanism of harvesting. Researchers have studied the

kinetics of flocculation with chitosan to understand how settling time affects the flocculation process. This involves monitoring the change in flocculation efficiency over time and analyzing the rate at which suspended particles aggregate and settle[424]. Additionally, the impact of settling time on the flocs' particle size distribution during chitosan-based flocculation has been studied. By analyzing the changes in particle size distribution over different settling times, researchers aim to understand the dynamics of floc formation, growth, and sedimentation. Settling time influences the structure and stability of flocs, which in turn affect their settling behavior[424]. Understanding these settling characteristics can aid in optimizing the flocculation process and designing efficient sedimentation units. In the DFC study, it was observed that using a combination of flocculants (Alum + Chitosan) requires a 20-minute settling period to coagulate around 94.0 ± 2% of microalgae cells effectively. Attempting to achieve this with shorter settling times (around 10 minutes) resulted in lower harvesting efficiency. This was due to the formation of smaller flocs, as there was insufficient time for flocs to grow. On the other hand, extending the settling time from 10 to 20 minutes led to the formation of larger flocs. This allowed for strong bridging and better adsorption which led better harvesting efficiency. The harvesting efficiency decline as the settling period was prolonged. This occurred as a result of bigger flocs impeding floc sedimentation. When employing a settling period of at least 20 minutes and at mixing rate at 200 rpm, Ahmad et al. were able to harvest microalgae cells with a harvesting efficiency of 99.0±0.4 %. In contrast, shorter mixing durations of 5 min and 250 rpm led to lower harvesting capacities[425].

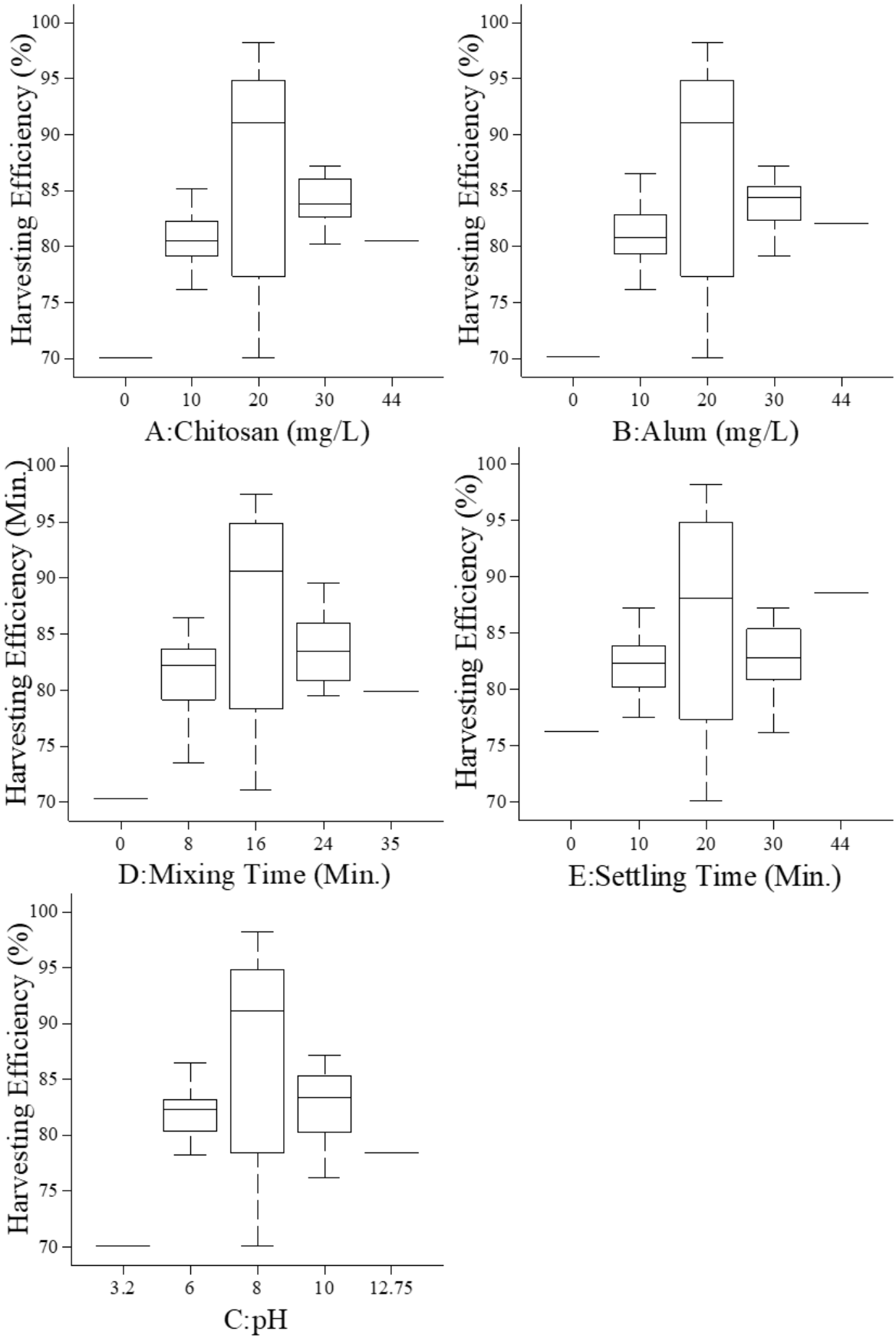


Figure 7.2. Input response relationship of individual parameters on harvesting efficiency of *D. mucosa* VSPA

