INTEGRATION OF RESPONSE SURFACE METHODOLOGY AND ARTIFICIAL NEURAL NETWORK FOR GROWTH OPTIMIZATION OF CHLORELLA VULGARIS IN PROSPECT OF BIODIESEL PRODUCTION

Mumayyiza Tahir [*], Muhammad Imtiaz Shafiq *

[*] School of Bio-Chemistry and Bio-Technology, University of the Punjab Canal Road, Quaid-e-Azam Campus, Lahore, Punjab, Pakistan 05422

Abstract- Biodiesel as an alternative fuel has shown the potential to mitigate the exceeding energy demands in automotive sector. In this scenario, third-generation biodiesel production from microalgae emerged as a clean and reliable method. In the current work, growth conditions for freshwater algae collected from Punjab University, Lahore, Pakistan were optimized. The collected samples were grown in a glass photobioreactor utilizing BG-11 and Basel Bold Medium (BBM). The optical density (OD) at 688 nm was recorded for 15 days, at the light/dark cycle of 12/24 and 16/24, and at two distinct temperatures of 25ºC and 28 ºC. The developed quadratic prediction model of OD using Response Surface Methodology (RSM) and diagnostic study showed the reasonable agreement between actual and predicted results. Moreover, RSM predicted results were incorporated for creating 12 different Artificial Neural Network Models (ANNMs) with hidden layer neurons in range of 9-20. The ANNM with 5 neurons in hidden layer and TRAINCGP training function showed best statistical performance with least percentage contribution to mean square error (MSE) chart (1.24%) and highest correlation coefficient (0.99005). Finally, the OD results were numerically optimized using multi objective optimization which designated 13^{th} day, photoperiod of 16/24, temperature of 25 ºC and BG11 media for maximum OD of 2.862. The optimization identified result was experimentally validated and absolute percentage error (APE) was 3.92%. Thus, the integrated use of RSM statistical results and ANN modelling could be effectively used for prediction and optimization of Chlorella Voulgaris growth for maximizing biodiesel yield.

Keywords- Artificial Neural Network, Biodiesel, Chlorella Voulgaris, Microalgae, Response Surface Methodology

1 INTRODUCTION

In recent decades, petroleum products have fueled social progress and economic success. However, they have exacerbated a double dilemma of energy scarcity and pollution. The technological innovations and modern-day machinery have collectively contributed towards the higher fossil fuel consumption and sheer damage to the environment [1-3]. In order to deal with the very issues, researchers have long been in a quest to look for alternative energy resources[4-7]. Biodiesel as a fuel is one of such potential sources with promising efficiency and less deteriorating effect on the environment in terms of greenhouse emissions[8]. It is a clean, sustainable, and low-cost green energy manufactured from natural plant-derived fats and acid monoesters[9, 10]. Through the whole advancement of biodiesel production, the research has always been focused on advanced reliable methods[11, 12]. In recent years, 3rd generation made biofuel from microalgae is gaining popularity but has inherent dependence on the growth optimization of incorporated resources [13]. Thus, the proper growth conditions for microalgae needs to be optimized for desirable large-scale biodiesel production and have been thoroughly addressed in this work.

Over the years, much work has been reported regarding utilization of various useful sources for biodiesel production. Waste cooking oil has proven to be a reliable source for large scale production of biodiesel. Xiangmei Meng et.al, produced biodiesel from waste cooking oil collected from restaurants in main cities of China using alkali catalyst and performed the engine testing. They deduced that biodiesel could be used as a green fuel without any engine modifications[14]. The energy and cost assessment of green fuel production, biodiesel, through traditional transesterification reaction and ultrasonic was conducted by Ahmad Mohammad shirazi et.al. They reported the cost ratio benefit of 2.081 and mean net return of 1.298 \$ L^{-1} for fuel production and rendered microwaves as promising method[15].

The recent advancements have shifted the focus of researchers towards use of microalgae as reliable source of biodiesel production [16]. In this context, Sivaramakrishnanet.al utilized microalgal lipids as useful feedstock for biodiesel and bioethanol production through novel biorefinery approach[17]. Similarly, Daroch, M et al., conducted comparative evaluation of various feedstocks for biodiesel production. They reported the use of alkali catalyst as the promising resource for conversion of algal oils to biodiesel [18]. Similarly, Mobinet al., found advantages of microalgae owing to its environmental friendly, a short cycle and a high doubling rate of approximately 24 hours [19]. Moreover, Sun et al., 2018 reported Chlorella Vulgaris, a green freshwater microalga, a better source in the manufacture of microalgae-based biodiesel because of its propensity to collect significant levels of lipids, up to 70% of dry cell weight[20].Similarly, Sakarika and kornaros found that the Oleic acid (C18:1), linoleic acid (C18:2), palmitic acid (C16:0), palmitoleic acid (C16:1), and stearic acid (C18:0) are the primary fatty acids generated by C. Vulgaris and are recommended for biodiesel synthesis [21]

All the research conducted so far has identified that algal growth is inherently affected by certain factors like pH, CO_2, light, salinity, temperature, media composition, nutritional deficiency, nitrogen, carbon, phosphate, aeration, and oxygen etc. [22, 23].However, for maximizing biomass and lipid productivity, the mentioned factors must be optimized. In this context, JC Nzayisenga et.al. investigated the effect of light density on growth of microalgae grown in wastewater. The reported that sustaining the light/dark cycle of microalgae is important factor as respiration proceeds even at night and could result in biomass and lipid loss[24]. Similarly, Edmundson et al., 2015 studied that throughout the night, 35 percent of biomass and 26% of triacyl glycerides are reduced, resulting in lipid productivity depletion. [25]. For efficiently controlling the issue, many scientists have proposed autotrophic daytime/heterotrophic night time culture as a reliable way to limit lipid loss at night by optimizing settings such as managing daytime and overnight temperatures [26].

