Mathematics

Research article

Comparative analysis of feed-forward neural network and second-order polynomial regression in textile wastewater treatment efficiency

Ali S. Alkorbi[1], Muhammad Tanveer[2], Humayoun Shahid[2], Muhammad Bilal Qadir[3], Fayyaz Ahmad[2,*], Zubair Khaliq[4,*], Mohammed Jalalah[5,6], Muhammad Irfan[5], Hassan Algadi[5,6] and Farid A. Harraz[1,6,*]

[1] Department of Chemistry, Faculty of Science and Arts at Sharurah, Najran University, Sharurah 68342, Saudi Arabia

[2] Department of Applied Sciences, National Textile University, Faisalabad, 37610, Pakistan

[3] Department of Textile Engineering, National Textile University, Faisalabad, 37610, Pakistan

[4] Department of Materials, National Textile University, Faisalabad, 37610, Pakistan

[5] Department of Electrical Engineering, College of Engineering, Najran University, Najran 11001, Saudi Arabia

[6] Advanced Materials and Nano-Research Centre (AMNRC), Najran University, Najran 11001, Saudi Arabia

* **Correspondence:** Email: fayyaz.ahmad@ntu.edu.pk; zubntu@yahoo.com; faharraz@nu.edu.sa; Tel: +92-3331665861.

Abstract: This study refines a single-layer Feed-Forward Neural Network (FFNN) for the treatment of textile dye wastewater, concentrating on percentage decolorization (%DEC) and percentage chemical oxygen demand (%COD) reduction. The optimized neural network configuration comprises four input and one output neuron, fine-tuned based on the mean squared error (MSE). The training phase demonstrates a consistent MSE decline, reaching its lowest at epoch 209 for %DEC and epoch 34 for %COD, with corresponding MSEs of 1.799×10^{-5} and 1.4×10^{-3}, respectively. The maximum absolute errors for %DEC and %COD were found to be 4.0787 and 2.4486, while the mean absolute errors were 0.4821 and 0.7256, respectively. In contrast to second-degree polynomial regression, the FFNN model exhibits enhanced predictive accuracy, as indicated by higher R^2 values of 0.99363 for %DEC and 0.99716 for %COD, and reduced error metrics.

1. Introduction

Water, a vital element for life and crucial in numerous domestic and commercial processes, is paradoxically one of the most exploited natural resources, disregarding its utmost importance. The exploitation arises from urbanization, rapid industrialization, improved living standards, and population growth, resulting in unregulated and excessive worldwide water consumption. As a result, the global community confronts an urgent water crisis, as the quantity and quality of water face compromises. Additionally, urbanization encroaches upon natural resources, with industries like textiles releasing effluents containing harmful components, notably colors, endangering water bodies and further diminishing water availability and safety. The significance of this crucial matter necessitates prompt action to tackle it, ensuring the sustainable management and preservation of water resources for the well-being of future generations.

The extensive use and poor removal rate during aerobic waste treatment make synthetic dyes a severe environmental threat. The textile industry alone contributes to this issue by utilizing approximately 8×10^5 tonnes of synthetic dyes among the over 10,000 distinct types of dyes produced globally [1]. The complex and ever-changing mixture of pollutants, including dyes responsible for color and organic load, disrupts the biological equilibrium of the receiving water system [2]. The occurrence of this disruption can be attributed to the substantial presence of organic compounds and heavy metals found in the dyes, leading to approximately 93% of the water used in textile production being discharged as colored effluent [3, 4]. In addition, the nonbiodegradable nature of dyes used in textile dye baths poses a severe environmental risk. The unpleasant color of wastewater negatively affects aquatic organisms and compromises the water's ability to sustain life, ultimately disturbing the entire aquatic environment and food chain [5]. To address this challenge, it is crucial to promote existing processes and explore innovative approaches to effectively decolorize a mixture of colors rather than solely focusing on individual dye solutions in textile effluents.

Researchers have devised various physiochemical methods to eliminate color efficiently and other pollutants found in effluents from textile dyeing operations. These techniques include adsorption, chemical oxidation and reduction, chemical precipitation, flocculation, and photolysis [6–8]. However, these techniques often suffer from inefficiency, high costs, and side effects such as the generation of significant sludge and byproducts. Moreover, they may only be universally effective in degrading some types of dyes [9, 10]. Consequently, the focus of researchers has shifted towards biological treatments, which offer the advantage of lower operational costs [11, 12]. The remarkable capability of various microorganisms, including bacteria, fungi, and actinomycetes, to decolorize dyes has been discovered [13, 14]. Among them, white rot fungus has emerged as one of the extensively studied bacteria due to its remarkable dye-decolorizing properties, ability to thrive in challenging environmental conditions, and production of multiple extracellular enzymes [15].

Sathian et al. [16] conducted a study where they utilized pleurotus floridanus in a batch reactor to treat textile dye effluent. The main objective was to optimize the %DEC process by studying the influence of different process variables, such as pH, temperature, agitation speed, and dye wastewater concentration. To analyze the impact of these factors, researchers utilized the response surface methodology and conducted optimization research. To streamline the process and reduce the number of experiments, they implemented the box-Behnken design, which allows for fitting a quadratic surface and facilitates the analysis of parameter interactions and effective parameter optimization [16].

In order to mathematically describe the response as a function of continuous factors, a regression design was utilized to estimate the model parameters accurately. The study relied on the methodology proposed by Sathian et al. to determine the second-order polynomial response function [16].

$$Y(x) = \beta_0 + \sum_{i=1}^{k}\beta_i x_i + \sum_{i=1}^{k}\beta_{ii} x_i^2 + \sum_{i=1, i<j}^{k-1}\sum_{j=2}^{k}\beta_{ij} x_i x_j,$$

here, Y represents the predicted response, while β_i and β_{ij} are coefficients estimated from the regression analysis.

The experimental settings for treating textile dye wastewater utilizing pleurotus floridanus, based on the box-Behnken design, are provided in Table 1. The authors [16] presented the following response surface models for the %DEC and %COD reduction:

$$\begin{aligned} Y_{\%DEC} =& 58.62 + 0.09x_1 - 0.59x_2 + 3.63x_3 + 14.07x_4 + 0.70x_1x_2 - 1.95x_1x_3 - 2.67x_1x_4 \\ & + 2.43x_2x_3 - 0.20x_2x_4 - 1.28x_3x_4 - 10.32x_1^2 - 2.02x_2^2 - 2.40x_3^2 - 4.30x_4^2, \\ Y_{\%COD} =& 69.35 + 0.30x_1 - 0.54x_2 + 2.68x_3 + 13.95x_4 - 0.80x_1x_2 - 1.75x_1x_3 - 3.45x_1x_4 \\ & + 2.83x_2x_3 - 0.80x_2x_4 - 1.55x_3x_4 - 9.92x_1^2 - 2.18x_2^2 - 1.88x_3^2 - 4.72x_4^2, \end{aligned} \tag{1.1}$$

where x_1, x_2, x_3, and x_4 represent the %COD values of the process variables (i.e., pH, temperature (°C), agitation speed (rpm), and wastewater concentration, respectively). Machine learning algorithms surpass traditional prediction methods [17], like ordinary least square methods. These algorithms are flexible enough to model the nonlinear phenomena hidden in the data. Machine learning algorithms are highly versatile due to their predictive power, finding use across various applications such as detecting depression through data analytics, forecasting gross domestic product, analyzing fractal-fractional mathematical models, and controlling the spread of waterborne diseases [18–21]. Using the FFNN, we want to make the prediction models for %DEC and %COD.

Table 1. Range and level of process variables [16].

Variables	%COD	Levels		
		−1	0	+1
pH	x_1	5	7	9
Temperature(°C)	x_2	23	28	33
Agitation speed (rpm)	x_3	100	150	200
Dye wastewater concentration	x_4	Raw	1:1	1:2

This study presents an optimized single-layer feedforward neural network (FFNN) model that excels in predicting %DEC and %COD in textile wastewater treatment. The model showcases exceptional accuracy, achieving R^2 values as high as 0.9998 in training and 0.9838 in testing phases. The log-sigmoid (`logsig`) activation function notably stands out, delivering the lowest maximum absolute error of 4.0787 and a mean absolute error of 0.4821 for %DEC predictions, with similarly impressive results for %COD. The robustness of the FFNN model is further reinforced through 10-fold cross-validation, consistently yielding low mean absolute errors between 0.4–1 for %DEC and 0.4–0.9 for %COD across different folds.