7.3.2. RSM Modelling

As outlined in section 3.1 of the study, five input parameters significantly impact the harvesting efficiency (%): Chitosan dosage (mg/L of microalgal suspension), Alum dosage (mg/L of microalgal suspension), pH, mixing time (min.), and settling time (min.). Therefore, optimizing these input parameters is essential. Model prediction was carried out utilising both Response Surface Methodology Analysis (RSMA) and Artificial Neural Networks for this purpose (ANN). The results of the ANOVA analysis performed during the flocculation of the D. mucosa VSPA are shown in **Table 7.3**. The statistical significance of the constructed model, effects of individual input parameters, their squares, and interactions were evaluated using p-values and F-values. At the 95 % confidence level, the p-values for all input parameters (A-chitosan dose; B-alum dosage; C-pH; D-mixing time; and E-settling time), their squares, and the two-way interaction (AB; AC; AD and CD) were all less than 0.05. This shows that all input factors and two-way interaction (AB; AC; AD and CD) have a statistically significant impact on harvesting effectiveness. Other investigations and similar optimization are consistent. Liang et al., 2022 used response surface methodology (RSM) and a Box-Behnken design (BBD) to enhance a combined flocculation technique for C. vulgaris utilising calcium hydroxide and chitosan. Setting the pH to 8.97, adding 2 g/L of calcium hydroxide, delivering 20 mg/L of chitosan, and maintaining a flocculation period of 60 minutes resulted in a maximum flocculation efficiency of 97.08 % [426].

The alum dose, followed by the chitosan dosage and mixing time, was the most important input parameter among the individual input parameters. While *D. mucosa VSPA* harvesting efficiency utilising the DFC mechanism is least affected by pH. The efficiency of harvesting was noticeably impacted by two interactions between the dosages of chitosan-alum, chitosan-pH, chitosan-mixing duration, and pH-mixing time. Chitosan dosage-

Mixing time and chitosan dosage-Alum dose were the two parameters that had the greatest impact on two-way interactions. The multiple regression approach was used to fit the experimental data to the RSM quadratic equation. The coded quadratic equation predicting the harvesting efficiency is given in Eq. (7.5):

$$\boldsymbol{Harvesting\ efficiency}\ (\%) = 95.2584 + 1.56\,A + 1.89\,B + 0.55\,C + 1.53\,D + 1.17E - 0.87AB + 0.65\,AC - 1.26\,AD + 0.22\,AE + 0.12\,BC - 0.18BD - 0.34BE + 0.65CD - 0.08CE + 0.45DE - 2.18\,A^2 - 1.94\,B^2 - 3.42C^2 - 3.36D^2 - 2.42E^2$$

$$R^2 = (0.96) \qquad (7.5)$$

Insignificant terms can be eliminated from equation (7.5) to further simplify it, as illustrated in equation (7.6):

$$\boldsymbol{Harvesting\ efficiency}\ (\%) = 95.2584 + 1.56\,A + 1.89\,B + 0.55\,C + 1.53\,D + 1.17E - 0.87AB + 0.65\,AC - 1.26\,AD + 0.65CD - 2.18\,A^2 - 1.94\,B^2 - 3.42C^2 - 3.36D^2 - 2.42E^2 \qquad (R^2 = 0.93) \qquad (7.6)$$

The model constructed showed high significance with a large F-value of 31.51 and a low p-value of less than 0.001, indicating its validity at a 95% confidence level. This signifies that the occurrence of such a substantial F-value due to experimental noise is only 0.01%. The predicted-adjusted R^2 values (0.96 and 0.93) indicate that there is a substantial relationship between input parameters and harvesting efficiency. Adeq Precision evaluates the signal-to-noise ratio, and a ratio greater than 4 is considered preferred. The signal has a model ratio of 19.804, which is more than enough [427]. This model is well-suited for navigating the design space, enabling the mathematical prediction of harvesting efficiency while considering variations in input parameters. Furthermore, an F-value of 2.04 indicates that the lack of fit is tiny and insignificant when compared to the pure error, confirming the model's ability to estimate harvesting efficiency [428].

Table 7.3 Results of ANOVA obtained after RSM modelling.

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	1970.19	20	98.51	31.51	< 0.0001	**significant**
A-Chitosan	105.82	1	105.82	33.85	< 0.0001	
B-Alum	156.17	1	156.17	49.95	< 0.0001	
C-pH	13.15	1	13.15	4.21	0.0494	
D-Mixing Time	102.20	1	102.20	32.69	< 0.0001	
E-Settling Time	60.14	1	60.14	19.24	0.0001	
AB	24.68	1	24.68	7.89	0.0088	
AC	13.65	1	13.65	4.37	0.0455	
AD	50.75	1	50.75	16.23	0.0004	
AE	1.67	1	1.67	0.5326	0.4714	
BC	0.4753	1	0.4753	0.1520	0.6995	
BD	1.02	1	1.02	0.3247	0.5732	
BE	3.71	1	3.71	1.19	0.2848	
CD	14.18	1	14.18	4.53	0.0418	
CE	0.2278	1	0.2278	0.0729	0.7891	
DE	6.57	1	6.57	2.10	0.1579	
A^2	264.49	1	264.49	84.60	< 0.0001	
B^2	209.79	1	209.79	67.10	< 0.0001	
C^2	652.90	1	652.90	208.83	< 0.0001	
D^2	629.63	1	629.63	201.39	< 0.0001	
E^2	325.52	1	325.52	104.12	< 0.0001	
Residual	90.67	29	3.13			
Lack of Fit	78.43	22	3.57	2.04	0.1686	not significant
Pure Error	12.23	7	1.75			
Cor Total	2060.86	49				

df= degree of freedom, cor = correlation

7.3.3. Response 3D Plots by RSM optimization:

Graphical representations such as three-dimensional (3D) response surface plots can be used to investigate the interaction effects of combined input parameters on the harvesting efficiency. As illustrated in **Fig. 7.3 and 7.4**, the surface plots were generated by changing any two variables in the design space while keeping the remaining independent component constant at its midpoint. The response surface displayed an arc containing the ideal region for the input parameters. The noticeable peaks in the 3D response plots show that the ideal

conditions for getting the best response variable performance occurred within the intended design space.