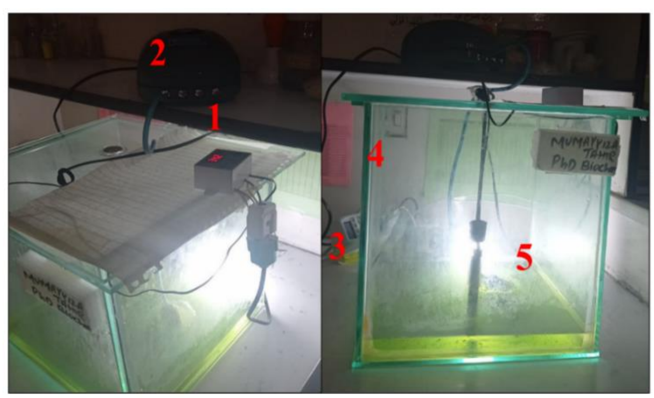
Process optimization using the traditional ways of conducting repetitive trials can be complex and time-consuming. Response Surface Methodology (RSM) is a statistical technique that has largely been used for the optimization of complex systems[27, 28]. In this process, based on experimental data, statistic-based models are developed that can interpret complex inter-relationships between operating parameters[29]. Subsequently, it facilitates the identification of optimum conditions for involved variables. Much research has been reported with valuable outcomes that made use of the very technique. Bajwa et al., 2019 investigated the optimization of certain physio-biochemical cellular components of chlorella pyrenoidosa for biodiesel production by using Box-Behnken response surface design methodology. They reported that FAME profile of Chlorella pyrenoidosa contained palmitic (C16:0, stearic acid), (C18:0), oleic (C18:1), linoleic (C18:2), and linolenic (C18:3) fatty acids, indicating that the chlorella pyrenoidosa has a favourable fatty acid profile that may be used to make biodiesel[30]. Similarly, Sirajunnisa Abdul Razack evaluated the statistical optimization of harvesting Chlorella vulgaris and identified 100 mg per liter bio flocculation concentration and 0.5 hours incubation time for maximum efficiency of 99.68% [31].

The world is getting reliable on computers owing to their accuracy and short time durations of solving complex problems. Nowadays, significant research is being done in the domain of Artificial Intelligence (AI)[32]. The large-scale implementation of AI in industrial and commercial prospects has necessitated the need for addressing the existing problems in the energy sector too. In this context, the application of Artificial Neural Networks for accurate predictions is most welcoming [33]. Vinoj Chamilka Liyanaarachchi designed ANN approach for optimizing growth conditions of C.V and found predicted and experimental values strongly correlated with R-square of 0.97[34].

In the light of the literature presented, much work has been done so far for optimizing the growth conditions of microalgae for high lipid extraction and biodiesel production. However, the combined use of the statistical tool and Artificial intelligence for developing accurate microalga growth predicting models and optimization of influencing varying factors has not been reported so far. In this work, the freshwater microalga – Chlorella Vulgaris growth has been predicted and optimized using ANN and RSM. First, the MATLAB NN toolbox was used for creating 12 different neural networks with different combinations of hidden layer neurons and training functions. Second, the best identified neural network predicted results were used for developing RSM quadratic model. It was used for studying the statistical interaction of light, temperature, days, growth medium, and photoperiod on optical density of *C.V.* Finally, the model was numerically optimized for getting the best combination of input variables in relation to maximum growth. Moreover, the obtained optimized parameters were validated through experimentation and absolute percentage error was checked. Thus, this work has developed the technique for optimizing growth conditions of *C.V.* through the collective use of RSM and ANN.

2 EXPERIMENTAL SETUP

The current study considers the water bodies within the premises of Punjab University Quaid-e-Azam Campus, Lahore. The water samples with enriched algal growth were collected in plastic bottles with sterilized media employing a glass pipette and sucker. The shaking incubator (SI6R-2 SHEL LAB) was utilized for maintaining the collected alga by designating the inside temperature to 25°C and constant stirring at 110 rpm. A glass photobioreactor was designed to investigate the effect of light and medium on algal growth. The experimental setup is shown in Figure 1. It comprises of a 10-liter glass tank filled to a depth of 3 liters with algal growth media and an aerator for facilitating the proper algal growth through constant stirring. A 13-watt LED bulb was used as a light source and a timer switch was incorporated for maintaining the desired photoperiod (Light/Dark cycle). The bulb was placed 15 cm vertically above the medium to eliminate possible temperature variations owing to illumination. Moreover, the entire glass tank was covered with aluminum foil to prevent the external light from seeping in the tank and affecting growth. The temperature inside the tank was adjusted between 22°C- 25°C by using an aquarium heater and was monitored using the thermostat.



1. Thermostat 2. Aerator 3. Timer Switch 4. Photobioreactor

5. Aquarium Heater

Fig. 1 Experimental Setup

2.1 ALGAL GROWTH MEDIA

Two Media i.e., Basal Bold Media (BBM) and BG-11 Media were used for cultivating and growth optimization of *Chlorella Vulgaris*. The composition of both mediums is comprehensively shown in Table 1. Unialgal culture of *Chlorella vulgaris* was obtained by repeated sub culturing in liquid and solid agar media. Moreover, morphology was validated by compound microscope and by comparing it with the published literature.

Table 1 Composition of mediums

Medium	Composition
Basal Bold Media (BBM)	$NaNO_3$ (250mg/L), K_2HPO_4 (75mg/L), $MgSO_4.7H_2O$ (75mg/L), $CaCl_2.2H_2O$ (25mg/L), KH_2PO_4 (175mg/L), NaCl (25mg/L),Alkaline EDTA solution (5.0g EDTA and 3.1g KOH in 100ml distilled water), Acidified Iron Solution 1mg/L (0.498g $FeSO_4$ and 0.1ml H_2SO_4 in 100ml distilled water), 1ml/L Trace metal solutions: $MnCl_2.4H_2O$ (1.44g/L), $ZnSO_4.7H_2O$ (8.82g/L), $(NH_4)_6Mo_7O_{24}.2H_2O$ (0.88g/L), $Co(NO_3)_2.6H_2O$ (0.49g/L), Cu $SO_4.5H_2O$ (1.57g/L). Final pH - 6.6.
BG-11 Media	$NaNO_3$ (1500mg/L), K_2HPO_4 (40mg/L), $MgSO_4.7H_2O$ (75mg/L), $CaCl_2.2H_2O$ (36mg/L), Citric acid (6mg/L), Trace metal solution 1ml/L (Ferric ammonium citrate 6g/L, Na_2.EDTA (1g/L), $MnCl_2$ (1.81g/L), $ZnSO_4.7H_2O$ (0.222g/L), $Na_2MoO_4.2H_2O$ (0.93g/L), $CuSO_4.5H_2O$ (0.08g/L). Final pH-7.5

2.2 DESIGN OF EXPERIMENT (DOE) AND ARTIFIFCAL NEURAL NETWORK MODELLING (ANNNM)

In this work, growth conditions for *Chlorella Vulgaris (C.V)* have been optimized using Response Surface Methodology (RSM) under the controlled design of experiment. RSM is a technique for studying complex systems from a statistical perspective. DESIGN EXPERT 13 was used for historical design whose detail attributes are enlisted in Table 2. The optical density (OD) of the C.V was constantly monitored for 15 days, with two media, light/dark cycles of 12 and 16 hours and two distinguished temperatures of 25 and 28 °C. The combination of all the stated parameters resulted in 60 experimental data points which were used for the Analysis of Variance (ANOVA) of OD and study of statistical parameters related to growth. The predicted results of RSM were validated through correlation coefficient and zoom-in plots over the entire data range.