Sensitivity analysis, conducted via input perturbation, highlights the relative influence of each input variable on the model's predictions, providing insights of model numerical stability. When juxtaposed with polynomial regression models, the FFNN demonstrates superior capabilities, evident from its higher R^2 values for both training (0.99941 vs 0.9714 for %DEC; 0.99446 vs 0.9747 for %COD) and testing (0.99363 vs 0.8354 for %DEC; 0.99716 vs 0.9495 for %COD). Moreover, the FFNN model shows lower maximum and mean absolute errors, underscoring its enhanced predictive precision.

Contour plots are effectively utilized to visualize the interactive effects of variables, offering a comprehensive view of their combined impact on treatment efficacy. In summary, this study leverages meticulous performance metrics, sensitivity analyses, and rigorous validation techniques to establish the optimized FFNN architecture as a highly accurate and reliable tool for predicting key parameters in textile wastewater treatment.

The current study's goal is to unveil a feed-forward neural network as a precise and dependable machine learning model to aid in treating textile dye wastewater. This is achieved by evaluating two key outputs: percentage of %DEC and %COD reduction. The feed-forward neural network's superior accuracy is highlighted through the comparison of its R^2-values and mean absolute errors against those of traditional second-order polynomial regression models. Our findings illustrate that neural network models possess greater flexibility in identifying the nonlinear patterns embedded within the data, in contrast to traditional polynomial regression models which exhibit relative weakness in this aspect. This conclusion is supported by the performance metrics we have reported.

2. Feed-forward neural network

The feedforward neural network (FFNN) architecture is inspired by the human brain's complex neural structure. Like the brain's interconnected neurons, FFNN consists of processing units, or neurons, organized in layers and connected in parallel. Each connection's strength is determined by weights, and in an FFNN, the output of one layer seamlessly flows forward as the input to the next, without any feedback loops. This network structure comprises an input layer with independent variables, an output layer with dependent variables, and hidden layers that function as feature detectors. Despite the potential for multiple hidden layers, the universal approximation theory suggests that even a single-layered network with ample neurons can accurately represent any input-output mapping [22, 23].

The FFNN undergoes a training phase where it learns to predict output variables from given input-output pairs. This process involves adjusting the connection strengths, or weights and biases, between neurons across layers, transforming input signals into desired outputs. An activation function, applied to the weighted sum of inputs, aids in minimizing the prediction error. The FFNN's mathematical representation can be expressed as:

$$Y = \mathbf{W}^{(2)}{}_{n \times 1} h \left(\mathbf{W}^{(1)}{}_{n \times q} \mathbf{X}^T{}_{q \times m} + \mathbf{b}^{(1)}_{n \times 1} \right) + \mathbf{b}^{(2)}_{1 \times 1}, \tag{2.1}$$

where Y and X correspond to output and input vectors, respectively. $\mathbf{W}^{(r)}$ and $\mathbf{b}^{(r)}, (r = 1, 2)$ denote weights and biases, with m representing input variables, and n the number of hidden layer neurons. The function h symbolizes the activation function at the hidden layer, typically including the log-sigmoid,

hyperbolic tangent sigmoid, and linear transfer function [24].

$$h(z) = \texttt{logsig}(z) = \frac{1}{1 + e^{-z}}, \qquad\qquad 0 \leq h(z) \leq 1, \qquad (2.2)$$

$$h(z) = \texttt{tansig}(z) = \frac{2}{1 + e^{-2z}} - 1, \qquad\qquad -1 \leq h(z) \leq 1, \qquad (2.3)$$

$$h(z) = \texttt{purelin}(z) = z, \qquad\qquad -\infty \leq h(z) \leq \infty. \qquad (2.4)$$

Prior to training, scaling inputs and targets to a consistent range is crucial. Input values X_i are scaled to X_{scaled} within $[-1, 1]$:

$$X_{scaled} = -1 + \frac{2(X_i - X_{min})}{X_{max} - X_{min}}, \qquad (2.5)$$

where X_{min} and X_{max} are the minimum and maximum input values. The dataset is divided into training and testing samples, with a typical split of 70%–30% or 80%–20%. When including validation, the 30% test data can be equally split into two parts of 15% each. This setup ensures the network is trained effectively and tested for accuracy.

The optimization of the single-layer feedforward neural network (FFNN) was meticulously conducted through an iterative trial-and-error methodology. The selection criteria for the optimal FFNN architecture hinged on maximizing the coefficient of determination (R^2) and minimizing the mean square error (MSE), as defined in Eqs (2.6) and (2.7):

$$R^2 = 1 - \frac{\sum_{i=1}^{N} \left(y_i^{\text{FFNN}} - y_i^{\text{exp}} \right)^2}{\sum_{i=1}^{N} \left(y_i^{\text{FFNN}} - \overline{y}^{\text{exp}} \right)^2}, \qquad (2.6)$$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} \left(y_i^{\text{FFNN}} - y_i^{\text{exp}} \right)^2. \qquad (2.7)$$

Here, N denotes the total data count, y_i^{exp} and y_i^{FFNN} are the normalized experimental and FFNN predicted data vectors, respectively, and $\overline{y}^{\text{exp}}$ is the experimental data mean. Error minimization within the network typically employs error backpropagation, demonstrated through the update rule:

$$x_{k+1} = x_k - \gamma_k g_k,$$

where x_k represents current weights and biases, g_k is the gradient, and γ_k is the learning rate. Adjustments in network parameters are predicated on the discrepancies between training outputs and actual outcomes. Figure 1 illustrates a standard backpropagation process.

To fine-tune the FFNN, various factors were evaluated, encompassing:

1) Training algorithms: Levenberg-Marquardt with Bayesian regulation and gradient descent with momentum,
2) Activation functions: *tansig*, `logsig`, and *purelin*.
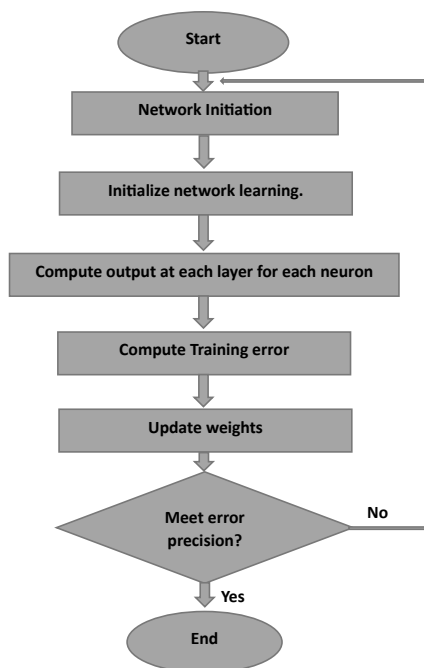
Figure 1. Flow chart diagram of the backpropagation process.

The dataset was partitioned into training (70%), validation (15%), and testing (15%) subsets. In the case of the Bayesian Regulation (*trainbr*) algorithm, the validation set was excluded, and a 80%-20% training-testing split was adopted. The MATLAB® software version 2023a was utilized for the MSE optimization process on a laptop equipped with an Intel® Xeon® E-2176M CPU at 2.70GHz and 16.0 GB of RAM, running on a 64-bit Windows 11 Pro system. A genetic algorithm ascertained the most efficacious factor combination for FFNN performance.

The Bayesian regularization algorithm (*trainbr*), particularly paired with the `logsig` activation function, yielded remarkable results. The optimal configuration, comprising four hidden neurons and the `logsig` function, minimized prediction errors, indicating that even a single hidden neuron can be effective. Figure 2 depicts the optimized FFNN model structure.
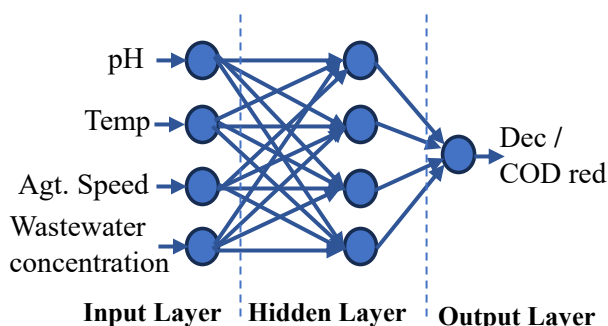
Figure 2. The architecture of the optimized FFNN model.

