Chemometrics and Intelligent Laboratory Systems 249 (2024) 105132

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemometrics

Application of ANN, hypothesis testing and statistics to the adsorptive removal of toxic dye by nanocomposite

Thamraa Alshahrani [a,**], Ganesh Jethave [b,*], Anil Nemade [b], Yogesh Khairnar [b], Umesh Fegade [c], Monali Khachane [b], Amir Al-Ahmed [d], Firoz Khan [d]

[a] Department of Physics, College of Science, Princess Nourah Bint Abdulrahman University, Riyadh, 11671, Saudi Arabia
[b] Dr. Annasaheb G. D. Bendale Mahila Mahavidyalay Jalgaon, Maharashtra, India
[c] Department of Chemistry, Bhusawal Arts, Science and P. O. Nahata Commerce College, Bhusawal, MH, 425201, India
[d] Interdisciplinary Research Center for Sustainable Energy Systems (IRC-SES), King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, 31261, Saudi Arabia

ARTICLE INFO

Keywords:
Trimetalic nanocomposite
Dye adsorption
Statistical analysis
Karl Pearson's correlation coefficient
Hypothesis testing
ANN modeling

ABSTRACT

Statistics can be used in a variety of ways to present, compute, and critically analyze experimental data. To determine the significance and validity of the experimental data, a variety of statistical tests are used. Using a synthesized CoO/NiO/MnO_2 Nanocomposite, the present study used adsorption to remove the dye Bromophenol Blue (BPB) from a contaminated aqueous solution. In order to (a) determine the optimal pH of the solution, (b) confirm the experiment's success, and (c) investigate the effect of adsorbent dose on BPB dye removal from aqueous solutions. The experimental data were statistically analyzed through hypothesis testing using the *t*-test, paired *t*-test, and Chi-square test. The null hypothesis that the optimal pH value is 7 is accepted since $t_{observed}$ (−1.979)<$t_{tabulated}$ (−2.262). Since $\chi^2_{observed}$ (1.052)< $\chi^2_{tabulated}$ (3.841), null hypothesis that the higher adsorbent dose helps in higher % removal of dye is accepted. Both the obtained Freundlich adsorption isotherm and the Langmuir isotherm's R^2 values, which were both close to 1, indicate that the isotherms are favorable. Karl Pearson's relationship coefficient values for Langmuir and Freundlich adsorption isotherms found to be 0.9693 and 0.9994 respectively, which show a more significant level of connection between's the factors. The ANN model predicted adsorption percentage with regression value R is 0.996. ANN model result predict 99.60 % BPB dye adsorption using optimized parametric conditions. The ANN model produced values that were more precise, reliable, and reproducible, demonstrating its superiority.

1. Introduction

Dyes possess the potential to jeopardize both human health and the environment due to their chemical composition and inherent toxicity [1]. Various hazardous effects are associated with dyes [2], primarily stemming from their inclusion of harmful chemicals like heavy metals, aromatic amines, and benzene derivatives, which can endanger human health by means of skin absorption, inhalation, or ingestion [3,4]. These chemicals may provoke a spectrum of health issues such as skin irritation, respiratory problems, carcinogenicity, mutagenic, reproductive toxicity, and endocrine disruption. Moreover, dyes can inflict adverse consequences on the environment, particularly when released into water bodies through manufacturing processes or improper disposal methods. The environmental hazards include water and soil pollution, as well as the non-biodegradability of many synthetic dyes, leading to long-term accumulation in ecosystems[3,5]. Despite the existence of regulations ensuring dye safety, it is imperative to handle and dispose of dyes responsibly, opting for sustainable practices like eco-friendly alternatives to mitigate their potential risks.

Additionally, adsorption emerges as a promising method for addressing dye contamination [6,7]. This process offers several advantages, including the effective removal of contaminants from air, water, and gases, versatile applications across various substances, high adsorption capacity [8], selectivity [9,10], regeneration and reusability of adsorbent materials, straightforward operation, rapid kinetics [11,12], minimization of waste generation, and potential resource recovery [13,14]. These benefits render adsorption a valuable technique for environmental remediation, water treatment, gas purification, and

* Corresponding author.
** Corresponding author.
E-mail addresses: thmalshahrani@pnu.edu.sa (T. Alshahrani), ganeshjethave@gmail.com (G. Jethave).

https://doi.org/10.1016/j.chemolab.2024.105132
Received 5 October 2023; Received in revised form 17 April 2024; Accepted 22 April 2024
Available online 23 April 2024

resource recovery efforts [15,16,17]. Moreover, artificial neural networks (ANNs) present a powerful tool for modeling and optimizing the adsorption of dyes onto metal oxide nanoparticles[18]. Leveraging experimental data, ANNs can facilitate data modeling, process optimization, prediction of adsorption behavior, understanding complex interactions, design optimization, and sensitivity analysis, thereby enhancing the efficiency and efficacy of dye adsorption processes [19–24].

Overall, the application of artificial neural networks in the adsorption of dyes on metal oxide nanoparticles offers a powerful tool for modeling, optimization, and prediction [5,25–29]. By leveraging the capabilities of ANNs, researchers can gain valuable insights into the adsorption process and develop efficient and sustainable strategies for dye removal and wastewater treatment.