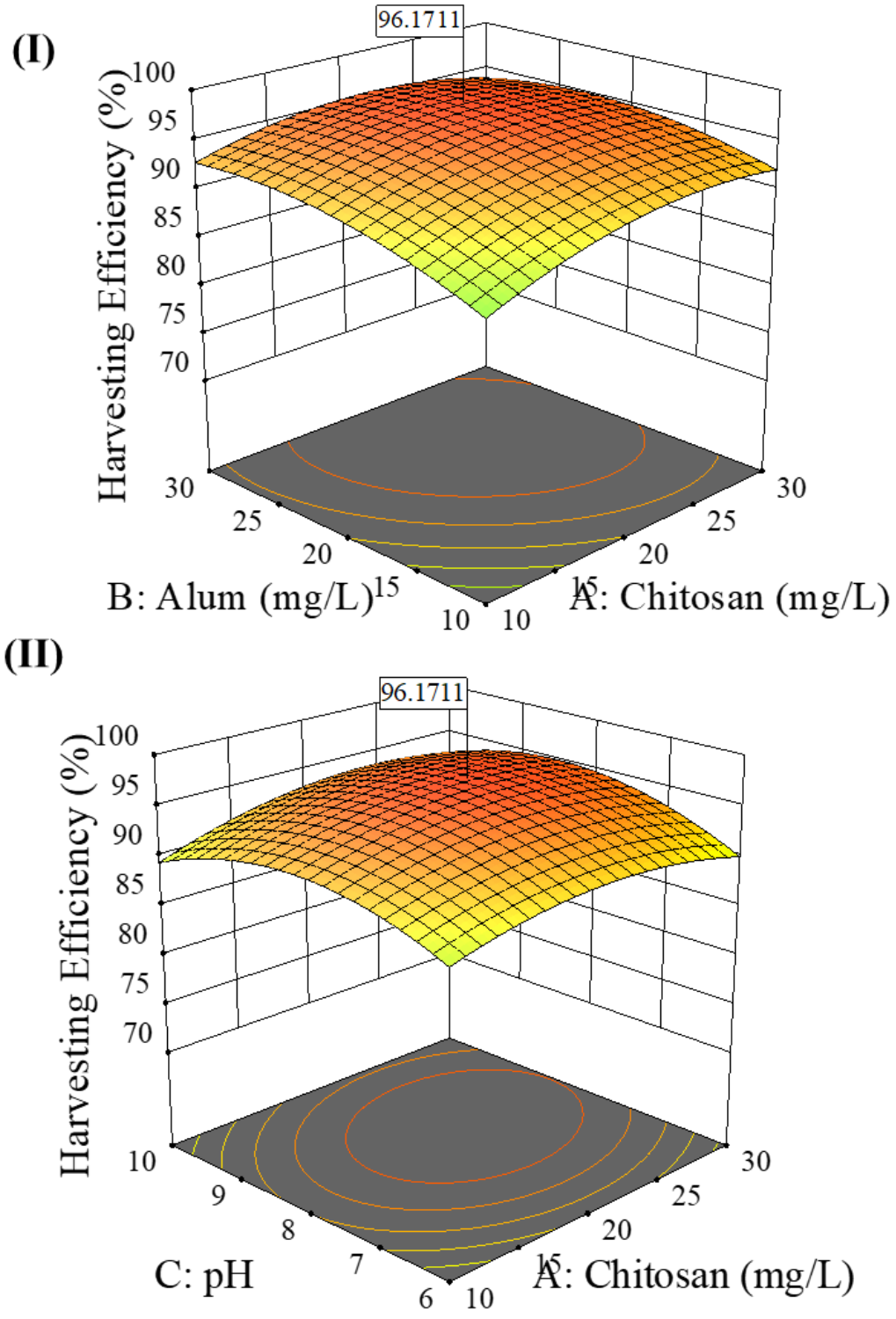


Figure 7.3. The two-way interaction is represented by a three-dimensional response surface map.: (I) chitosan dosage- alum dosage, (II) chitosan dosage- pH