2.1. Neural network training and testing performance

The mean squared error (MSE) is an effective objective function for optimizing the selection of weights and biases. However, the optimal solution achieved can vary based on the initial conditions, potentially converging to local optima. As depicted in Figure 3, the MSE values exhibit a declining trend with each iteration. Specifically, the training phase MSE for the parameters under investigation diminishes consistently with the number of epochs, reaching a nadir at epoch 209 for %DEC and epoch 34 for %COD. Training and testing curves corroborate the mitigation of overfitting, evidenced by the testing curve's interruption of overfitting progression. Both curves display a synchronous decrease over successive epochs, suggesting a parallel improvement in model performance. Optimal validation performances for %DEC and %COD were attained at 1.799×10^{-5} and 0.0014 respectively, corresponding to epochs 209 and 34.
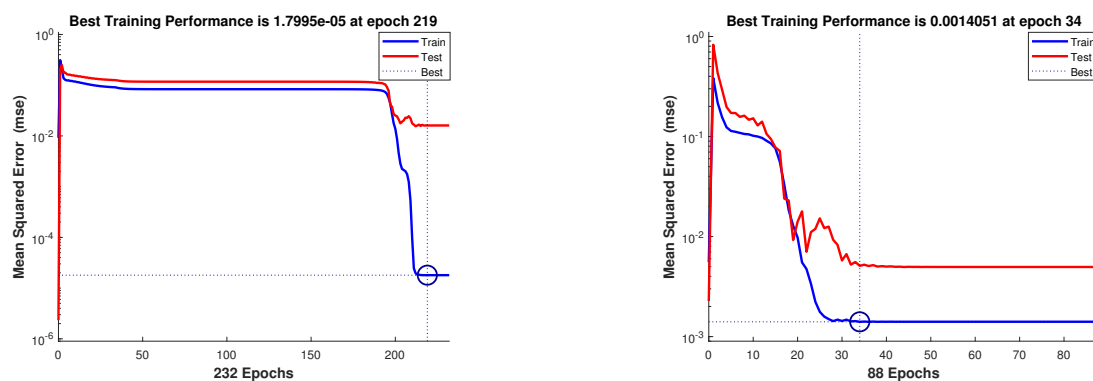
(a) %DEC training performance measurement with mean squared error.

(b) %COD reduction training performance measurement with mean squared error.

**Figure 3.** Analyzing training performance using mean squared error.

The error histograms, represented in Figures 4 and 5, illustrate the deviation between target and predicted values post-training of the neural network, with each bin denoting a range of error magnitudes. The y-axis quantifies the count of samples within each error range. The maximum observed error for %DEC was 4.0787, with an average error of 0.4821, an average percentage error of 1.0802%, and a maximum percentage error of 11.2448%. In the case of %COD reduction, the respective errors were a maximum of 2.4486, an average of 0.7256, an average percentage of 1.1460%, and a maximum percentage error of 3.4946%. The comparative analysis of the anticipated versus experimental data is graphically presented in Figure 6. These figures elucidate the close congruence between the FFNN predictions and the experimental results, visually underscoring the model's predictive accuracy.

2.2. Comparison of activation functions in FFNN

To examine the impact of different activation functions on the performance of a feed-forward neural network, three specific functions were selected: `logsig`, `tansig`, and `purelin`. The `logsig` function, with its sigmoidal response, is ideal for binary outputs, whereas `tansig` produces outputs from -1 to 1, beneficial for gradient-based learning methods. The `purelin`, being a linear function, is typically used for continuous outcomes. Their effects are graphically represented in Figure 7.
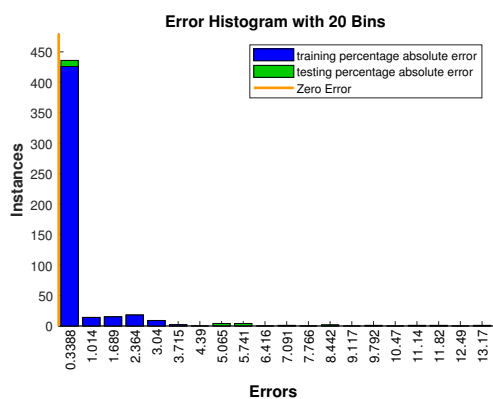
Error Histogram with 20 Bins

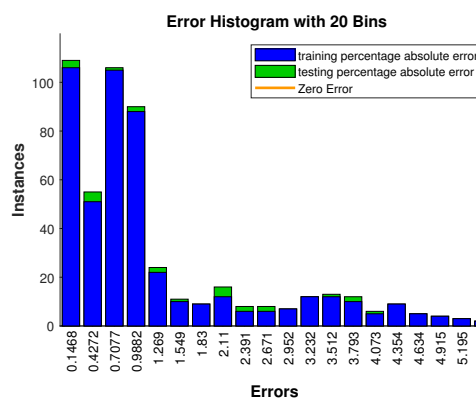(a) Absolute error histogram (%DEC).

Error Histogram with 20 Bins
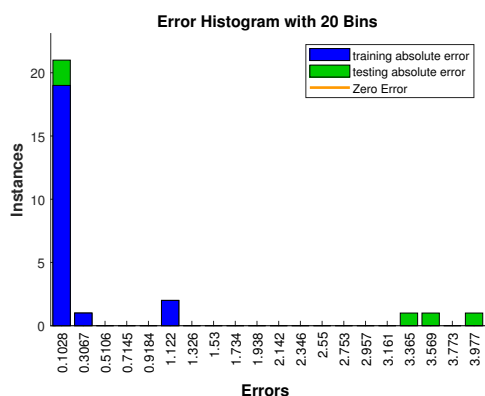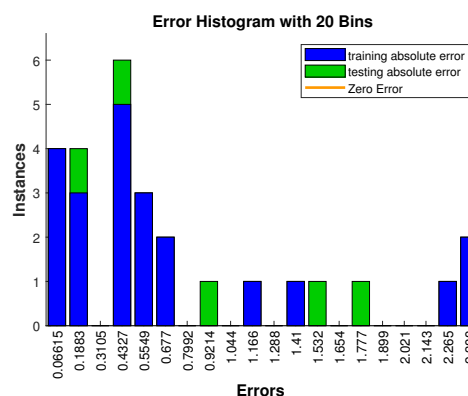
(b) Absolute error histogram (%COD).

Figure 4. Absolute error histogram for performance evaluation.

Error Histogram with 20 Bins

(a) Percentage absolute error histogram (%DEC).
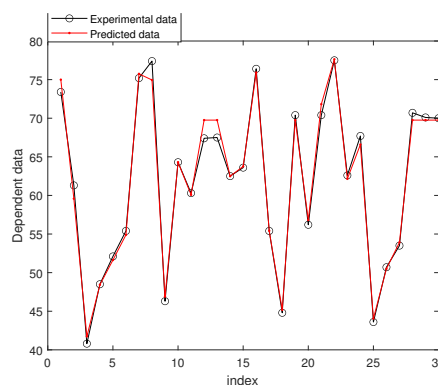
Error Histogram with 20 Bins

(b) Percentage absolute error histogram (%COD).

Figure 5. Absolute percentage error histogram for error distribution.
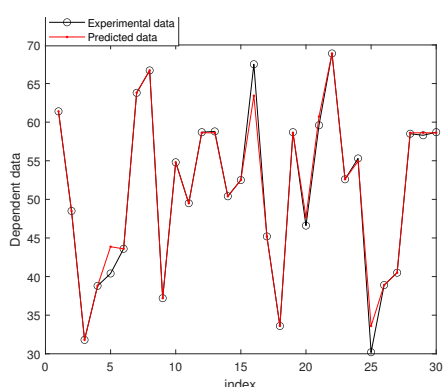
(a) %DEC prediction comparison: experimental vs. model.

(b) %COD reduction prediction comparison: experimental vs. model.

Figure 6. Graphical representation of experimental and predicted data.