Our focus extends to the utilization of artificial neural networks (ANNs) for modeling and optimizing the adsorption of BPB dyes onto CoO/NiO/MnO_2 nanoparticles, with the ultimate aim of enhancing the efficiency and effectiveness of dye removal processes. The choice of BPB as the pollutant for the adsorption study likely reflects its availability, well-characterized properties, representatives of organic pollutants, ease of detection, standardization, and relevance to environmental remediation efforts [30]. And the choice of CoO/NiO/MnO_2 nanocomposite material as the adsorbent likely stems from its potential synergistic effects, high surface area, tunable properties, stability, environmental compatibility [31], and previous research [32] demonstrating its effectiveness for adsorption applications. By elucidating these aims, we lay the foundation for the subsequent sections of our study, which delve into experimental methodologies, results, and discussions aimed at addressing these objectives.

2. Karl Pearson's correlation coefficient (r)

Equation (1) displays the widely used Karl Pearson correlation coefficient, which measures the degree of connection between two variables.

$$r = \frac{[\sum (X_i - X) \times (Y_i - Y)]}{n \times \sigma_X \times \sigma_Y} \tag{1}$$

where X and Y are the variables of the Karl Pearson correlation coefficient. X and Y are the values of c and c/W, respectively, c for the Langmuir isotherm and W for the Freundlich adsorption isotherm, respectively while X_i and Y_i are the ith values of the independent variables.

3. Statistical analysis of experimental data

Statistical analysis of experimental data involves applying various statistical techniques to analyze and interpret the results obtained from an experiment. The goal is to draw meaningful conclusions, identify patterns or trends, assess the reliability of the data, and make statistical inferences about the population from which the data were collected. Hypothesis testing is used to make inferences about population parameters based on sample data. The process involves formulating null and alternative hypotheses, selecting an appropriate statistical test (e.g., ANOVA, chi-square test, *t*-test), calculating a test statistic, and determining the statistical significance of the results. Hypothesis testing helps determine whether observed differences or associations in the data are statistically significant or due to random chance.

The null hypothesis (H_o) is assuming that it is true in order to statistically describe the process qualitatively and quantitatively. While simultaneously defining an alternative hypothesis (H_a), consideration is given to whether the alternative hypothesis is accepted if the null hypothesis is rejected or the other way around. Before testing the hypothesis, the level of significance, or maximum probability of rejecting the null hypothesis, is established [33]. Hypothesis tests, also known as tests of significance for analyzing experimental data, can be broken down into two categories: (1) parametric tests, such as the Chi-square test, *t*-test, z-test, and the F-test; and (2) distribution-free tests, which are hypotheses that are tested without making assumptions based on the characteristics of the original population.

z-test contrasts the example mean and the conjectured populace mean worth. For large samples, $n > 30$, it is used to determine the significance of the mean.

When the population is normal and infinite, the population variance is unknown, and the sample size is small ($n < 30$), then the *t*-test is used. With unknown population variance, it is an appropriate test for determining the significance of the difference between two sample means.

For small values of n, the paired *t*-test compares two related samples under the assumption that they are normal. It is not necessary for two populations to have the same variance. Typically, this test is applied to data from both treatment studies and controls.

A useful statistical method for determining the independence of population variance, significance, and goodness of fit is the Chi-square test [34–38].

4. Adsorption isotherms

Adsorption isotherm equations are used to analyze experimental data ([39]; McCabe and co. 1993). Isotherm not entirely set in stone and the coefficient of relationship is determined. We can obtain a correlation coefficient of the linear fit of the experimental data [37]. Langmuir and Freundlich are two commonly used adsorption isotherms that describe the relationship between the concentration of adsorbate (dye solute) and the amount adsorbed on a solid adsorbent. These isotherms are widely employed to understand and characterize adsorption processes. Equation (2) provides a favorable form of the Langmuir isotherm.

$$W = W_{max}\left(\frac{K_c}{1 + K_c}\right) \tag{2}$$

To test the Langmuir isotherm model, c values are plotted against c/W. The plot shows the values of the Langmuir isotherm constants K and W_{max}. The isotherm is strongly favorable if K is large and Kc is greater than 1. When $Kc < 1$, the isotherm is nearly linear. The Langmuir isotherm assumes that adsorption sites are independent and do not interact with each other. It predicts that the adsorption capacity reaches a maximum value (q_{max}) at high concentrations, indicating saturation of the adsorbent surface.

Equation (3) describes the Freundlich isotherm, which is also a favorable type of isotherm. It is an empirical model that describes heterogeneous adsorption on surfaces with varying adsorption energies and sites.

$$W = b.c^m \tag{3}$$

Regarding, log c is plotted Vs log W to put the Freundlich isotherm model to the test. The plot is used to determine the b and m Freundlich isotherm constants. A better fit for adsorption from liquids is provided by $m < 1$.

Both the Langmuir and Freundlich isotherms have their applications and limitations, and their appropriateness depends on the specific adsorption system and experimental conditions. These isotherms provide valuable insights into adsorption phenomena and are often used to characterize adsorbents, evaluate adsorption processes, and design adsorption systems.