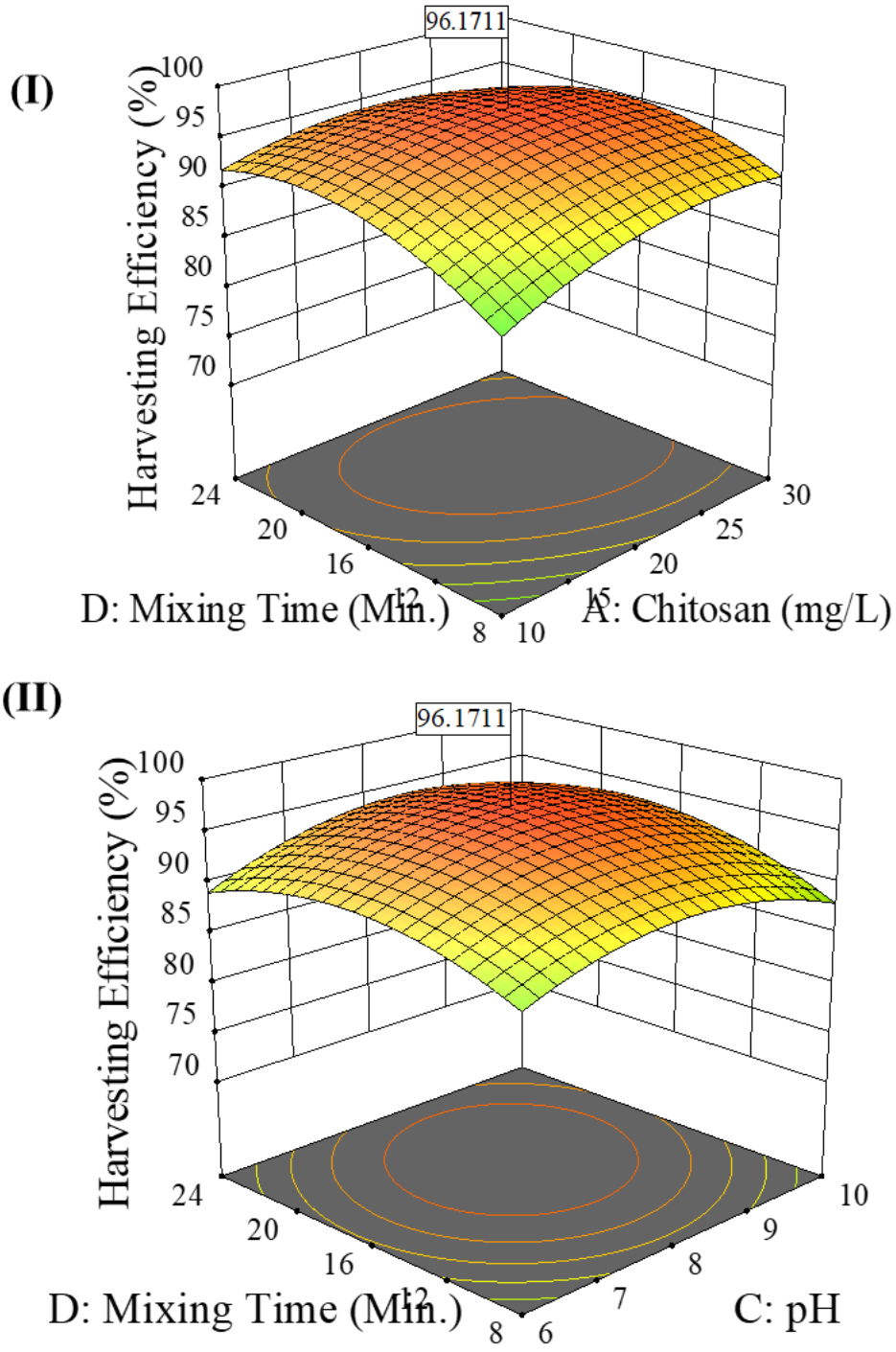


Figure 7.4. The two-way interaction is represented by a three-dimensional response surface map.: (I) chitosan dosage-Mixing time, and (II) and pH-mixing time.

Figure 7.3 and 7.4 reveals that chitosan dosage and alum dosage exhibited a synergistic and significant impact on harvesting efficiency. The highest harvesting efficiency (96.17%) was attained with approximately 22 mg of chitosan per liter of microalgal suspension and around 24 mg of alum per liter of microalgal suspension. Chitosan dosage also exhibited a

two-way interaction with pH and mixing time. The optimal harvesting efficiency was achieved at a pH of around 8 and a mixing time of approximately 16-18 minutes. The interaction between mixing time and pH similarly yielded the same outcome as depicted in **Figure 7.4.** A comprehensive explanation of the two-way interaction between chitosan dosage, alum dosage, pH, and mixing time can be found in detail in section 7.3.1.

The optimum values of input parameters, chitosan dosage (A), alum dosage (B), pH (C), mixing time (D), settling time (E) in RSM optimization were 22.5 mg of chitosan per liter of microalgal suspension (equivalent to 11.84 mg chitosan/g dry biomass), 24.06 mg of alum per liter of microalgal suspension (equivalent to 12.66 mg chitosan/g dry biomass), 8.25, 17.6 minutes and 22.43 minutes respectively (**Fig. 7.5**). Under these optimum conditions, the flocculation efficiency reached 96.17%. The flocculation performance of the suggested approach was also compared to other research in the study. An experiment found that harvesting Chaetoceros muelleri was more than 99 % efficient with an ideal dosage of chitosan (80 mg/L) at pH 9.6 and a settling duration of 40 minutes [429]. This is greater than the flocculation efficiency of previous research, such as one that used a dual coagulation approach with 5 mg/L FeCl3 and 0.5 mg/L chitosan (97.2 %) and another that used 10 mg/L aluminium sulphate and 1 mg/L chitosan (10 mg/L aluminium sulphate and 1 mg/L chitosan) (97.6 %) [415]. In another investigation, the lowest chitosan and aluminium sulphate dosages necessary to achieve more than 90% harvesting efficiency in C. vulgaris were 250 mg/L and 2500 mg/L, respectively [430]. The flocculation efficiency in this study was likewise comparable to that of another study (96.3%)[415] and greater than that of two other investigations (84.6% and 90%) [251], [431]. In comparison to these investigations, dual flocculation-coagulation (DFC) was efficient and clearly superior than monoflocculation tests.