(a) `logsig`   **(b)** `tansig`   **(c)** `purelin`
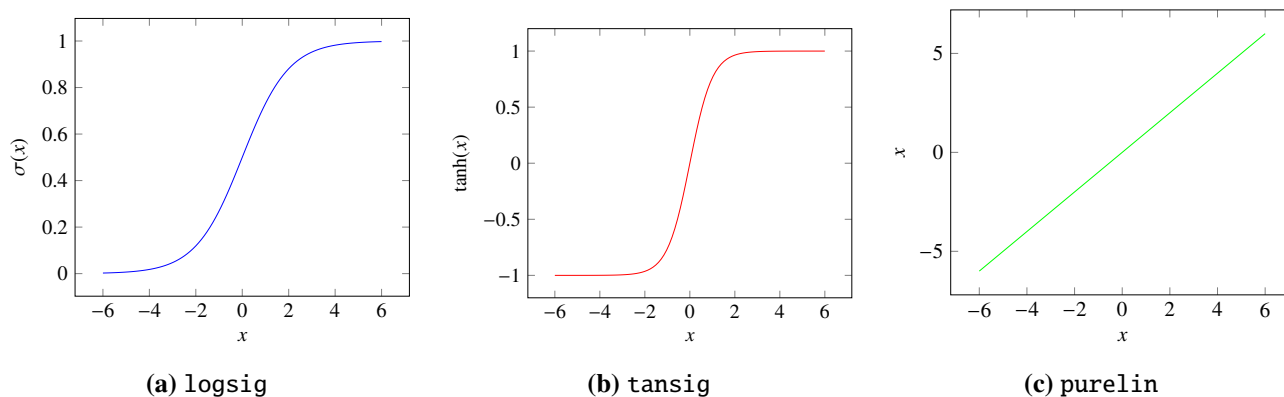
Figure 7. Comparative plots of neural network activation functions: logistic sigmoid (`logsig`), hyperbolic tangent (`tansig`), and linear transfer (`purelin`).

The performance of FFNN with different activation functions `logsig` , `tansig`, and `purlin` was evaluated based on their maximum error, average error, and coefficient of determination (R^2) values for %DEC and %COD reduction. The comparative results are detailed in Table 2. The `logsig` function exhibited commendable performance with the lowest maximum and average errors for DEC and COD, alongside high R^2 values, indicating precise predictions and strong model fitness. Notably, the Tansig function also performed well, especially in the COD category, with the highest R^2 values across the test, train, and complete datasets. However, it showed a higher error percentage compared to `logsig`. Conversely, the `purlin` function demonstrated significantly higher errors and lower R^2 values, suggesting less accuracy and model adequacy.

Upon comparison, the `logsig` activation function emerges as the most effective, combining lower error rates with high R^2 values, thus ensuring both accuracy and reliability in the predictive modeling of wastewater treatment parameters. The bounded nature of the `logsig` function makes it well-suited for classification tasks because it normalizes its output between 0 and 1. The `tansig` function, also bounded, is capable of producing negative values, which can be beneficial in regression scenarios similar to the one in this study. Although `logsig` and `tansig` perform comparably, `purelin`(being unbounded and with a consistent gradient) tends to underperform for this particular application as it operates as an identity function. Simulations indicate a preference for `logsig` based on its performance metrics; however, `tansig` may also be a viable option depending on specific use cases.

Table 2. Performance metrics of FFNN activation functions.

| Activation Function	Max Error (DEC)	Avg Error (DEC)	Max Error (COD)	Avg Error (COD)
logsig	4.079	0.482	2.449	0.726
tansig	4.488	0.760	3.592	0.781
purlin	10.882	4.956	10.484	5.192

2.3. K-fold cross validation

To evaluate the model's performance, we employed k-fold cross-validation on a dataset comprising 30 observations, which were divided into ten approximately equal-sized groups. During the cross-validation process, the neural network was trained on nine groups (constituting 90% of the data), while the remaining group (10%) was used for testing the model's accuracy. This division and

testing sequence was repeated ten times with unique group combinations. The maximum absolute error and mean absolute error for each iteration were recorded, as illustrated in Figures 8 and 9. This method is crucial as it provides insights into the artificial neural network's (ANN) overall predictive accuracy.
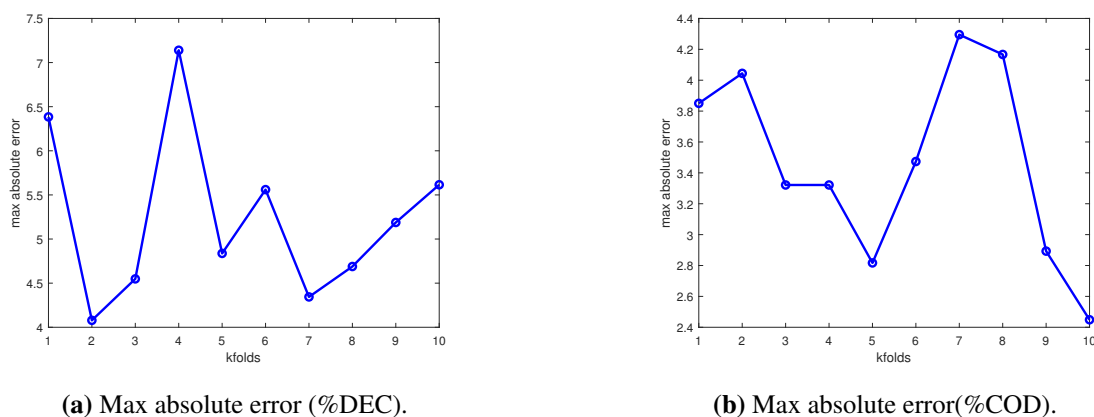
(a) Max absolute error (%DEC).　　　　(b) Max absolute error(%COD).

Figure 8. Error plot using max absolute errors k-fold models.
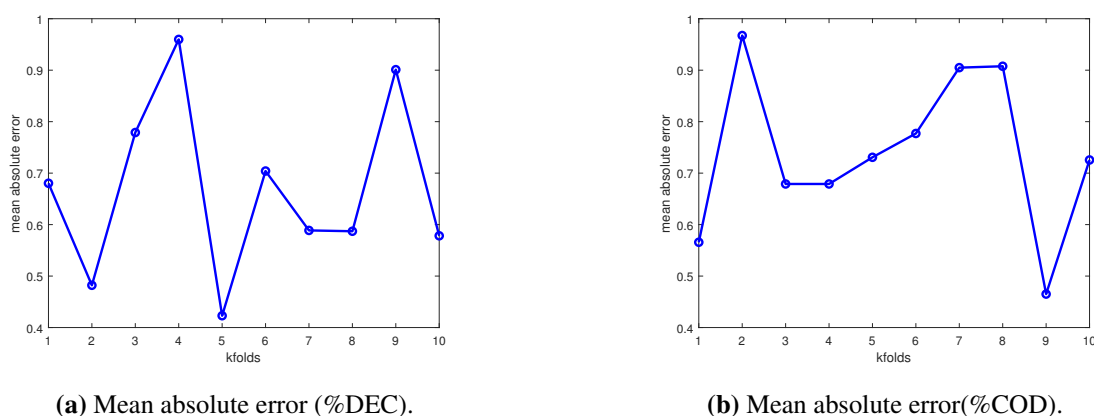
(a) Mean absolute error (%DEC).　　　　(b) Mean absolute error(%COD).

Figure 9. Error plot using mean absolute errors k-fold models.

A neural network's effectiveness hinges on its internal parameters: weights and biases. These parameters are refined during training via an optimization algorithm like backpropagation, which adjusts them based on prediction errors compared to actual outcomes. Properly calibrated weights and biases are crucial for the neural network's task performance and accuracy on new data. The trained model's weights and biases for %DEC, shown below (Eqs (2.8) and (2.9)), represent these scaled parameters

$$\mathbf{W}^{(1)} = \begin{bmatrix} -2.3898 & -0.1492 & -0.2641 & -0.9848 \\ -0.4323 & 0.4641 & -0.0365 & 1.6058 \\ -1.9563 & 0.9855 & -0.7559 & -0.0556 \\ 0.5595 & 1.3666 & -0.9700 & -0.2274 \end{bmatrix}, \quad \mathbf{b}^{(1)} = \begin{bmatrix} -1.4689 \\ 1.4982 \\ 0.8996 \\ -0.5276 \end{bmatrix}, \quad (2.8)$$

$$\mathbf{W}^{(2)} = \begin{bmatrix} -2.3345 & 1.2586 & 1.6142 & -1.6828 \end{bmatrix}, \quad \mathbf{b}^{(2)} = \begin{bmatrix} -0.6462 \end{bmatrix}. \quad (2.9)$$

Similarly, the weights and biases of the trained model for %COD reduction are given below (Eqs (2.10) and (2.11)).

$$\mathbf{W}^{(1)} = \begin{bmatrix} -0.9953 & -0.2235 & -0.0654 & 0.6878 \\ 2.2070 & -0.7526 & 0.6212 & -0.2701 \\ 1.8252 & 0.2532 & 0.1390 & 1.6896 \\ 0.5798 & 1.3836 & -1.0326 & -0.8945 \end{bmatrix}, \quad \mathbf{b}^{(1)} = \begin{bmatrix} -0.6459 \\ -1.1744 \\ 1.5634 \\ -0.7813 \end{bmatrix}, \tag{2.10}$$

$$\mathbf{W}^{(2)} = \begin{bmatrix} -1.0946 & -2.0369 & 2.4652 & -1.5414 \end{bmatrix}, \quad \mathbf{b}^{(2)} = \begin{bmatrix} -0.1190 \end{bmatrix}. \tag{2.11}$$