Moreover, the RSM predicted resulted were utilized to design ANN models capable of predicting the growth of C.V. In the process, first, MATLAB NN toolbox was used for creating 12 different neural networks each with a varying number of neurons and training functions. The comprehensive description of the used literature is shown in Table 3. Second, the best ANN model was selected on the statistical performance capacity in terms of Mean Square Error (MSE) and correlation coefficient (R-square). Finally, the predicted results of the nominated network were processed for optimization and subsequent validation.

Table 2 Attributes of RSM historical design

Factors	Type	Levels	Extremes
A: Days	Numerical	15	1,15
B: Temperature (°C)	Numerical	2	25,28
C: Photoperiod (Hours)	Numerical	2	12,16
D: Media	Categorical	2	--

Table 3 Literature of ANN Models

Network literature	Training Function	Neurons
N1N9	TRAINBFG	N9:N20 Number of neurons in hidden layer
N2N10	TRAINBR	
N3N11	TRAINCGB	
N4N12	TRIANCGF	
N5N13	TRIANCGP	
N6N14	TRIANGD	
N7N15	TRAINGDM	
N8N16	TRAINGDA	
N9N17	TRAIGDX	
N10N18	TRAINLM	
N11N19	TRAINOSS	
N12N20	TRAINR	

3 RESULTS AND DISCUSSION

This section explains the ANOVA results for RSM modelling of *(C.V)* growth along with the statistical interactions of involved variables and response. Prior to analysis, appropriate transformation and model selection for analysis are integral as they relate to overall performance [35]. The fit summaries of the various models are shown in Table 4. Based on adjusted and predicted correlation coefficients of 0.8806 and 0.8645, the quadratic model is deemed appropriate. Moreover, the square root transformation with a constant of 0.04 was applied after the iterative selection of available options

Table 4 Fit Summaries of Models

Source	Sequential p-value	Adjusted R^2	Predicted R^2
Mean	< 0.0001		
Linear	< 0.0001	0.7096	0.6789
2FI	0.0006	0.7780	0.7335
Quadratic	< 0.0001	0.8806	0.8645

3.1 DIAGNOSTIC STUDY

ANOVA model for the growth of C.V needs to be comprehensively validated using statistical tools. This step is inevitable as any errors or defective data can be easily identified in the initial stages and removed before further processing the results. Among the plenty of diagnostic graphs available in Design Expert, three are shown in Figure 2 (a-c). Graph 3(a) shows the normal plot of externally studentized residuals against the normal probability. With the exception of few, all the data points are falling on the straight line which evidenced the desirable statistical performance. Similarly, the residuals vs. run plot in Figure 2(b) with the symmetrical pattern above and below the mean gives the confidence to safely employ model for optimization. Lastly, the predicted vs. actual plot of optical density of Chlorella Vulgaris in Figure 2(c)shows that both values are comparable over the entire experimental range.

Based on these effective diagnostic results, the model could be more extensively studied using the quadratic equation given below:

Optical Density = 1.12213 + 0.578406 * A + 0.0416124 * B + 0.021965 * C -0.0563446 * D + 0.178922 * AB + 0.12067 * AC - 0.0696444 * AD + 0.01245 * BC - 0.0103582 * BD + 0.03654 * CD- 0.361467 * A^2 + 0.43124 * B^2 +0.001234 * C^2

The prediction ability of the RSM model designed for growth could be also checked through plot over the entire data range, as shown in Figure 3**.** The yellow zoomed-in portion demonstrates a reasonable overlapping of actual and RSM predicted results. Therefore, the results from the diagnostic assisted us in validating the designed response surface model.

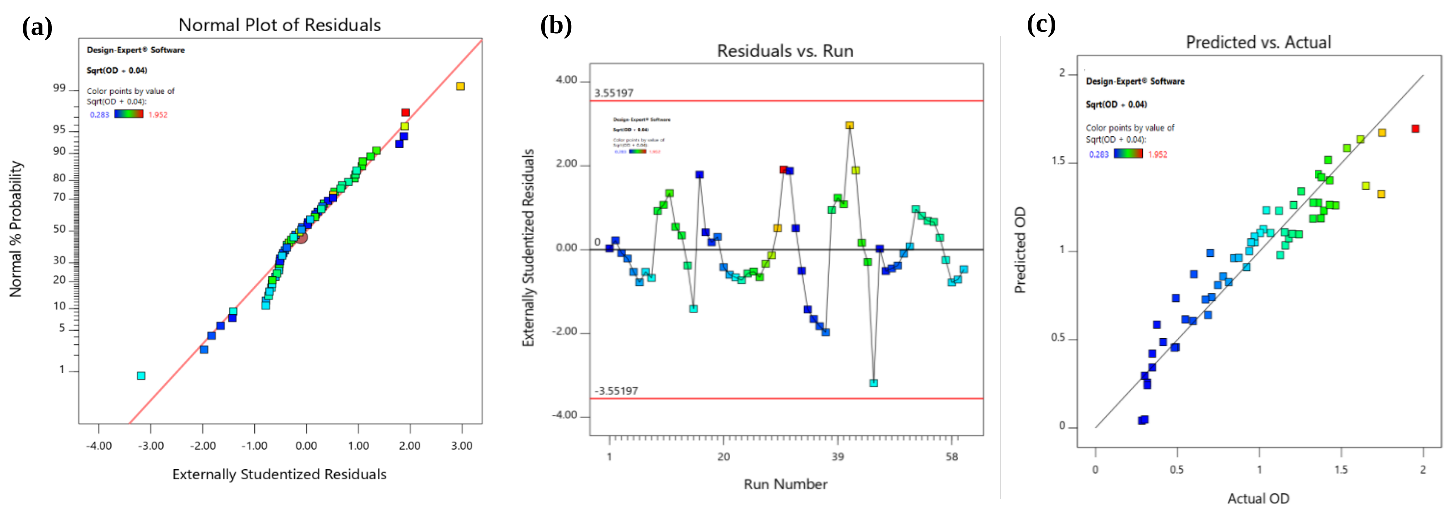


Fig.2 (a) Normal plot of residuals for OD (b) Residual vs. Run plot for OD and (c) Predicted vs. Actual plot for OD