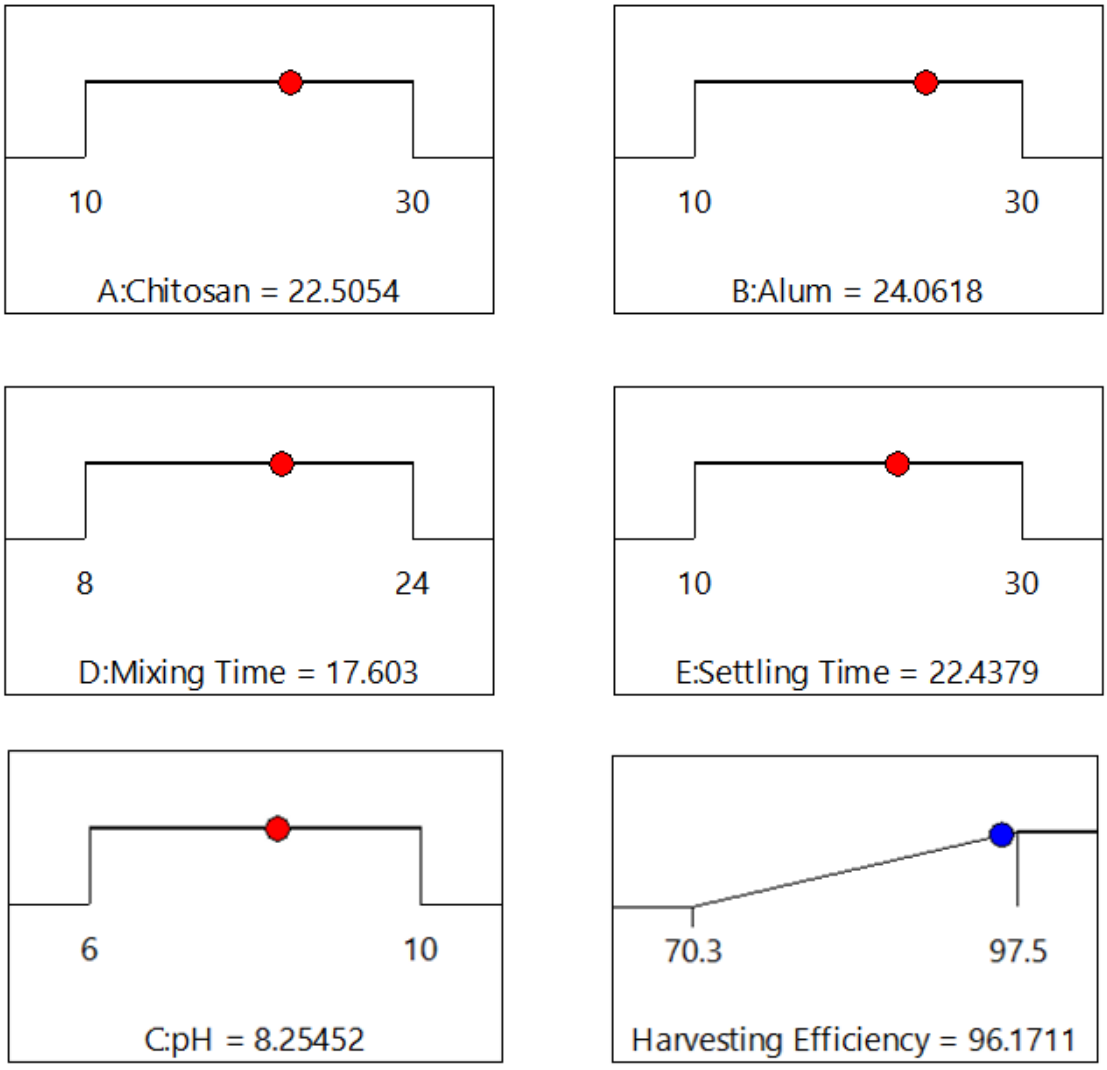


Figure 7.5 Optimum value of Input parameters achieving optimum harvesting efficiency using RSM optimisation.

7.3.4. ANN Modelling

The non-linear link between input parameters and harvesting efficiency was established using Artificial Neural Network (ANN) models. The ANN model was created using MATLAB with a 5-10-1 topology. As demonstrated in **Table 7.4**, these ANN models successfully predicted harvesting efficiency values for D. mucosa VSPA. **Figure 7.6** displays simulation curves for ANN models after training, validation, and testing for harvesting efficiency. The models were trained on 34 data points across 5 epochs (R2 -

0.99), verified on 8 data points (R2 - 0.98), and tested on additional 8 data points (R2 - 0.96). The correctness of the ANN models is shown further by residual histogram plots (**Figure 7.7**). The figures show a high frequency of residuals near 0 for the training, validation, and testing datasets, highlighting the models' accuracy in predicting new values.