2.4. Regression plot

The R^2-value between the outputs and targets of the designed network is shown in Figure 10 for %DEC case (for %COD reduction see Figure 11). The dashed line depicts the line of ideal results, while the solid line shows the line of best fit. With R^2-values ranging from 0.9982 to 0.9998 for the training phase and 0.9456 to 0.9838 for the test step, a good match between the output and the target data was found for all examined parameters for the training and test stages. The R^2-value varied from 0.9970 to 0.9990 overall for all parameters tested, demonstrating the model's accuracy and reliability.
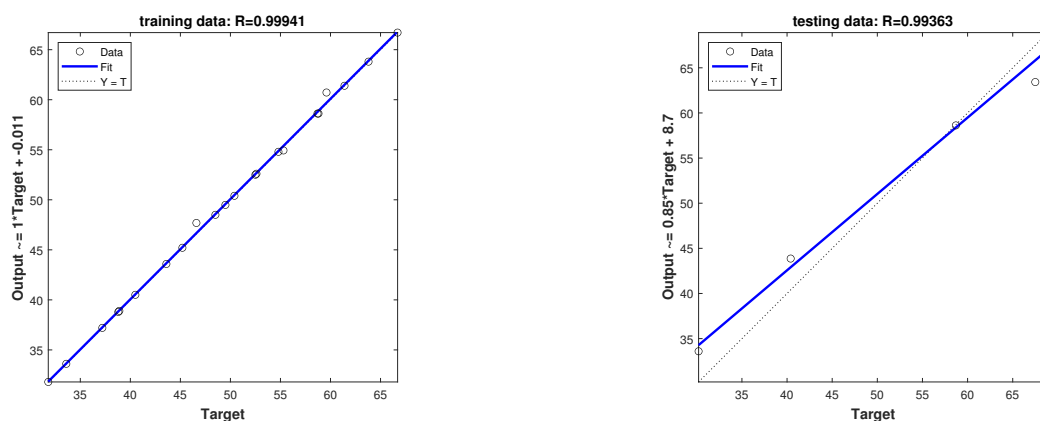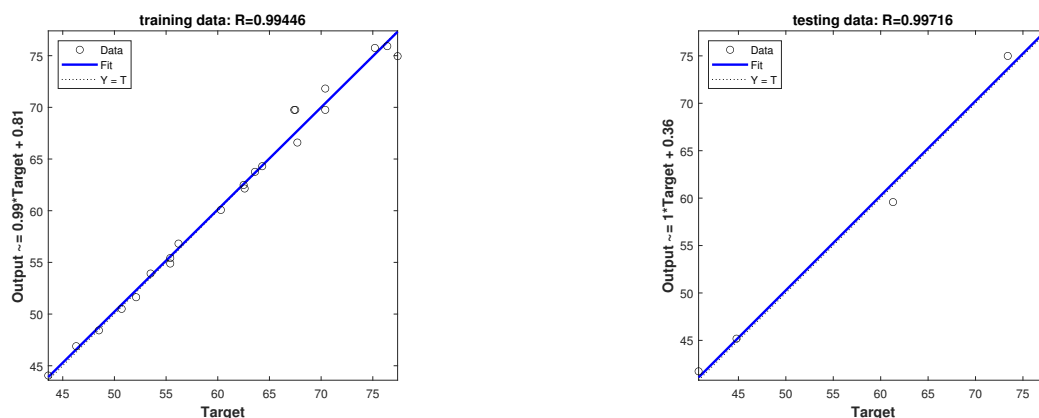

(a) Regression model using training data set for %DEC.

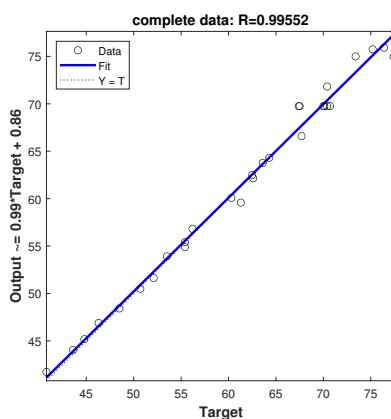(b) Regression model using testing data set for %DEC.
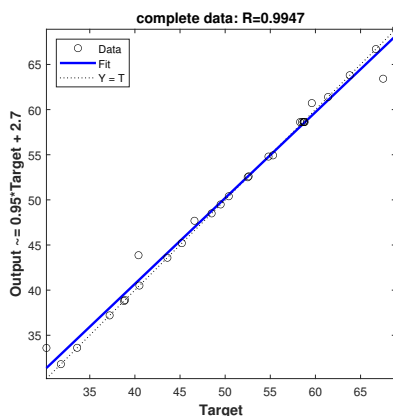
(c) Regression model using complete data for %DEC.

Figure 10. Regression models for %DEC.

(a) Regression model using training data set for %COD reduction.

(b) Regression model using testing data set for %COD reduction.

(c) Regression model using complete data for %COD reduction.

Figure 11. Regression models for %COD reduction.

2.5. Second-order approximation of FFNN

A single-layer perceptron, a basic form of an FFNN, processes inputs by multiplying them with weights, adding biases, and applying an activation function to produce the hidden layer output. This output is further transformed to produce the final output as shown in Eq (2.1). The second-order approximation of Eq (2.1) using Taylor's series expansion is given by:

$$\hat{Y}_{\%\text{DEC}} = 58.62 - 1.96x_1 - 1.35x_2 + 4.20x_3 + 13.98x_4 + 2.65x_2x_1 - 5.69x_3x_1 - 8.56x_4x_1$$
$$+ 4.48x_2x_3 - 1.59x_2x_4 - 1.53x_3x_4 - 17.98x_1^2 - 3.44x_2^2 - 1.85x_3^2 - 5.12x_4^2,$$
$$\hat{Y}_{\%\text{COD}} = 69.76 - 2.08x_1 - 0.70x_2 + 3.30x_3 + 15.09x_4 + 1.82x_2x_1 - 4.69x_3x_1 - 8.79x_4x_1$$
$$+ 4.73x_2x_3 + 0.49x_2x_4 - 2.43x_3x_4 - 16.79x_1^2 - 3.35x_2^2 - 1.94x_3^2 - 7.41x_4^2.$$

Here, $\hat{Y}_{\%\text{DEC}}$ and $\hat{Y}_{\%\text{COD}}$ predict %DEC and %COD reduction, respectively, with x_1, x_2, x_3, x_4 representing pH, temperature, agitation speed, and dye wastewater concentration.

The model's sensitivity to predictors is analyzed through partial derivatives at optimal conditions ($x_1 = -0.2, x_2 = 0.16, x_3 = 0.66, x_4 = 1$):

$$\frac{dY_{\%\text{DEC}}}{dx_1} = 0.3630, \quad \frac{dY_{\%\text{DEC}}}{dx_2} = 0.1669,$$

$$\frac{dY_{\%\text{DEC}}}{dx_3} = -0.1462, \quad \frac{dY_{\%\text{DEC}}}{dx_4} = 5.0106,$$

$$\frac{dY_{\%\text{COD}}}{dx_1} = -0.4650, \quad \frac{dY_{\%\text{COD}}}{dx_2} = -0.3646,$$

$$\frac{dY_{\%\text{COD}}}{dx_3} = -0.6051, \quad \frac{dY_{\%\text{COD}}}{dx_4} = 3.9230.$$

These derivatives indicate how a 1% change in each predictor affects %DEC and %COD. Similar analysis is performed for the FFNN model at optimal conditions ($x_1 = -0.3309, x_2 = 0.2588, x_3 = 0.4337, x_4 = 1$):

$$\frac{dY_{\%\text{DEC}}^{\text{FFNN}}}{dx_1} = -0.5119, \quad \frac{dY_{\%\text{DEC}}^{\text{FFNN}}}{dx_2} = -2.7096,$$

$$\frac{dY_{\%\text{DEC}}^{\text{FFNN}}}{dx_3} = 3.8914, \quad \frac{dY_{\%\text{DEC}}^{\text{FFNN}}}{dx_4} = 7.8107,$$

$$\frac{dY_{\%\text{COD}}^{\text{FFNN}}}{dx_1} = -0.3353, \quad \frac{dY^{\text{FFNN}}}{dx_2} = 0.2229,$$

$$\frac{dY_{\%\text{COD}}^{\text{FFNN}}}{dx_3} = 1.5798, \quad \frac{dY_{\%\text{COD}}^{\text{FFNN}}}{dx_4} = 4.5630.$$