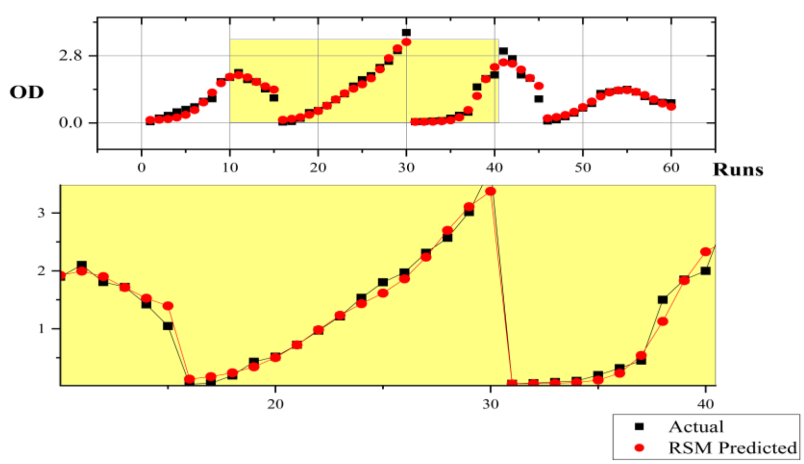


Fig.3 Comparison of actual and predicted results

3.2 RESPONSE SURFACES

In this section, the experimentally obtained results are studied using designed response surfaces. Moreover, the statistical interactions of input factors (A, B, C and D) and response (OD) is also comprehensively discussed. Figure 4 (a) shows the variation of C.V. growth for BG11 medium at a temperature of 25 ^{0}C. The color-coded surface demonstrates the rising-falling pattern over 15 days. Up to 11th day, the growth continued to increase and later experienced a sudden decline on the 12th day. Moreover, photoperiod is found significantly affecting the response and higher duration of light exposure appeared reasonable for maximizing growth. The dots above and below the surface designate the design points which are higher and lower than the reference data. Similarly, OD of basel bold medium (BBM) over the same test conditions is shown in Figure 4 (b). Although the earlier stated direct relations continue to exist, but the OD for BBM appears significantly lower than that for BG11.

The growth variations at a test temperature of 28 ^{0}C for BG11 and BBM are shown in Figure 5 (a) and (b), respectively. Once again, the variations are identified similar to those for lower temperature. However, the declining pattern of both cases is uniquely distinguished. The curve for BBM shows a greater downward curve than BG11; indicating that growth started to decline earlier for BBM when evaluated at par with BG11.

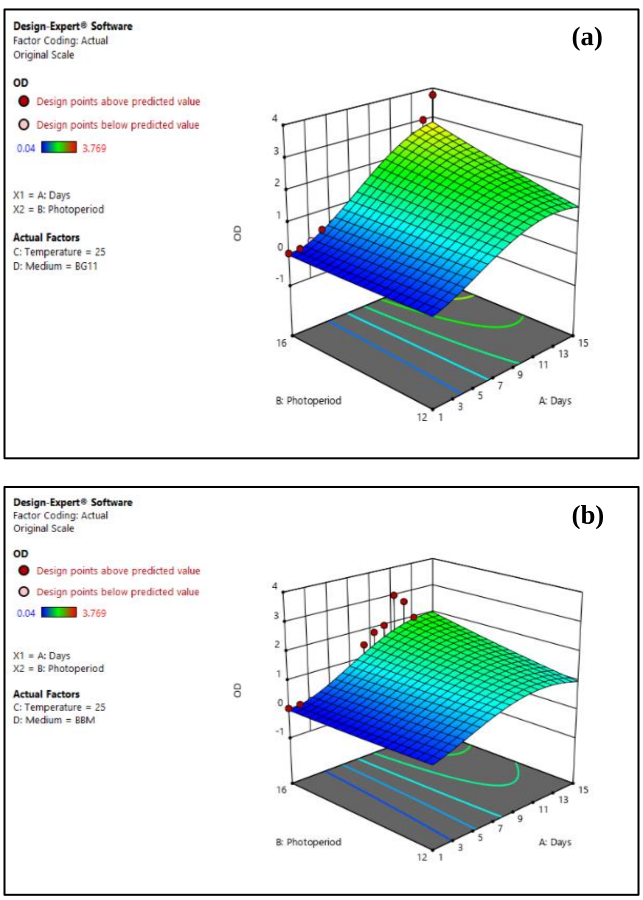


Figure 4 (a) OD variation with photoperiod and days for BG11 and 25 °Cand (b) OD variation with photoperiod and days for BBM and 25°C

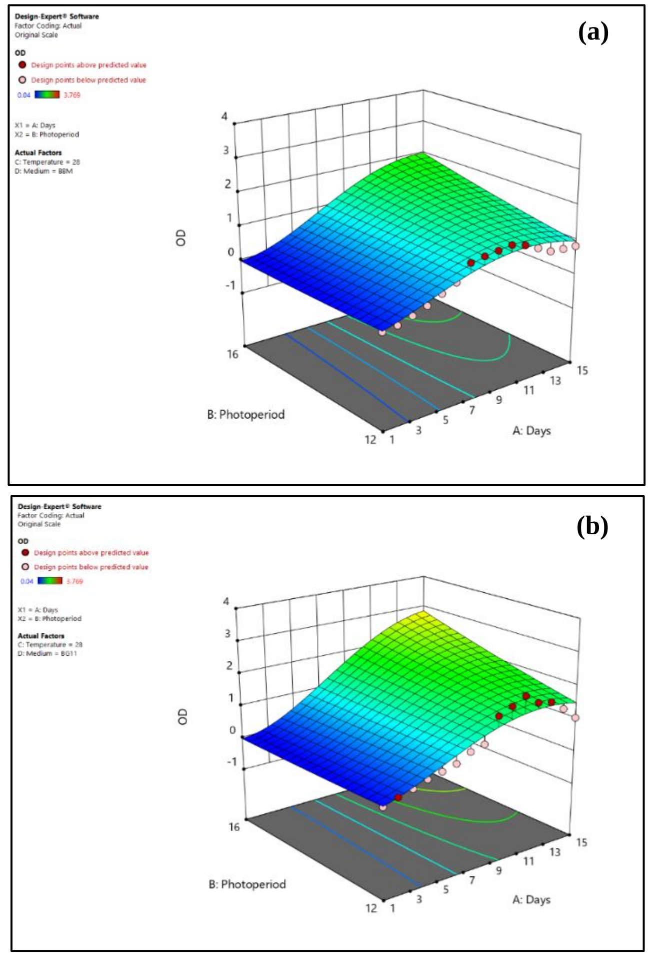


Figure 5 (a) OD variation with photoperiod and days for BG11 and 28 °C and (b) OD variation with photoperiod and days for BBM and 28 °C

3.3 APPLICATION OF ANN

This section utilizes artificial neural network for prediction of Chlorella Vulgaris growth. The process and results are explained in following sections.