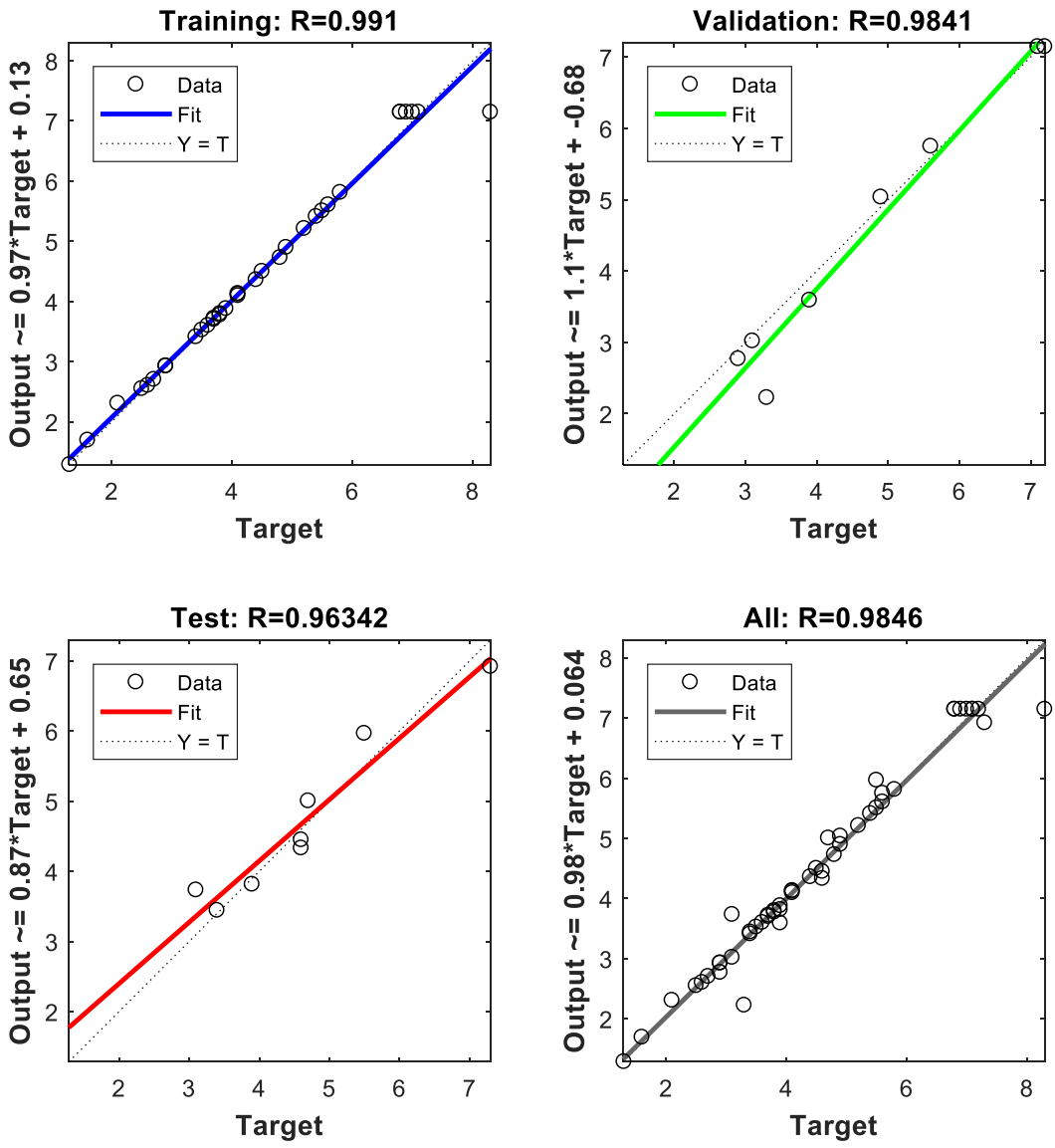


Figure 7.6 Simulation results from *D. mucosa VSPA* harvesting following training, validation, and testing of ANN models