The partial derivatives of the second-order approximation of the FFNN model are:

$$\frac{d\hat{Y}_{\%\text{DEC}}}{dx_1} = -0.4073, \quad \frac{d\hat{Y}_{\%\text{DEC}}}{dx_2} = -3.6537,$$

$$\frac{d\hat{Y}_{\%\text{DEC}}}{dx_3} = 4.1129, \quad \frac{d\hat{Y}_{\%\text{DEC}}}{dx_4} = 5.5041,$$

$$\frac{d\hat{Y}_{\%\text{COD}}}{dx_1} = -1.3121, \quad \frac{d\hat{Y}_{\%\text{COD}}}{dx_2} = -0.5003,$$

$$\frac{d\hat{Y}_{\%\text{COD}}}{dx_3} = 1.9685, \quad \frac{d\hat{Y}_{\%\text{COD}}}{dx_4} = 2.2511.$$

These derivatives demonstrate the rate of change in %DEC and %COD reduction with respect to changes in the predictors (x_1, x_2, x_3, x_4).

2.6. Effects of predictors on responses through contour plots

Contour plots offer an effective method to study the interactive effects of parameters on %DEC and %COD reduction. In this study, artificial neural network models were employed to generate the contour plots, enabling a comprehensive examination of the interaction impact between two parameters on dye wastewater's %DEC and %COD reduction. These plots represent two elements simultaneously while keeping the other components constant, making them highly valuable for understanding the direct and indirect impacts of the two factors.

Figures 12–23 illustrate the contours for the %DEC and %COD reduction of dye wastewater. The nature of these contours vividly illustrates how the variables interact, providing valuable insights into their combined influence on the treatment process. The curve's elliptical shape denotes a positive interaction between the two variables, while its round shape denotes a negative interaction. The circular form of the contour in graphs, as seen from the pictures, represents how all the variables interact. Every pair of variables exhibits a significant interaction, and the surface included inside the contour diagram's smallest ellipse denotes the most significant expected yield.

Figure 12 shows the interaction between pH and temperature in textile dye %DEC. pH significantly influences how microorganisms treat dyeing effluents, which is vital in %DEC. The contour plot in Figure 12 visually illustrates the combined impact of pH and temperature on dye %DEC, offering insights for optimizing treatment conditions. The efficacy of dye %DEC is increased by increasing pH, according to Figure 12. Following that, the %DEC's effectiveness declines. Figures 13 and 14 showed a similar trend. The pH significantly impacts the effectiveness of dye %DEC, and for most dyes, the ideal pH range for color removal is between 6.0 and 7.0.
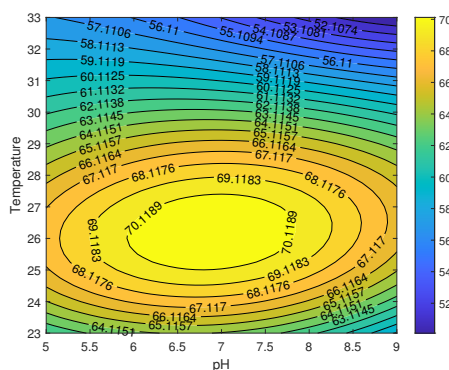

Figure 12. Relationship between pH, temperature, and %DEC with fixed agitation speed and wastewater concentration.
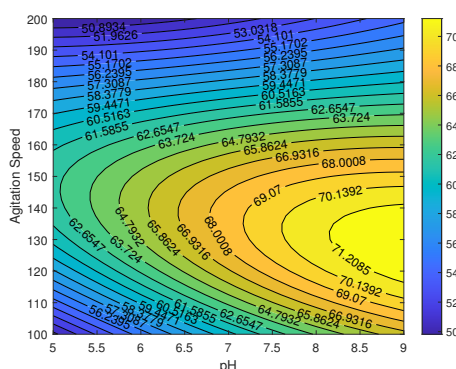

Figure 13. Relationship between pH, agitation speed, and %DEC with fixed temperature and wastewater concentration.
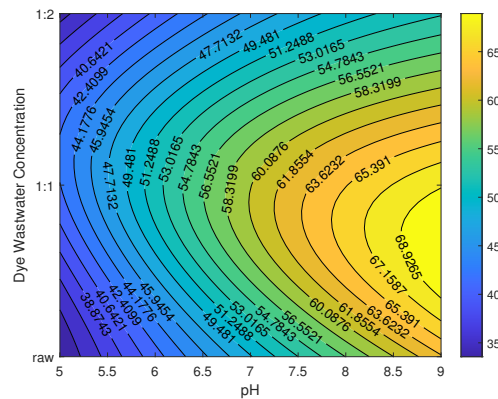
Figure 14. Relationship between pH, wastewater concentration, and %DEC with fixed temperature and agitation speed.

The %DEC is likewise seen in Figure 12 to increase with temperature up to 29.2 $°C$ and then decrease. At higher temperatures, decolorizing action was dramatically reduced. This could be caused by a decrease in cell viability or by the deactivation of the %DEC-causing enzymes. According to Figure 13, the %DEC percentage rises with increasing agitation speed, reaching a maximum of 171.68 rpm. The efficacy of dye removal declines after that. This was seen in Figures 15 and 17, too. A drop in dye concentration accelerates %DEC. Figures 14, 16 and 17 demonstrate this in great detail. The findings indicate that the percentage of dye elimination increases as dye concentration decreases. This is attributed to the higher presence of chemicals and contaminants in dye wastewater at higher concentrations, which hinders the growth and activity of microorganisms responsible for dye degradation.
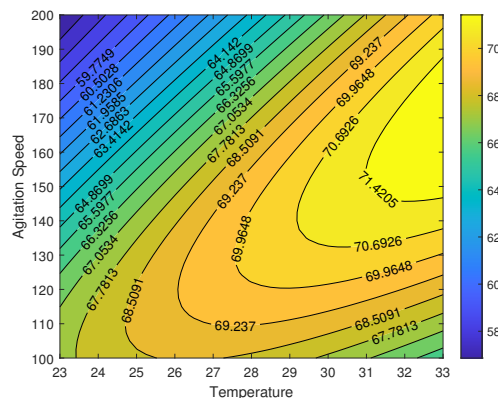

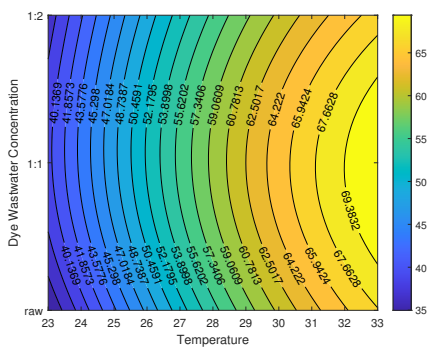**Figure 15.** Relationship between temperature, agitation, and %DEC with fixed pH and wastewater concentration.

Figure 16. Relationship between temperature, wastewater concentration, and %DEC with fixed pH and agitation speed.
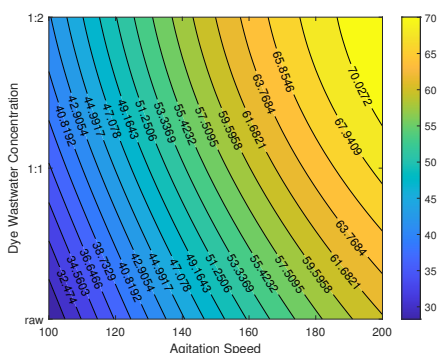

Figure 17. Relationship between agitation speed, wastewater concentration, and %DEC with fixed temperature and pH.