3.3.1 PRE-PROCESSING OF EXPERIMENTAL DATA

The experimental input data (days, photoperiod, temperature and medium) and output (optical density) were normalized prior to ANN application in the range of 0 and 1. Doing so provides an easy way for the faulty data identification. The following empirical relations were used for normalizing the data sets. The results identified the two out-of-range data which were removed and not considered for further analysis.

$$\mathbf{I_N = \frac{I - least(I)}{Max(I) - Least(I)}} \quad (1)$$

$$\mathbf{O_N = \frac{O - Least(O)}{Max(O) - Least(O)}} \quad (2)$$

3.3.2 ARTIFICIAL NEURAL NETWORK MODELS (ANNMS)

The processing system of the human brain is the base concept of Artificial Neural Network[36, 37]. It assesses performance based on the given input conditions and the defined output variables. Over the years, more and more researchers resorted to this technique to solve and predict complex linear and nonlinear systems[37]. ANN structures are essentially composed of input, output and hidden layers. The neurons in the input and output layers are used for storing input and the target data. However, the central layer consists of processing neurons and serve the function of connecting different layers in the structure. The whole set of input and target data is first trained by the model and is later repetitively validated using specific algorithms until the desired performance parameters are achieved. Thus, efficient modeling of real-life problems is complex and requires iterative combinations of different models and algorithms for desirable results[32].

The ANN models were developed using four input layers, one hidden and the output layer for the present case. The non-numeric variable, growth medium for *C.V*, was coded in 0s and 1s; with 0 designated as BG11 while 1 representing BBM. Moreover, the network type used was feedforward backpropagation in combination with Trainlm training function, owing to its historically accurate results. The number of neurons in hidden layers were checked in the range of 9-20, starting from the low extreme. Similarly, the default training functions were also varied with each neuron number and consequently, network names were derived from the combination of these two. The whole set of experimentally obtained data was divided as follows: 70% training, 15% validation and 15% testing. The adapted criterion for training was ceasing when the error started escalating. The output of each network was exported and compared with the actual data on Mean Square Error and correlation coefficient as a base. The detailed neural network base modelling approach is shown in Figure 6, while the summary of characteristics associated with ANNMs are listed in Table 5.

Table 5 Characteristics of ANNMs

Characteristics	Description
Layers	Input: numeric (03), non-numeric (01) Hidden (01) Output (01)
Hidden Layer Neurons	Iterative selection in the range of 9-20
Network category	Feedforward backpropagation
Data Distribution	Training (70%), Validation (15%), Testing (15%)
Stopping Criterion	Mean Square error stars escalating
Statistical Evaluation Parameters	Mean Square Error (MSE) Coefficient of Determination

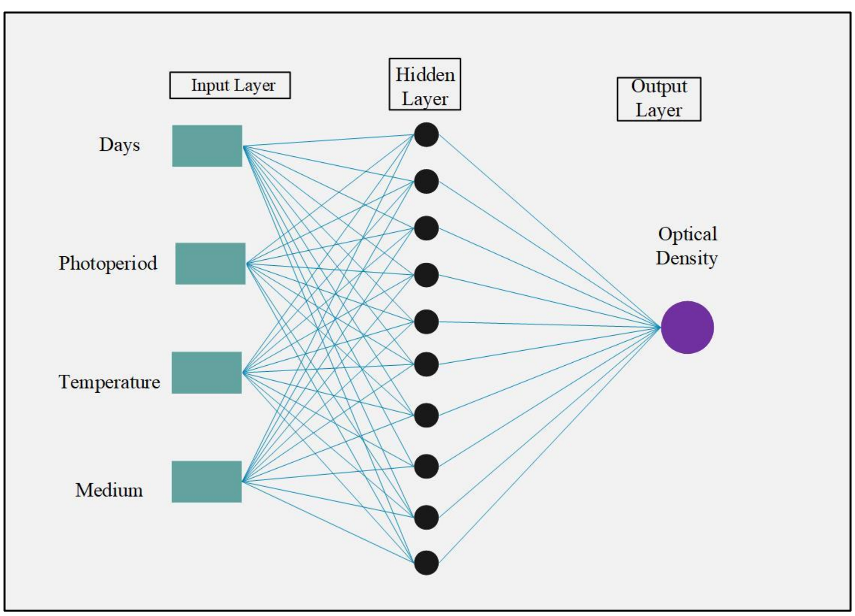


Fig. 6 Artificial Neural Network Base Model

3.4 ANN PREDICTIONS COMPARISON

This section presents the comparative statistical performance of twelve neural networks for growth prediction for Chlorella Vulgaris. Figure 7 shows the bar chart for correlation coefficients of different networks with incrementing neurons on the ordinate. For all the trained networks, the actual and predicted data relations are highly accurate except for the value of 0.77423, which is for the case of 12 neurons and the training function of TRAINCGF. The training algorithm TRAINCGP with thirteen processing neurons showed a distinguished correlation coefficient near to 1. This brings forward that for the accurate modelling of C.V. growth considering the given experimentally defined numerical ranges, N5N13 should be used for efficient results. Similarly, for further prediction ability assessment of ANNMs, MSEs were evaluated and plotted as cumulated percentage contribution in Figure 8. On the comparative scale, the most deviated results were for the case of N4N12, as shown by the light purple portion with a percentage of 49.29. Moreover, the best model (N5N13) selected previously is endorsed owing to the least square error, shown with the bulged part of the chart. The chart gives an important insight that models N6N14 and N9N17 also have reliable ability to predict growth as their comparative percentage of MSE are below 2%. However, for accuracy in real-life applications, the most accurate model should be considered.