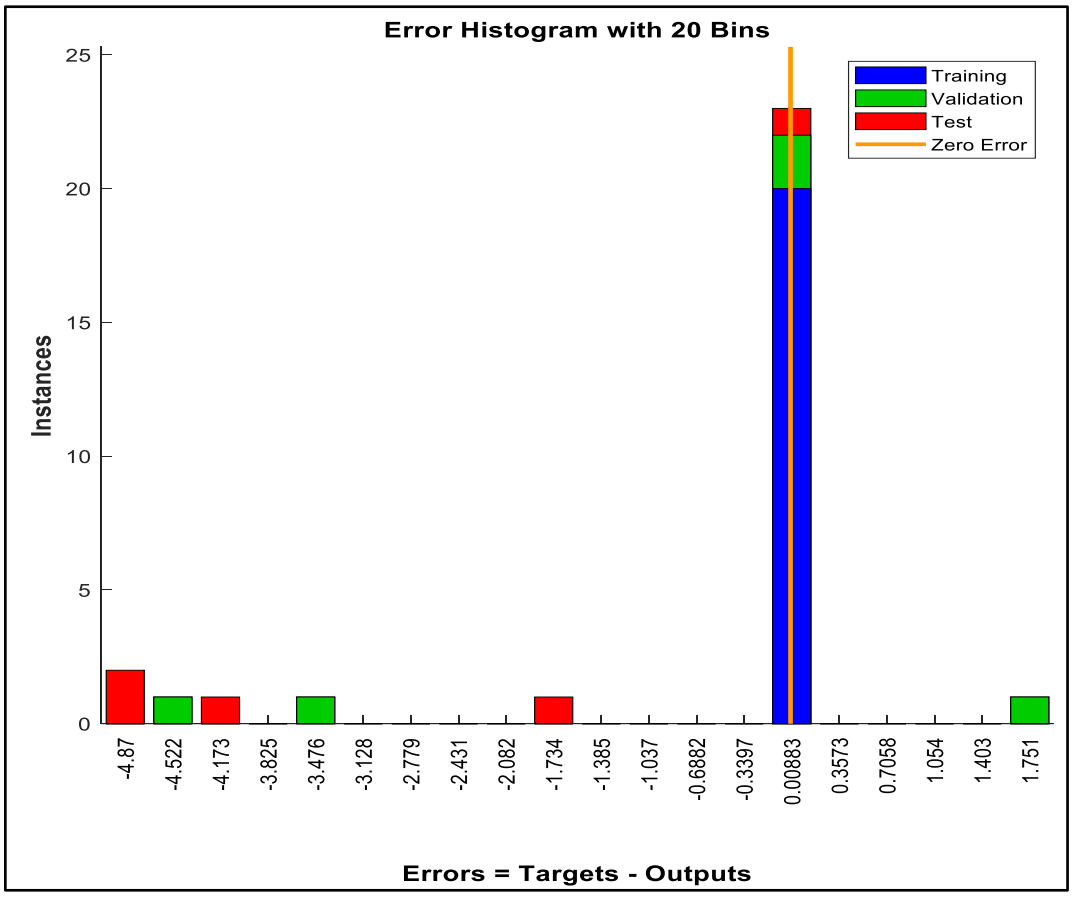


Figure 7.7 Residual histogram plots were generated after training, validating, and testing ANN models during the harvesting of *D. mucosa VSPA*.

RSM emphasizes statistical significance and interaction through ANOVA, while ANN excels at capturing non-linear input-output relationships [432]. The generalisation capability of the produced response surface methodology (RSM) and artificial neural network (ANN) models was evaluated. This was performed by calculating their respective R^2, MSE, and root mean square error (RMSE). MSE and RMSE for ANN and RSM predicted models are 0.61, 0.78, and 1.81, 1.35, respectively. The R^2 values obtained for ANN models (0.99) were higher and closer to 1 than for RSM models (0.96), indicating that the ANN models matched the data better. ANN surpassed Response Surface Methodology (RSM) in terms of prediction accuracy, as evidenced by reduced Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) values. Table 7.4 displays experimental data as well as anticipated values plotted against experimental runs. A

detailed examination of the produced data revealed that the ANN predictions closely matched the experimental values in the harvesting of *D. mucosa VSPA*. This lends credence to ANN's superior generalisation capabilities over RSM. Numerous studies in the literature have repeatedly demonstrated that ANN outperforms RSM in terms of prediction ability. For example, Hossain et al. (2022) improved *Chlorella kessleri* culture process parameters and reported ANN's excellent prediction performance ($R^2 = 0.99$) and low error (0.7)[215]. Garg and Jain (2020) observed similar patterns in lipid and biodiesel productivity optimization for microalgal species. This study found that ANN models provided better and more accurate predictions than RSM, as stated in the literature [433], [434]–[437].

Table 7.4. Experimental design and predicted values of harvesting efficiency by RSM and ANN.

Factor 1 Chitosan Dosage (mg/L)	Factor 2 Alum Dosage (mg/L)	Factor 3 pH	Factor 4 Mixing time (min.)	Factor 5 Settling time (min.)	Response Harvesting Efficiency (%)	RSM predicted	Sq. Error (RSM)	ANN Predicted	Sq. Error (ANN)
30	30	10	24	10	84.2	85.16	0.92	85.16	0.92
10	10	6	8	30	79.1	76.42	7.18	78.42	0.46
10	10	6	24	30	82.5	81.94	0.31	81.94	0.31
10	30	6	24	30	85.5	86.21	0.50	85.21	0.08
30	30	10	8	30	86.5	85.59	0.83	85.59	0.83
30	10	10	24	10	80.5	82.55	4.20	80.55	0.00
30	10	6	8	10	82.1	80.21	3.57	82.21	0.01
20	20	8	16	20	93.1	95.26	4.67	93.26	0.03
20	20	8	16	20	95.2	95.26	0.00	95.26	0.00
20	20	8	16	20	94.6	95.26	0.44	94.26	0.12
20	0	8	16	20	78.2	79.75	2.40	79.75	2.40
10	30	6	24	10	84.5	83.92	0.34	83.92	0.34
10	30	6	8	30	83.1	81.40	2.89	82.40	0.49
10	30	10	24	30	87.2	87.42	0.05	87.42	0.05
20	20	8	16	20	97.5	95.26	5.02	97.26	0.06
30	30	6	8	30	84.5	84.44	0.00	83.44	1.12
30	30	10	24	30	87.2	88.03	0.69	87.03	0.03
20	20	8	16	20	96.2	95.26	0.88	96.26	0.00
10	10	10	24	10	80.1	79.34	0.58	79.84	0.07
20	43.78	8	16	20	88.1	88.78	0.46	88.08	0.00
10	10	6	8	10	74.2	74.57	0.14	74.27	0.00

10	10	10	24	30	81.2	82.66	2.13	81.66	0.21
10	30	10	8	10	80.5	79.80	0.49	81.80	1.69
20	20	8	16	20	94.3	95.26	0.92	94.26	0.00
30	30	6	24	10	82.5	81.01	2.22	82.01	0.24
30	10	10	8	10	83.2	81.22	3.92	84.22	1.04
20	20	8	16	20	94.5	95.26	0.58	94.26	0.06
20	20	8	16	0	76.3	78.77	6.10	76.77	0.22
30	10	6	24	30	82.7	83.45	0.56	82.45	0.06
20	20	8	16	20	95.2	95.26	0.00	94.26	0.88
10	10	10	8	10	73.5	72.96	0.29	73.96	0.21
30	10	6	24	10	80.2	78.89	1.72	81.89	2.86
30	10	10	24	30	89.5	86.78	7.40	89.58	0.01
20	20	3.24	16	20	71.1	74.55	11.90	72.55	2.10
10	30	10	8	30	79.2	79.94	0.55	78.94	0.07
20	20	8	16	43.78	84.6	84.37	0.05	84.27	0.11
20	20	12.76	16	20	78.4	77.18	1.49	77.88	0.27
30	10	10	8	30	82.3	83.63	1.77	83.33	1.06
10	10	10	8	30	74.2	74.47	0.07	74.27	0.00
10	30	10	24	10	86.5	85.46	1.08	85.06	2.07
20	20	8	0	20	70.3	72.56	5.11	71.56	1.59
43.78	20	8	16	20	85.5	86.63	1.28	85.23	0.07
30	10	6	8	30	82.9	82.97	0.00	82.47	0.18
10	10	6	24	10	79.5	78.28	1.49	77.28	4.93
0	20	8	16	20	78.1	79.20	1.21	79.10	1.00
10	30	6	8	10	81.2	80.92	0.08	81.92	0.52
20	20	8	35.024	20	79.9	79.87	0.00	79.87	0.00
30	30	6	24	30	85.1	84.21	0.79	84.21	0.79
30	30	10	8	10	84.5	84.54	0.00	83.54	0.92
30	30	6	8	10	84.2	83.05	1.32	84.05	0.02