5. Materials and method

The CoO/NiO/MnO_2 were synthesized according to a co-precipitation method followed by magnetic stirring. Firstly 8.4 gm of Co Acetate, 4.99 gm $MnSO_4$ and 4 gm of Ni Nitrate was added to 200 ml DW. 2 M NaOH solution was added to this slurry slowly with constant

stirring until the pH becomes 12 ± 0.2 and stirring was further continued for 2 h at 60 °C temperature (Fig. 1). After completion of reaction, precipitate is filter by Whatman filter paper then washed the precipitate several time with distilled water till pH become neutral. After that drying of precipitate at 80 °C is carried out followed by calcination at 500 °C.

This dried powder in the range of 50–200 mg was used to calculate the % removal of BPB dye from the solution. Batch experiment was carried out in the 'Orbital' constant temperature variable speed shaker holding 12 Erlenmeyer flasks at a time at room temperature and 1200 rpm for 90 min in the pH range 1–10. The solutions were centrifuge before each batch experiment analysis. Analysis of concentration of dye in the filtrate was done by UV–Vis spectrophotometer (UV 1800, Schimadzu) at the wavelength 590 nm. Percentage removal of dye from the solution after the batch adsorption was calculated as.

$$\%\ Removal = \left[\frac{C_i - C_o}{C_i}\right] \times 100 \tag{4}$$

where C_o and C_i represent the final and initial concentration of BPB dye (mg/L) in the solution, respectively.

$$q_e = \frac{[(C_i - C_o)V]}{m} \tag{5}$$

where q_e is the equilibrium adsorption capacity (mg adsorbate/g adsorbent), V is the volume of the solution (liters), and m is the adsorbent's mass (mg) [40].

100 ppm stock solution of BPB dye is prepared by taking 100 mg dye in 1000 ml of distilled water. From this stock solution, series of various concentration solutions are prepared for the calibration curve.

6. Characterization of adsorbent

Selected Area Electron Diffraction (SAED) is a technique commonly used in electron microscopy to analyze the crystal structure of materials. SAED analysis allow researchers to gain insights into their crystal structure, lattice parameters, orientation, texture, and size. SAED provides valuable information about the arrangement of atoms within a sample. When analyzing metal oxide nanoparticles using SAED, the spacing between crystal planes in the metal oxide nanoparticles can be determined from the diffraction pattern. SAED can provide information about the orientation and texture of metal oxide nanoparticles. HRTEM (JEOL JEM 2100 plus) provides atomic-scale details of nanoparticle structure, crystal lattice, defects, and interfacial characteristics. HRTEM is particularly suitable for studying smaller nanoparticles and individual nanoparticles and EDAX provides qualitative and quantitative information about the elemental composition and distribution within nanoparticles. The FE-SEM (Make-Bruker Modal-S-4800) with a driving voltage of 15 kV provides high-resolution, three-dimensional images of nanoparticles and offers information about their size, shape, and surface morphology. FESEM is particularly useful for studying larger nanoparticles or bulk samples. These techniques provide valuable information about their size, shape, structure, composition, and elemental distribution. The advanced spectrophotometric method using a UV–Visible spectrophotometer (Shimadzu, Japan model, UV-1800) was used to determine the BPB dye concentration in the experimental solution at a wavelength of 590 nm. Additionally, pH measurement was conducted using a digital pH meter.

7. Artificial neural network (ANN)

All calculations in this study to predict the adsorption efficiency, and computer programs written in MATLAB and are used to create the modular artificial neural networks with an NN toolbox. The Mean Square Error (MSE) of the training and prediction set were also used to determine the variation of the models' parameters. We randomly divided 1000 experimental runs into training (80 %) and testing (20 %) dataset. Training set randomly divided into Training, validation, and test sets to test the model's ability to predict unobserved experiments that were not used in the training process.

There are three layers: an input layer with four neurons (dye concentration, solution pH, contact time, and adsorbent dose), an output layer with one neuron, and a hidden layer with 18 different modes. In nonlinear and complex chemical engineering systems, including the adsorption process, the Multilayer ANN model is used to forecast inputs and outputs. The BPB dye adsorption percentage is predicted using the

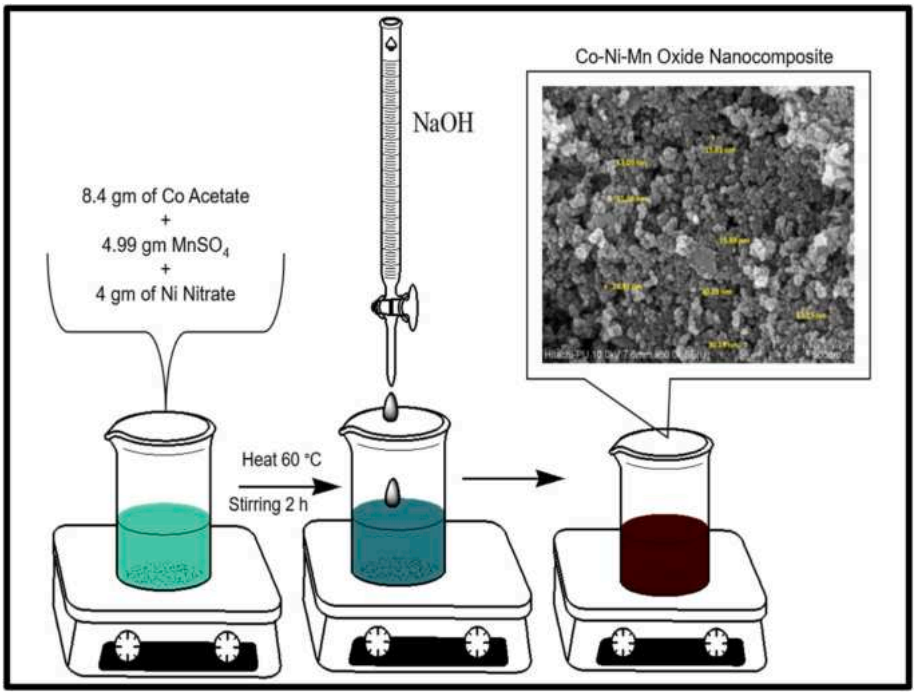

Fig. 1. Schematic representation of synthesis of Nanocomposite.