Similar trends were observed for the influence of process factors on %COD reduction in textile dye wastewater, as depicted in Figures 18–23. The optimal conditions of pH 6.3, temperature 29.29 °C, agitation speed 171.68 rpm, and dye wastewater concentration 1:2 resulted in the highest %DEC and %COD reduction. These optimal conditions were further validated through experiments, confirming the effectiveness of the ANN predictions. Under these optimal settings, the maximum color removal and %COD reduction were determined to be 70.69% and 80.50% , respectively.
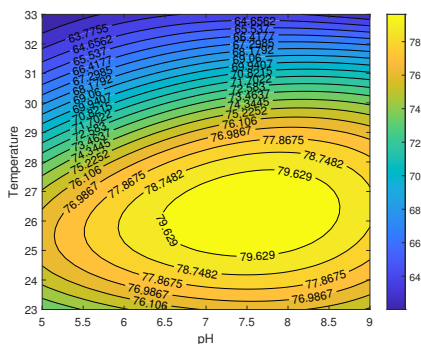

Figure 18. Relationship between pH, temperature, and %COD reduction with fixed agitation speed and wastewater concentration.
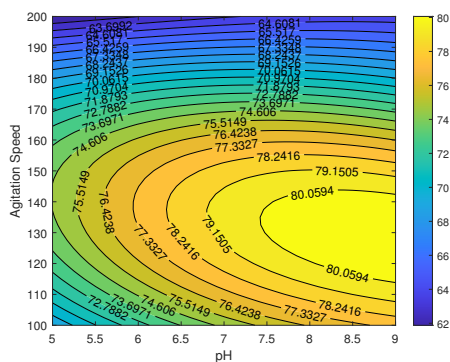
Figure 19. Relationship between pH, agitation speed, and %COD reduction with fixed temperature and wastewater concentration.
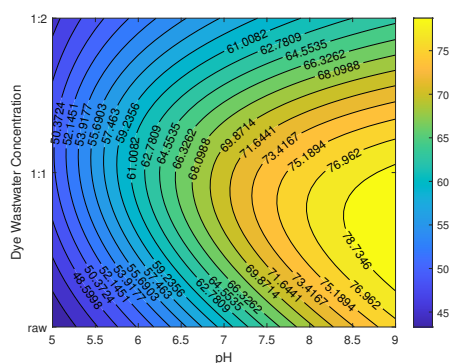

Figure 20. Relationship between pH, wastewater concentration, and %COD reduction with fixed agitation speed and temperature.
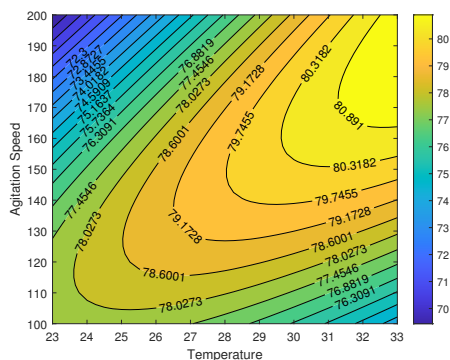

Figure 21. Relationship between temperature, agitation speed, and %COD reduction with fixed pH and wastewater concentration.
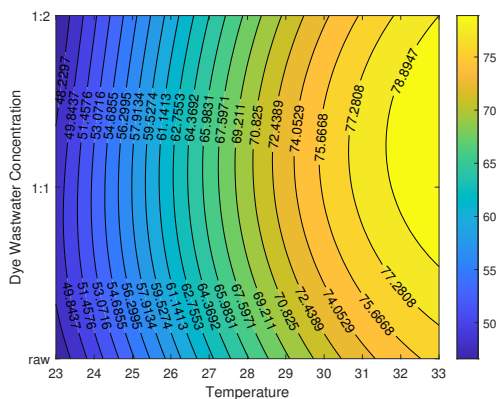
Figure 22. Relationship between temperature, wastewater concentration and %COD reduction with fixed agitation speed and pH.
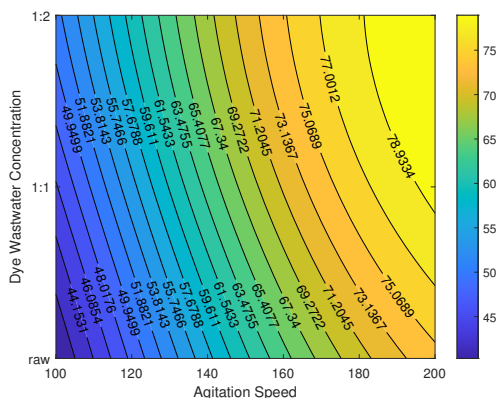

Figure 23. Relationship between agitation speed, wastewater concentration, and %COD reduction with fixed pH and temperature.

2.7. Sensitivity analysis

Sensitivity analysis was conducted on the neural network models to evaluate the robustness of %COD and %DEC predictions. This entailed adjusting the ANN model parameters slightly to observe the response variation. Stability was inferred from the boundedness of the resulting errors. In our study, a perturbation of 0.01 was introduced to the FFNN model's weights and biases. The perturbed model's outcomes were then measured against the original FFNN model's performance, with the deviations captured in Figure 24. The observed maximum absolute errors in Figure 24 were relatively low, with 2.3852 for %DEC and 1.1949 for %COD, demonstrating the model's stability against small internal variations.
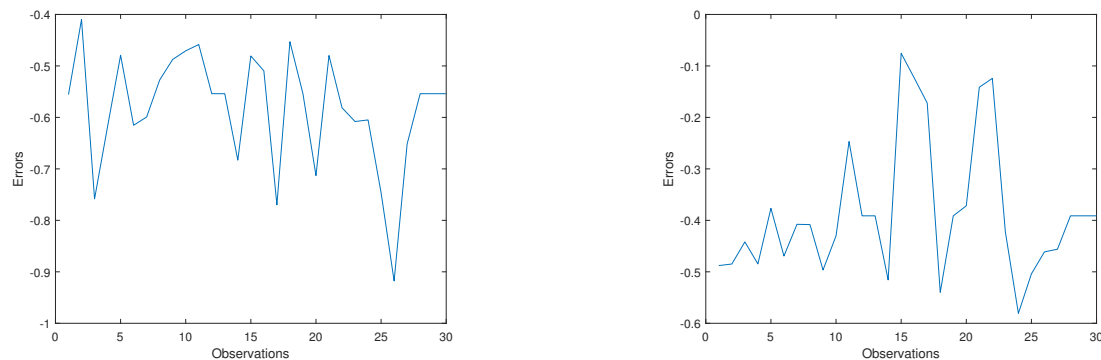
(a) Error plot of sensitive analysis for %DEC.

(b) Error plot of sensitive analysis for %COD reduction.

Figure 24. Error plots using input perturbation method for the sensitive analysis of predicting FFNN.

3. Optimization

The authors [16] used second-order polynomial equations from Response Surface Methodology (RSM) on experimental data to determine optimal conditions for %DEC and %COD removal. Optimization in MATLAB$^®$ yielded optimal conditions: pH 6.6, temperature 28.8 °C, agitation speed 183 rpm, and dye wastewater concentration 1:2, with experimental validation achieving 71.2% %DEC and 80.5% %COD removal. We utilized the `fminmax` constrained optimization function from the optimization toolbox in MATLAB$^®$ to invoke the ANN models for the percentage of dye removal (%DEC) and chemical oxygen demand reduction (%COD). The necessity to simultaneously maximize %DEC and %COD dictated the use of the `minmax` optimization strategy. The ANN predicted optimal conditions were a pH of 6.3, a temperature of 29.2 °C, an agitation speed of 171 rpm, and a dye concentration ratio of 1:2. These conditions were experimentally validated, achieving a %DEC of 71.69% and a %COD removal of 80.50%. Under these optimized conditions, the ANN's second-order approximation yielded a %DEC of 69.91% and a %COD of 79.98%.

4. Comparative assessment of FFNN and second-degree polynomial regression models

We analyze multiple performance indicators when assessing the efficacy of the second-degree polynomial regression model (1.1) and the FFNN model for predicting %DEC and %COD reduction. For %DEC, the FFNN model exhibits superior performance over the second-degree polynomial regression in terms of both training ($R^2 = 0.99941$ vs. $R^2 = 0.9714$) and testing ($R^2 = 0.99363$ vs. $R^2 = 0.8354$) datasets, signifying a more substantial proportion of variance explained by the FFNN model. Moreover, the FFNN model attains a lower maximum absolute error (4.0787 vs. 4.64) and mean absolute error (0.4821 vs. 1.44), indicating enhanced predictive precision for %DEC.

For % COD reduction, the trend continues with the FFNN model surpassing the polynomial regression with higher training ($R^2 = 0.99446$ vs. $R^2 = 0.9747$) and testing ($R^2 = 0.99716$ vs. $R^2 = 0.9495$) R^2 values. Error metrics further corroborate the FFNN's superiority, with a reduced maximum absolute error (2.4486 compared to 3.79) and mean absolute error (0.7256 versus 1.36).

The FFNN model outstrips the second-degree polynomial regression model in both the explanation

and prediction of %DEC and %COD reduction. The FFNN model not only elucidates a more significant fraction of the data's variability but also forecasts with higher accuracy, as reflected by the diminished error magnitudes.