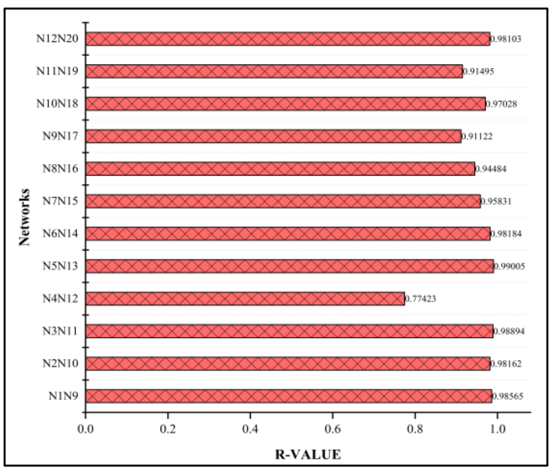


Fig.7 R-square values of ANNMs

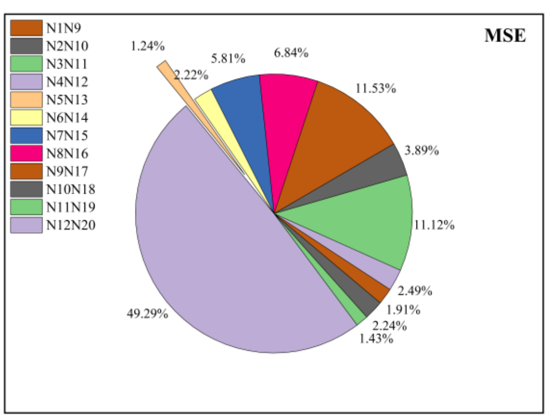


Fig.8 Mean square of ANNMs

The N5N13 model is now comprehensively investigated for further evaluation and point-to-point study variations in predicted and experimental data. The detailed analysis of selected ANNM is divided into two stages: in-training and post-training, for minimizing the possible errors. The summary of graphical results of in-training phase is shown in Figure 9. The gradient plot with epochs shows that the difference of actual and predicted data is constantly decreasing with increment along abscissa. The gradient is always desirable to be as low as low as possible and thus, the value of 0.0089113 at epoch number 300 is reliable and vouch for the efficiency of N5N13.

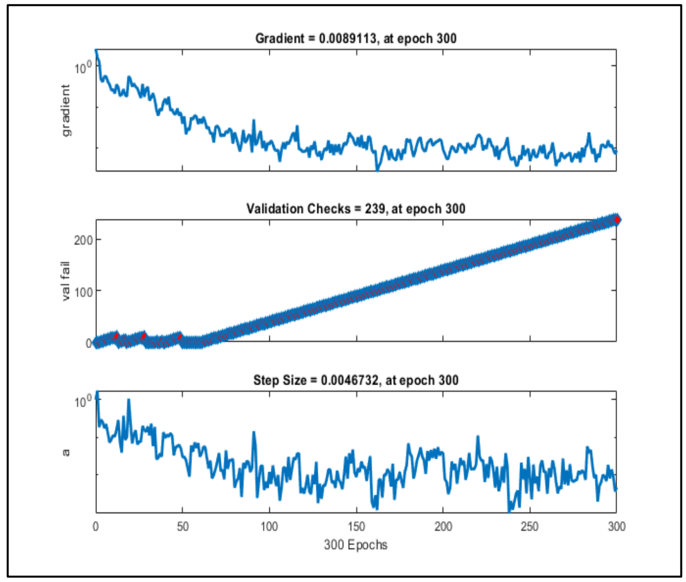


Fig.9 In-training results summary

Similarly, the post-training attributes are shown in Figure 10 (a-c). The results in this stage are more descriptive comparable to previous. In Figure 10 (a), the plotted squared deviation on the ordinate show that error for training, testing and validation are comparable and therefore, the chances of over-fitting of ANN model is potentially negated. Moreover, the best validation occurred at run 60 and after that this phase, red, green and blue lines continue to extend rightwards but in parallel relation with best-fitted line. Similarly, the statistical performance of three phases of ANNMs modelling are shown in Figure 10 (b). The coefficients of determination for training, testing, validation and overall model come out to be 0.98182, 0.99784,0.99894, and 0.98711, respectively. The R-square values sufficiently close to 1 demonstrate the Chlorella Vulgaris model's viability in the prospect of effectual prediction. Moreover, variations between predicted and actual results over the entire data range are shown in Figure 10 (c). The overlapping pattern for the data points 10-50 is zoomed and shown in the yellow highlighted region. The visual observation makes it evident that all actual and predicted results perfectly overlap in the selected range. A few individual circles and rectangles seem slightly deviated; however, they could be ignored due to the insignificant cumulative effect on overall performance. The thorough evaluation of N5N13 models during in-training and post-training stages governs that a developed neural network is worthy and could be used for accurate growth prediction.