ANN model in the current study.

The feedforward ANN model is composed with input, hidden and output layer. The ANN layer contains one input layer with 04 inputs called as neurons. pH, concentration (ppm), Adsorbent dose (mg) and Time (min) are inputs for the ANN network. The hidden layers with different number of neurons from 1 to 20 are experimented and output layer with one neuron namely % adsorption. An ANN model is designed using MATLAB 2022(a).

8. Results and discussion

As visible in the FE-SEM and HRTEM image presented in Fig. 2a and c, the $CoO/NiO/MnO_2$ are consists of crystalline particles sized in 13–30 nm of scale. The SAED pattern displayed in Fig. 2b portrays well-defined

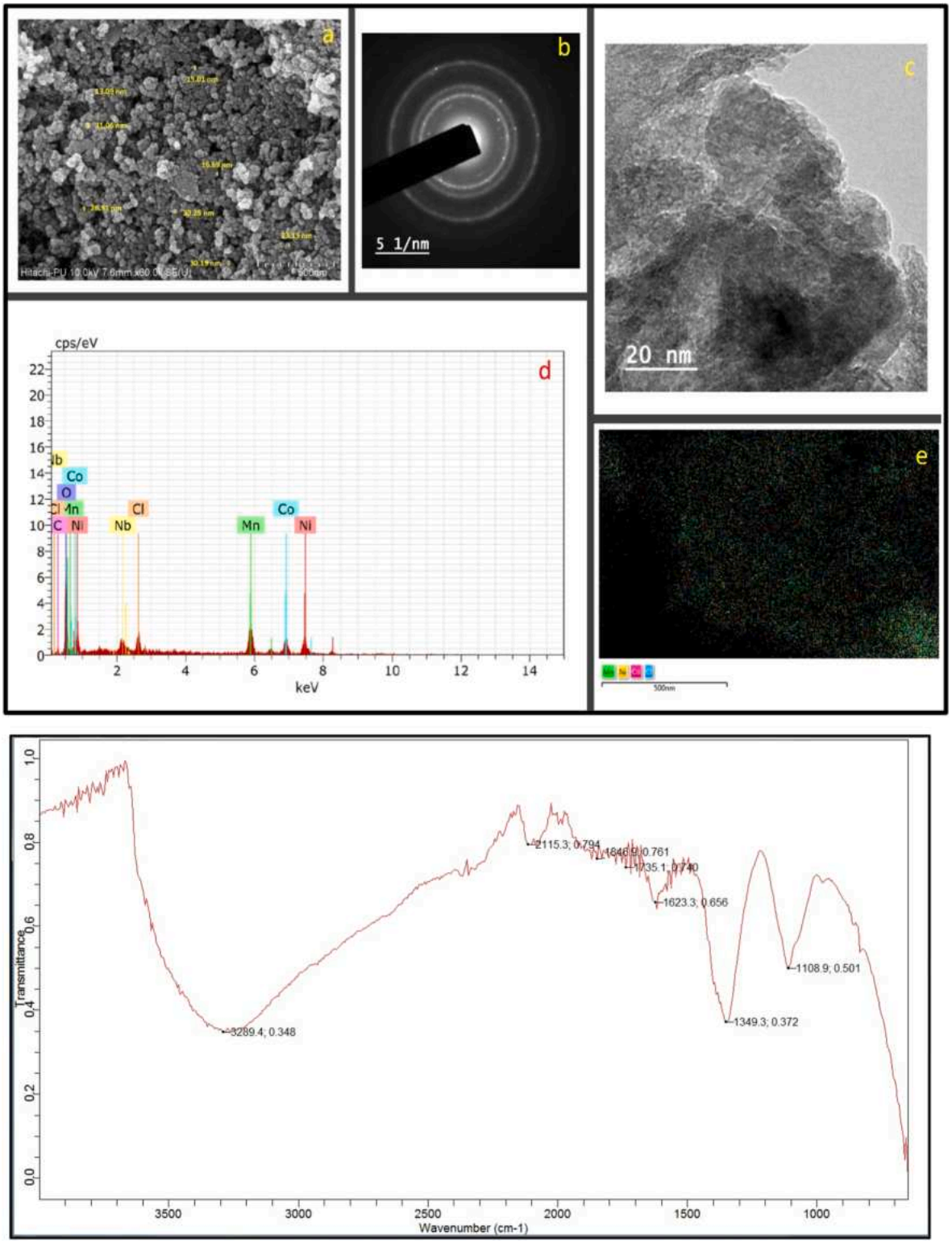

Fig. 2. a) FESEM image b) SAED image c) HRTEM image (d) EDX spectra and e) Elemental mapping and f) FT-IR spectra of $CoO/NiO/MnO_2$ Nanocomposite.