5. Conclusions

The comprehensive analysis of the single-layer Feed-Forward Neural Network (FFNN) model in this study underscores its efficacy in treating textile dye wastewater, particularly in reducing %decolorization (%DEC) and chemical oxygen demand (%COD). The model's structure, involving a single input layer with four neurons and one output neuron, adeptly captures the complex relationships within the data. The training and testing phases, guided by the mean squared error (MSE) objective function, have demonstrated impressive accuracy, as reflected in the high coefficient of determination (R^2) values. We summarizes the study's critical findings, which are as follows:

- **High model accuracy:** The FFNN model exhibits superior performance with R^2 values ranging from 0.9982 to 0.9998 during training and 0.9456 to 0.9838 in testing. These high R^2 values indicate a strong correlation between the model's predictions and the target data, thereby validating its accuracy and reliability.
- **Effective activation functions:** Among various activation functions, `logsig` stands out, demonstrating the lowest maximum and average errors for %DEC and %COD, with values of 4.079 and 0.482 for %DEC, and 2.449 and 0.726 for %COD, respectively. This finding establishes `logsig` as the most effective activation function for the FFNN model.
- **Optimal treatment conditions:** The optimal conditions predicted by the FFNN model for %DEC and %COD reduction were experimentally validated, resulting in a significant %DEC of 71.69% and %COD reduction of 80.50%. These results highlight the model's practical utility in real-world applications.
- **Comparative analysis with polynomial regression:** The FFNN model outperforms the second-degree polynomial regression model in both explanation and prediction of %DEC and %COD reduction. The FFNN model achieves a lower maximum absolute error and mean absolute error compared to the polynomial regression, thereby demonstrating enhanced predictive precision.

The study effectively demonstrates the potential of single-layer FFNN models in environmental applications, particularly in the context of wastewater treatment in the textile industry.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

Conflict of interest

The authors declare no conflict of interest.

References

1. G. M. Walker, L. R. Weatherley, Adsorption of acid dyes on to granular activated carbon in fixed beds, *Water Res.*, **31** (1997), 2093–2101. https://doi.org/10.1016/S0043-1354(97)00039-0

2. A. H. Hassani, S. Seye, A. H. Javid, M. Borgheei, Comparison of adsorption process by GAC with novel formulation of coagulation–flocculation for color removal of textile wastewater, *Int. J. Environ. Res.*, **2** (2008), 239–248.

3. S. Wijannarong, S. Aroonsrimorakot, P. Thavipoke, S. Sangjan, Removal of reactive dyes from textile dyeing industrial effluent by ozonation process, *APCBEE procedia*, **5** (2013), 279–282. https://doi.org/10.1016/j.apcbee.2013.05.048

4. V. K. Gupta, A. Nayak, S. Agarwal, I. Tyagi, Potential of activated carbon from waste rubber tire for the adsorption of phenolics: Effect of pre-treatment conditions, *J. Colloid Interf. Sci.*, **417** (2014), 420–430. https://doi.org/10.1016/j.jcis.2013.11.067

5. X. R. Xu, H. B. Li, W. H. Wang, J. D. Gu, Decolorization of dyes and textile wastewater by potassium permanganate, *Chemosphere*, **59** (2005), 893–898. https://doi.org/10.1016/j.chemosphere.2004.11.013

6. A. A. Ansari, B. D. Thakur, Bio-chemical reactor for treatment of concentrated textile effluent, *Colourage*, **49** (2002), 27–30.

7. F. Zhang, A. Yediler, X. Liang, A. Kettrup, Ozonation of the purified hydrolyzed azo dye reactive red 120 (CI), *J. Environ. Sci. Health Part A*, **37** (2002), 707–713. https://doi.org/10.1081/ESE-120003248

8. R. Rajeshkannan, M. Rajasimman, N. Rajamohan, Sorption of acid blue 9 using *Hydrilla verticillata* biomass—optimization, equilibrium, and kinetics studies, *Bioremediat. J.*, **15** (2011), 57–67. https://doi.org/10.1080/10889868.2010.548002

9. R. Krull, M. Hemmi, P. Otto, D. C. Hempel, Combined biological and chemical treatment of highly concentrated residual dyehouse liquors, *Water Sci. Technol.*, **38** (1998), 339–346. https://doi.org/10.1016/S0273-1223(98)00517-4

10. P. Verma, D. Madamwar, Decolourization of synthetic dyes by a newly isolated strain of *Serratia marcescens*, *World J. Micro. Biotechnol.*, **19** (2003), 615–618. https://doi.org/10.1023/A:1025115801331

11. V. Arutchelvan, D. J. Albino, V. Muralikaishnan, S. Nagarajan, Decolourization of textile mill effluent by Sporotrichum pulverulentum, *Indian J. Environ. Ecoplaning*, **7** (2003), 59–62.

12. J. P. Jadhav, S. P. Govindwar, Biotransformation of malachite green by *Saccharomyces cerevisiae* MTCC 463, *Yeast*, **23** (2006), 315–323. https://doi.org/10.1002/yea.1356

13. J. S. Chang, C. Chou, S. Y. Chen, Decolorization of azo dyes with immobilized pseudomonas luteola, *Process Biochem.*, **36** (2001), 757–763. https://doi.org/10.1016/S0032-9592(00)00274-0

14. M. S. Khehra, H. S. Saini, D. K. Sharma, B. S. Chadha, S. S. Chimni, Decolorization of various azo dyes by bacterial consortium, *Dyes Pigments*, **67** (2005), 55–61. https://doi.org/10.1016/j.dyepig.2004.10.008

15. S. Pointing, Feasibility of bioremediation by white-rot fungi, *Appl. Microbiol. Biotechnol.*, **57** (2001), 20–33. https://doi.org/10.1007/s002530100745

16. S. Sathian, G. Radha, V. Shanmugapriya, M. Rajasimman, C. Karthikeyan, Optimization and kinetic studies on treatment of textile dye wastewater using *Pleurotus floridanus*, *Appl. Water Sci.*, **3** (2013), 41–48. https://doi.org/10.1007/s13201-012-0055-0

17. E. Pérez-Santín, L. de-la-Fuente-Valentín, M. G. García, K. A. S. Bravo, F. C. L. Hernández, J. I. L. Sánchez, Applicability domains of neural networks for toxicity prediction, *AIMS Mathematics*, **8** (2023), 27858–27900. https://doi.org/10.3934/math.20231426

18. H. Najafi, A. Bensayah, B. Tellab, S. Etemad, S. K. Ntouyas, S. Rezapour, et al., Approximate numerical algorithms and artificial neural networks for analyzing a fractal-fractional mathematical model, *AIMS Mathematics*, **8** (2023), 28280–28307. https://doi.org/10.3934/math.20231447

19. N. Ruttanaprommarin, Z. Sabir, R. A. S. Núñez, S. Salahshour, J. L. G. Guirao, W. Weera, et al., Artificial neural network procedures for the waterborne spread and control of diseases, *AIMS Mathematics*, **8** (2023), 2435–2452.https://doi.org/10.3934/math.2023126

20. S. D. Mourtas, E. Drakonakis, Z. Bragoudakis, Forecasting the gross domestic product using a weight direct determination neural network, *AIMS Mathematics*, **8** (2023), 24254–24273. https://doi.org/10.3934/math.20231237

21. N. Alruwais, H. Alamro, M. M. Eltahir, A. S. Salama, M. Assiri, N. A. Ahmed, Modified arithmetic optimization algorithm with Deep Learning based data analytics for depression detection, *AIMS Mathematics*, **8** (2023), 30335–30352. https://doi.org/10.3934/math.20231549

22. J. E. F. Moraes, F. H. Quina, C. A. O. Nascimento, D. N. Silva, O. Chiavone-Filho, Treatment of saline wastewater contaminated with hydrocarbons by the photo-Fenton process, *Environ. Sci. Technol.*, **38** (2004), 1183–1187. https://doi.org/10.1021/es034217f

23. V. K. Pareek, M. P. Brungs, A. A. Adesina, R. Sharma, Artificial neural network modeling of a multiphase photodegradation system, *J. Photoch. Photobio. A: Chem.*, **149** (2002), 139–146. https://doi.org/10.1016/S1010-6030(01)00640-2

24. A. M. Ghaedi, A. Vafaei, Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: A review, *Adv. Colloid Interfac. Sci.*, **245** (2017), 20–39. https://doi.org/10.1016/j.cis.2017.04.015