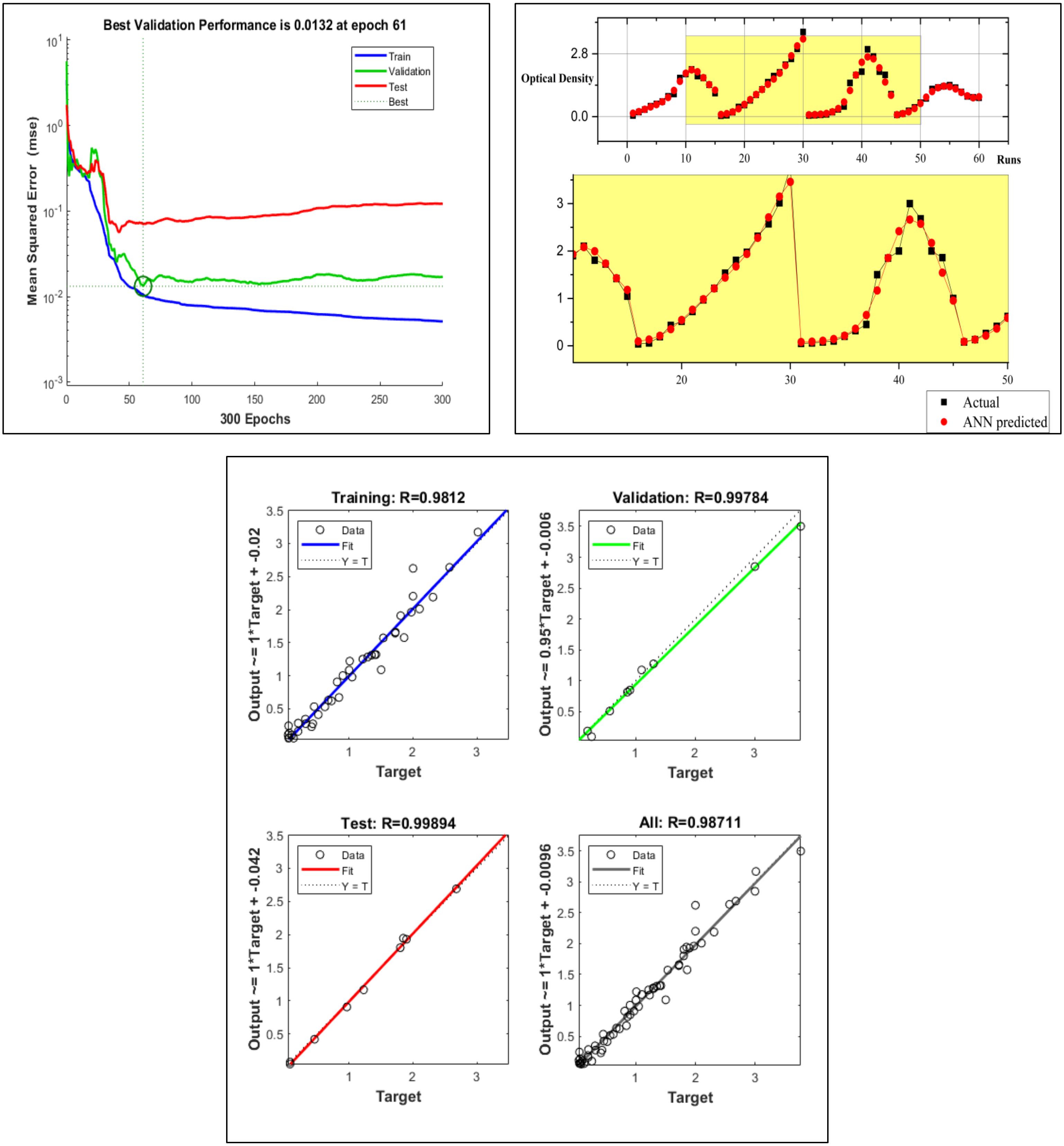


Fig. 10 (a) Validation performance (b) correlation coefficients of training, validation and testing and (c) Actual vs. predicted plot

4 OPTIMIZATION RESULTS

This optimization study aims at finding the appropriate growth conditions for Chlorella Vulgaris. The number of days, photoperiod, temperature and medium were design factors and optical density was the response. The optimization requires certain constraints to be defined for varying factors and responses. In the present scenario, the goal is to maximize growth under the given experimental conditions. The setup is shown in Table 6**.**

Table 6 Optimization setup

Name	Goal	Lower Limit	Upper Limit	Lower Weight	Upper Weight	Importance
A: Days	is in range	1	15	1	1	3
B: Photoperiod	is in range	12	16	1	1	3
C: Temperature	is in range	25	28	1	1	3
D: Medium	is in range	BG11	BBM	1	1	3
OD	maximize	0.04	3.769	1	1	3

The optimum growth conditions found were: 14th day, exposure to light and dark for 16 and 8 hours, temperature of 25 ^{0}C and BG11 media. The reliability of these optimized conditions could be validated using the concept of composite desirability and experimentation corresponding to obtained conditions. Composite desirability (D) is a unitless number in the range of 0-1 and stands for the degree of accuracy to which the design factors have collectively optimized the overall objective. The value in the proximity of 1 symbolizes the accuracy. In the present case, the composite desirability of 0.846768, shown in Figure 11, is near to 1, which is a testimonial to the fact that obtained optimized results are reliable and accurate.

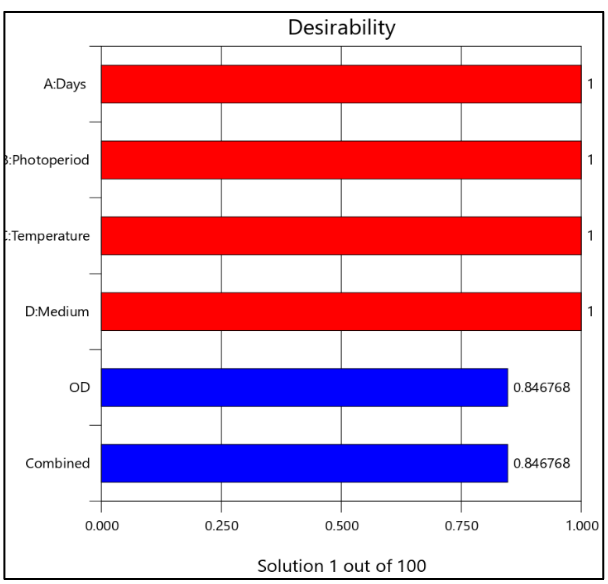


Fig.11 Desirability chart

5 CONCLUSION

This work has optimized the conditions for Chlorella Vulgaris growth through the integrated use of Artificial Neural Network predicted results and RSM-based numerical optimization. The results could be concluded as:

- The comparative analysis of fit summaries of various models deemed quadratic model as appropriate with an R-square value of 0.8806.
- Predicted vs. actual and residual plots accurately validate the selected quadratic model.
- Zoomed plot of RSM prediction showed a reasonable agreement between predicted and recorded results.
- Neural network N5N13 showed the least mean square error (1.24%) and maximum R-square value of 0.99005 among all networks.
- The coefficients of determination for training, testing, validation, and overall, the ANN model turns out to be 0.9812,0.99894,0.99784 and 0.98711 respectively.
- The best validation performance of ANN was recorded at the 61st epoch.
- Optimization identified 15th day, the temperature of 25 ^{o}C, the light cycle of 16 hours, and BG 11 medium for maximum OD of 2.862.
- The composite desirability (D) of 0.846768 validated the efficient multi-objective optimization.
- The experimental and optimized value showed an absolute percentage error of 3.92%.

The application of deep learning and the statistical concept has provided with the technique which could be effectively used for simultaneous data-based prediction and optimization and could save considerable amount of time and capital.

6 REFERENCES

[1] Huang, C.-W., et al., Cofiring characteristics of coal blended with torrefied Miscanthus biochar optimized with three Taguchi indexes. 2019. 172: p. 566-579.

[2] Marchetti, J., et al., Possible methods for biodiesel production. 2007. 11(6): p. 1300-1311.

[3] Usman, M., et al., Experimental assessment of regenerated lube oil in spark-ignition engine for sustainable environment. Advances in Mechanical Engineering, 2020. 12(7): p. 1687814020940451.

[4] Choi, J., et al., Impacts of local vs. trans-boundary emissions from different sectors on PM2. 5 exposure in South Korea during the KORUS-AQ campaign. 2019. 203: p. 196-205.