rings consisting of distinct spots and several isolated spots, suggesting the nanoscale polycrystalline nature of the CoO/NiO/MnO_2. The homogeneously distributed Mn on the CoO/NiO further validates the spectra. Furthermore, the successful mixture of Mn with the CoO/NiO surface and the formation of CoO/NiO/MnO_2 are indicated by the well-defined elemental mapping of Co, Ni, O, and Mn, which has been shown by different colors (as shown in Fig. 2e). In Fig. 2d, the EDX spectra of the CoO/NiO/MnO_2 show the composition of the sample with peaks corresponding to Co, Ni, O, and Mn. Additionally, the EDX area scanning also confirms the presence of Co (12 %), Ni (17 %), O (44 %), Mn (14 %) and traces of other elements (13 %) in the CoO/NiO/MnO_2.

In FTIR analysis 3289 cm^{-1} peak likely corresponds to the stretching vibration of hydroxyl (-OH) groups, which are commonly found in metal hydroxides and metal oxide surfaces. The presence of this peak suggests the presence of surface hydroxyl groups on the nanocomposite. 1349 cm^{-1} peak could correspond to the bending vibration of hydroxyl (-OH) groups. It may indicate the presence of metal hydroxides or metal carboxylates on the nanocomposite surface. 1108 cm^{-1} peak may correspond to the stretching vibration of metal-oxygen (M − O) bonds in metal oxides or hydroxides. Its presence confirms the presence of metal oxide components in the nanocomposite.

9. Statistical analysis

9.1. Hypothesis testing

Data collected from analysis of the filtrates for BPB dye concentration was used for

A. Testing the hypotheses: (a) The data gathered from the analysis of the filtrates for the concentration of BPB dye 1) to determine the optimal pH for maximum BPB dye removal from solution; 2) to determine the experiment's success; 3) to deduce that a higher adsorbent dose contributes to a higher percentage of BPB dye removal; 4) to determine the K value for favorability of the Langmuir isotherm; and 5) to determine the m value for favorability of the Freundlich isotherm.
B. Using the Karl Pearson correlation coefficient equation, to determine whether or not the data in the Langmuir and Freundlich adsorption isotherms are reliable.

9.2. I. hypothesis testing to determine the optimal pH value for maximizing BPB dye removal from a solution containing 10 ppm BPB dye

The optimal pH level was determined using a two-tailed *t*-test within a 5 % level of significance (Table 1) (see Table 2).

Null hypothesis Ho: Optimum pH = 7

Alternate hypothesis Ha: Optimum pH = 7.

Table 1
% removal of BPB dye w.r.t. pH

n	pH	% removal (X_i)	$(X_i-X_{avg})^2$
1	1	40	864.36
2	2	52	302.76
3	3	60	88.36
4	4	68	1.96
5	5	75	31.36
6	6	84	213.16
7	7	97	761.76
8	8	81	134.56
9	9	78	73.96
10	10	59	108.16
		$\Sigma X_i = 694$	
		Average $X_i = 69.4$	

Table 2
Dye concentration in the solution before (X_i) and after (Y_i) the experiment.

n	X_i	Y_i	$D_i = X_i - Y_i$	$(Di-Davg)^2$	% Adsorption
1	10	1	9	−24.625	90
2	20	3	17	−16.625	85
3	30	6	24	−9.625	80
4	40	9	31	−2.625	77.5
5	50	13	37	3.375	74
6	60	16	44	10.375	73.33
7	70	20	50	16.375	71.43
8	80	23	57	23.375	71.25

For (n - 1) degree of freedoms is the calculated value of t,

$$t = \frac{(X_{avg} - \mu)}{\frac{\sigma_s}{\sqrt{n}}} = t_{obs} \quad (5)$$

where σs = standard deviation = $\sqrt{\sum \left[\frac{(X_i - X_{avg})^2}{n-1}\right]}$.

X_{avg} = Average value of X and n = sample size.

From Eq. (5), the standard deviation value is 17.15, the calculated value of t is −1.979, and the tabulated value is −2.262 for 9 degrees of freedom (d.f.) at a 5 % level of significance. for t-distribution with two tails [37]. The null hypothesis that the optimal pH value is 7 is accepted since $t_{observed} < t_{tabulated}$ (Fig. 3).

$$t_{calculated} = \frac{(D_{avg} - 0)}{\frac{\sigma_{diff}}{\sqrt{n}}} \quad (6)$$

Where n = number of matched pairs, D_{avg} = Mean of differences, σ_{diff} = Standard deviation of differences.

D_{avg} (mean of differences) = (∑Di)/n = 33.62

Standard deviation $\sigma_{diff} = \sqrt{\left[\sum \left\{\frac{D_i - (D_{avg})^2 . n}{n-1}\right\}\right]} = 16.54$

From Eq. (6), the calculated value of t is $t_{observed}$ = 5.75, whereas $t_{tabulated}$ = 2.365 at 5 % level of significance for 9 degrees of freedom (d.f.) for 2-tailed t-distribution [37]. Since $t_{calculated} < t_{observed}$, The alternate hypothesis that the experiment in removing BPB dye from the contaminated solution was successful is accepted, while the null hypothesis that the experiment did not result in any change in the concentration of BPB dye in the solution is rejected (Fig. 4).