7.3.5. RSM-GA and ANN-GA based optimization

As objective functions, the response surface methodology (RSM) and artificial neural network (ANN) models were used to build the best-fit model of the input parameters for optimal harvesting efficiency. The RSM and ANN models were then combined using the genetic algorithm (GA) to generate optimal input parameter values. Hyperparameters such as beginning population size, elite size, and mutation rate impacted the performance of the GA technique. Trial and error were used to find the following parameters: population size (100), elite size (10) and mutation rate (0.1). After 100 generations, the optimization operation was terminated when the global solution was obtained. The optimal conditions

for the following input parameters are shown in **Table 7.5**: chitosan dosage (A), alum dosage (B), pH (C), and mixing duration (D). This suggests that the ANN-GA-based technique to optimization is both resilient and dependable. The ANN-GA optimization findings revealed that the optimal values of the input parameters were 20.5 mg of chitosan per liter of microalgal suspension (equivalent to 10.78 mg chitosan/g dry biomass), 20.4 mg of alum per liter of microalgal suspension (equivalent to 10.73 mg alum/g dry biomass), 8.21, 16.8 minutes and 20.5 minutes respectively. This suggests that a lower chitosan/alum dosage is needed to achieve the maximum flocculation efficiency (98.87%) in dual flocculation-coagulation mechanism. The optimization process increased harvesting efficiency compared to those previously reported literatures [251], [431]. The ANN-GA hybrid strategy for optimising additional process parameters such as flocculent type, mixing rates, ionic strength, floc size, Zeta potential, and other significant parameters impacting microlage harvesting efficiency may also be used.

Table 7.5 The optimal global solution as a result of the genetic algorithm hybridisation technique and experimental validation

Input Parameters (Coded Value)	**RSM**	**RSM-GA**	**ANN-GA**
A: Chitosan (mg/L)	22.5	21.5	20.5
B: Alum (mg/L)	24.06	22.16	20.4
C: pH	8.25	8.11	8.21
D: Mixing Time (min.)	17.6	17.4	16.8
E: Settling Time (min.)	22.43	21.57	20.05
Output Parameter (Prediction)			
Harvesting efficiency (%)	96.17	96.47	98.87
Experimental Verification			
Harvesting efficiency (%)	92.31	94.61	98.19
% Error	4.18	1.96	0.68

In summary, the ANN-GA hybrid approach was found to be a robust and reliable method for optimizing the input parameters for high harvesting efficiency of *D. mucosa VSPA*. The strategy may be expanded to improve other process parameters as well. However, it is vital to remember that the optimization method used will be determined by the application. If the objective function is complex or non-linear, then GA may be a better choice than other optimization methods. Additionally, if the experimental time and cost are limited, then GA may be a better choice as it can search a wider range of potential solutions.

7.4. Implications for the Future

The application of coupled RSM-GA and ANN-GA hybrid models has great potential for simplifying and automating both pilot-scale and large-scale harvesting. These models provide an accurate forecast of the flocculating mechanism, as well as insights into ideal operational settings for maximising microalgal harvesting via the DFC mechanism. Moreover, this hybrid methodology has the potential to extend its optimization capabilities to key parameters such as flocculent type, mixing rates, ionic strength, floc size, Zeta potential and shear stress. Adherence to such parameters via the use of RSM-ANN-GA hybrid models promotes ecologically conscious and compliant flocculation, which contributes to successful harvesting.

7.5. Conclusion

This study emphasized the viability and optimization of dual flocculation coagulation (DFC) mechanism to enhance microalgae harvesting. By using alum/chitosan in combination showed synergetic effect, and leads flocculation efficiency exceeding 98.8% was achieved. RSM and ANN was used to generated best predictive model for enhanced harvesting. ANOVA analysis revealed that all input parameters, Chitosan-alum dosage, Chitosan-pH, Chitosan-Mixing time and pH-Mixing time significantly influenced the harvesting efficiency. The ANN (MSE-0.61, RMSE 0.78) model demonstrated superior

generalization ability compared to the RSM model (MSE-1.81, RMSE 1.35). For the optimization, a multi-input approach was pursued using RSM, RSM-GA and ANN-GA techniques. The RSM, and RSM-GA projected higher values for dependent variables, achieving 96.17% and 96.47% harvesting efficiency, respectively. In contrast, the ANN-GA optimized input variable values, achieving 98.87% harvesting efficiency. Furthermore, optimization capabilities of ANN-GA (error: 0.68) were robustly established through this study. Additionally, ANN-GA was found superior in optimizing. These methods are useful for solving issues with multiple inputs and objectives, such as microalgal harvesting based on DFC.