[5] Shahare, V.V., B. Kumar, and P. Singh, Biofuels for sustainable development: a global perspective, in Green technologies and environmental sustainability. 2017, Springer. p. 67-89.

[6] Szybist, J.P., et al., Biodiesel combustion, emissions and emission control. 2007. 88(7): p. 679-691.

[7] Usman, M. and N. Hayat, Lubrication, emissions, and performance analyses of LPG and petrol in a motorbike engine: A comparative study. Journal of the Chinese Institute of Engineers, 2020. 43(1): p. 47-57.

[8] Mahmudul, H., et al., Production, characterization and performance of biodiesel as an alternative fuel in diesel engines–A review. Renewable and Sustainable Energy Reviews, 2017. 72: p. 497-509.

[9] Ye, Y., et al., Optimizing culture conditions for heterotrophic-assisted photoautotrophic biofilm growth of Chlorella vulgaris to simultaneously improve microalgae biomass and lipid productivity. 2018. 270: p. 80-87.

[10] William, P. and L.J.E.E.S. Laurens, Microalgae as biodiesel and biomass feedstock: review and analysis of the biochemistry, energetic and economics. 2010. 3: p. 554-590.

[11] Correa, D.F., et al., Towards the implementation of sustainable biofuel production systems. 2019. 107: p. 250-263.

[12] Tiwari, A.K., et al., Biodiesel production from jatropha oil (Jatropha curcas) with high free fatty acids: an optimized process. 2007. 31(8): p. 569-575.

[13] Foteinis, S., et al., Used-cooking-oil biodiesel: Life cycle assessment and comparison with first-and third-generation biofuel. 2020. 153: p. 588-600.

[14] Meng, X., G. Chen, and Y. Wang, Biodiesel production from waste cooking oil via alkali catalyst and its engine test. Fuel processing technology, 2008. 89(9): p. 851-857.

[15] Mohammadshirazi, A., et al., Energy and cost analyses of biodiesel production from waste cooking oil. Renewable and sustainable energy reviews, 2014. 33: p. 44-49.

[16] Hossain, A.S., et al., Biodiesel fuel production from algae as renewable energy. American journal of biochemistry and biotechnology, 2008. 4(3): p. 250-254.

[17] Sivaramakrishnan, R. and A.J.F. Incharoensakdi, Utilization of microalgae feedstock for concomitant production of bioethanol and biodiesel. 2018. 217: p. 458-466.

[18] Daroch, M., S. Geng, and G.J.A.E. Wang, Recent advances in liquid biofuel production from algal feedstocks. 2013. 102: p. 1371-1381.

[19] Mobin, S.M., H. Chowdhury, and F.J.E.P. Alam, Commercially important bioproducts from microalgae and their current applications–A review. 2019. 160: p. 752-760.

[20] Sun, X.-M., et al., Microalgae for the production of lipid and carotenoids: a review with focus on stress regulation and adaptation. 2018. 11(1): p. 1-16.

[21] Sakarika, M. and M.J.B.t. Kornaros, Chlorella vulgaris as a green biofuel factory: comparison between biodiesel, biogas and combustible biomass production. 2019. 273: p. 237-243.

[22] Oziel, L., et al., Environmental factors influencing the seasonal dynamics of spring algal blooms in and beneath sea ice in western Baffin Bay. 2019. 7.

[23] Schnurr, P.J. and D.G. Allen, Factors affecting algae biofilm growth and lipid production: A review. Renewable and Sustainable Energy Reviews, 2015. 52: p. 418-429.

[24] Nzayisenga, J.C., et al., Effects of light intensity on growth and lipid production in microalgae grown in wastewater. 2020. 13(1): p. 1-8.

[25] Edmundson, S.J. and M.H.J.A.r. Huesemann, The dark side of algae cultivation: characterizing night biomass loss in three photosynthetic algae, Chlorella sorokiniana, Nannochloropsis salina and Picochlorum sp. 2015. 12: p. 470-476.

[26] Qi, F., et al., Improving productivity and quality of biodiesel from Chlorella vulgaris SDEC-3M through customized process designs. 2016. 129: p. 100-107.

[27] Khuri, A.I. and S. Mukhopadhyay, Response surface methodology. Wiley Interdisciplinary Reviews: Computational Statistics, 2010. 2(2): p. 128-149.

[28] Giovanni, M., Response surface methodology and product optimization. Food technology, 1983.

[29] Hill, W.J. and W.G. Hunter, A review of response surface methodology: a literature survey. Technometrics, 1966. 8(4): p. 571-590.

[31] Bajwa, K., et al., Response surface methodology as a statistical tool for optimization of physio-biochemical cellular components of microalgae Chlorella pyrenoidosa for biodiesel production. 2019. 9(5): p. 1-16.

[32] Razack, S.A., et al., Statistical optimization of harvesting Chlorella vulgaris using a novel bio-source, Strychnos potatorum. Biotechnology Reports, 2015. 7: p. 150-156.

[33] Bartman, J., Z. Gomółka, and B. Twaróg, ANN training–the analysis of the selected procedures in Matlab environment. Computing in science and technology, 2015: p. 88-101.

[33] Teng, S.Y., et al., Catalytic thermal degradation of Chlorella vulgaris: Evolving deep neural networks for optimization. 2019. 292: p. 121971.

[34] Liyanaarachchi, V.C., et al., Artificial neural network (ANN) approach to optimize cultivation conditions of microalga Chlorella vulgaris in view of biodiesel production. Biochemical Engineering Journal, 2021: p. 108072.

[35] Bhardwaj, B., R. Kumar, and P.K. Singh, An improved surface roughness prediction model using Box-Cox transformation with RSM in end milling of EN 353. Journal of Mechanical Science and Technology, 2014. 28(12): p. 5149-5157.

[36] Usman, M., et al., Response Surface Methodology Routed Optimization of Performance of Hydroxy Gas Enriched Diesel Fuel in Compression Ignition Engines. Processes, 2021. 9(8): p. 1355.

[37] Usman, M., et al., Artificial Neural Network Led Optimization of Oxyhydrogen Hybridized Diesel Operated Engine. Sustainability, 2021. 13(16): p. 9373.

7 AUTHORS

First Author – Mumayyiza Tahir, Ph.D. Scholar, School of Biochemistry and Biotechnology.

2nd Author – Prof. Dr. Muhammad Imtiaz Shafiq, P.hD, School of Biochemistry and Biotechnology

Correspondence Author – Mumayyiza Tahir, Ph.D. Scholar, School of Biochemistry and Biotechnology,