The data from the experiments that were carried out for BPB dye removal from samples of initial BPB dye concentrations of 10 and 100

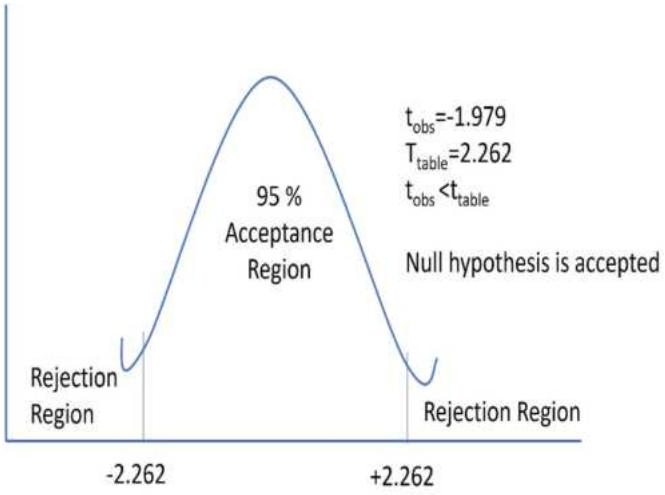

Fig. 3. Probability chart for t-distribution for two-tailed test.

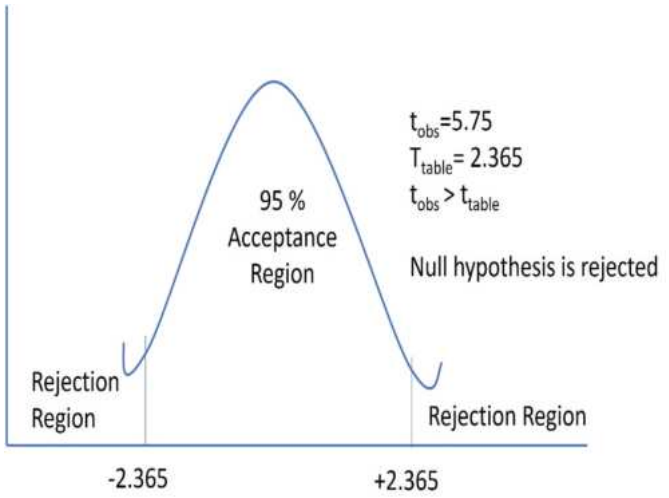

Fig. 4. Probability chart for t-distribution for two-tailed test.

ppm with adsorbent dosages of 100 and 200 mg at pH = 7 were subjected to hypothesis testing using the Chi-square test to infer that the higher adsorbent dose helps in higher % removal of BPB dye (Tables 3 and 4).

Null hypothesis H_o: % removal is higher with higher adsorbent dose

Alternate hypothesis H_a: % removal decreases with adsorbent dose

If Oij = Observed frequencies of the cell in the ith row and jth column.

Eij = Expected frequencies of the cell in the ith row and jth column.

$$\chi^2 = \sum \left\{ (O_{ij} - E_{ij})^2 \Big/ E_{ij} \right\} \quad (7)$$

From Eq. (7), the calculated value of χ^2 is χ^2 observed = 1.052

whereas χ^2 tabulated = 3.841 at 5 % level of significance for 1 degrees of freedom (d.f.) for χ^2-distribution [37].

Since χ^2 observed< χ^2 tabulated, null hypothesis that the higher adsorbent dose helps in higher % removal of dye is accepted (Fig. 5).

9.3. Adsorption isotherms

Experimental data were analyzed using Langmuir and Freundlich adsorption isotherm equations (Eqs. (2) and (3)).

9.4. Langmuir isotherm

To determine whether the Langmuir isotherm equation was favorable or not, the experimental data is used. An isotherm curve is produced on the arithmetic graph by plotting the BPB dye concentration c (ppm) against c/W, where W is the adsorbent loading (g adsorbed per g solid) (Table 5; Fig. 6). The isotherm's slope is 220, and the Langmuir isotherm constant K is 0.8491, as shown in the plot. A favorable isotherm is indicated by the value of K (1) as being of the linear type. The value 0.9693 for the Karl-Pearson correlation coefficient (Eq. (1)), indicating

Table 3
Observed frequencies for dye removal for initial dye concentration of 10 and 100 ppm and adsorbent dosages 100 and 200 mg at pH 7.

Adsorbent Dose (mg)	% Dye Removal		Total
	initial conc 10 ppm	initial conc 100 ppm	
100	45	20	65 = A
200	95	40	135 = a
Total	**140=B**	**60=b**	**200=N**

Table 4
Calculation for Chi square.

Group	Oij	Eij	$(Oij-Eij)^2/Eij$
AB	45	45.5	0.005495
Ab	15	19.5	1.038462
aB	95	94.5	0.002646
ab	40	40.5	0.006173

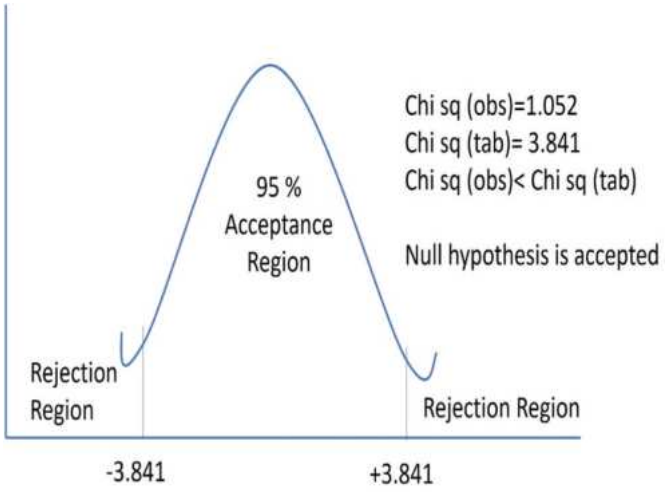

Fig. 5. Probability chart for Chi-square distribution for effectiveness of the adsorption experiment.

Table 5
Experimental data for Langmuir isotherm.

c, ppm	Adsorbent Dose, mg	Quantity of dye adsorbed	W = mg dye adsorbed/ mg adsorbent	c/w = Y
10	200	9	0.045	222.22
20	200	17	0.085	235.29
30	200	24	0.12	250
40	200	31	0.155	258.06
50	200	37	0.185	270.27
60	200	44	0.22	272.73
70	200	50	0.25	280
80	200	57	0.285	280.70

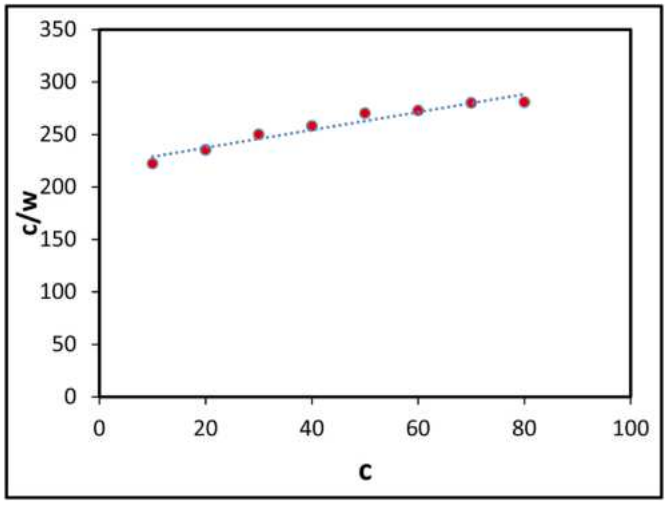

Fig. 6. Langmuir Isotherm plot.

that the two variables are highly correlated.

9.5. Freundlich isotherm

The Freundlich isotherm describes adsorption processes where the adsorption capacity continuously increases with increasing concentration, suggesting a non-saturation behavior. The Freundlich exponent, m, indicates the degree of adsorption intensity: $m > 1$ implies favorable adsorption, $m = 1$ represents linear adsorption, and $m < 1$ indicates less favorable adsorption. An isotherm curve was produced by plotting logarithmic values of c (BPB dye concentration, ppm) against logarithmic values of W (adsorbent loading, mg adsorbed/g solid) (Table 6; Fig. 7). The plot provided a slope (m) with a value of 2.22. The Freundlich adsorption isotherm was found to be a good fit and favorable for the adsorption data since $m > 1$. The fact that Karl Pearson's correlation coefficient was 0.9993 indicates that the two variables are highly correlated.

9.5.1. ANN modeling results

The dataset was randomly divided into training set with 588 tuples, validation dataset of 126 tuples and testing set with 126 tuples. Training algorithm Levenberg-Marquardt was employed on the data. The mean square error and regression values summarized in Table 7. Network with additional test dataset of 210 tuples were tested on neural network designed with 18 neurons. It was found that the 18 neurons regression value is comparatively high than other with minimum mean square error. The ANN diagram with three layers is shown in Fig. 8.

9.6. Prediction by ANN modeling

Utilizing the input neurons network topology, the developed ANN architecture is utilized to optimize BPB dye adsorption. In order to improve the ANN topology's generalizability, a number of ANN topologies are trained and the best one is chosen based on the minimization of mean square error (MSE) to determine the number of neurons in the hidden layer [49].

The optimal topology of the ANN model in this investigation has four input variables and an 18-neuron hidden layer (Fig. 8), as well as one output layer (fourteen-one) (Table 7). BPB dye adsorption (percent) has been shown to be better predicted by the developed ANN model. Using a multilayered neural network with the lowest MSE and a R value of 0.9930, the best network for predicting BPB dye adsorption conditions was chosen.

9.7. Optimization of the number of hidden neurons

The accuracy and prediction of optimal conditions were significantly influenced by the hidden layer's neurons. The trained networks are unable to effectively learn if the network topology is straightforward. As a result, the neurons were optimized so that the model's best predictive ability and accuracy could be used to select the best neurons. Based on the model's performance, such as R, and MSE values, the neurons that were predicted were obtained. Figure depicts the optimal structure of

Table 6
Experimental data for Freundlich isotherm.

c, ppm	W = mg dye adsorbed/ mg adsorbent	remaini dye in solution (Y_i)	log c	log w
10	0.045	1	1	−1.34679
20	0.085	3	1.30103	−1.07058
30	0.12	6	1.477121	−0.92082
40	0.155	9	1.60206	−0.80967
50	0.185	13	1.69897	−0.73283
60	0.22	16	1.778151	−0.65758
70	0.25	20	1.845098	−0.60206
80	0.285	23	1.90309	−0.54516

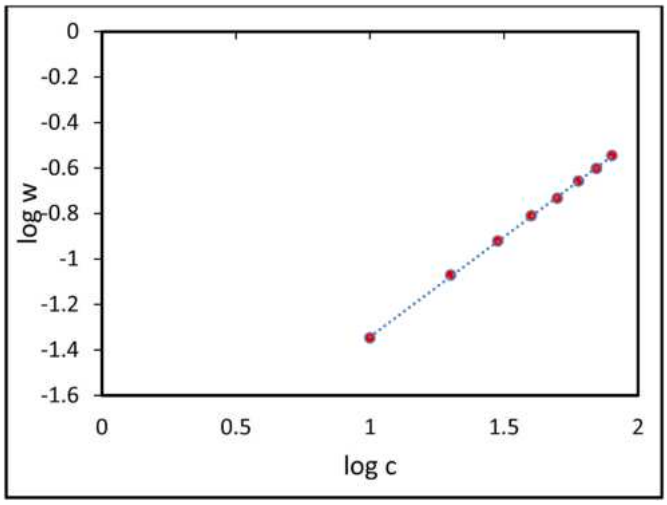

Fig. 7. Freundlich isotherm plot.

Table 7
Dependence among neuron number at hidden layer, MSE and R.

Number of neurons in hidden layer	Train		Validation		Test	
	MSE	R	MSE	R	MSE	R
1	215.3078	0.8091	257.3302	0.8166	210.9578	0.8406
2	61.9957	0.9502	63.3665	0.9523	74.8789	0.9481
3	37.9299	0.9718	39.3643	0.9705	34.7058	0.9716
4	30.1652	0.9765	22.5255	0.9822	28.0030	0.9805
5	29.9207	0.9776	27.1016	0.9766	34.0244	0.9793
6	18.9665	0.9855	18.0478	0.9873	22.4624	0.9812
7	17.3397	0.9866	19.1507	0.9865	20.2688	0.9846
8	20.5426	0.9843	27.3125	0.9798	21.9270	0.9840
9	17.4005	0.9866	20.7532	0.9847	23.3479	0.9822
10	24.0961	0.9816	20.5911	0.9843	43.8281	0.9655
11	15.1186	0.9887	18.2362	0.9844	15.4941	0.9885
12	26.3344	0.9801	26.1530	0.9816	29.3750	0.9735
13	14.4015	0.9891	14.1745	0.9892	16.9411	0.9863
15	12.5456	0.9908	11.1263	0.9915	18.4059	0.9833
16	9.8726	0.9923	10.4352	0.9924	13.5380	0.9904
17	10.4382	0.9921	13.5265	0.9898	17.1804	0.9864
18	**5.2033**	**0.9961**	**7.4697**	**0.9942**	**9.5043**	**0.9930**
19	8.9165	0.9931	15.6868	0.9885	16.2452	0.9877
20	9.3438	0.9930	14.2785	0.9896	13.5528	0.9887

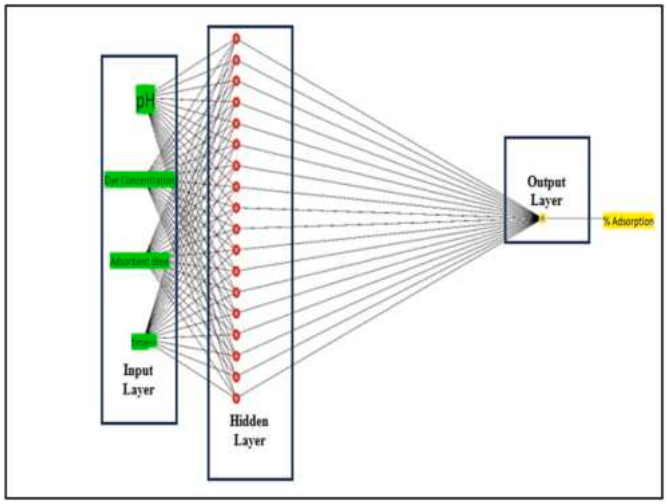

Fig. 8. ANN model with 18 neurons hidden layer.

the feed-forward network model for the neural network. The best-predicted capability and high accuracy of the model for BPB dye adsorption were observed in the optimal 18 neurons. The results showed higher values for R^2 (0.996), which strongly suggests that the developed ANN model is significant and can be used to predict the optimal neuron that has produced maximum BPB dye adsorption, as evidenced by the decrease in both MSE values (5.2) [13,28].

Additional 20 % dataset with 210 tuples were tested on the trainined model achieved 0.994 regression rate (Fig. 10).

9.8. Training, validation and testing of the model

For the purpose of developing the model, the input data were divided into three subcategories: training (70 percent), validation (15 percent), and testing (15 percent). Fig. 9 (a,b,c) is the ANN model with better training, validation, and testing R^2 values (0.996, 0.994, and 0.992, respectively). Fig. 9d indicates that a linear equation with an R^2 value of 0.995 was the best fit for the entire model [50,51]. As a result, the developed ANN model was able to more precisely reproduce experimental results and accurately simulate BPB dye adsorption (target). The algorithm-trained, multilayered feed-forward ANN with significant R^2 values was used to precisely achieve BPB dye adsorption (target).

9.9. External validation of the ANN dataset

According to Choi and Park [29], the probability of overfitting can be avoided when the number of neurons is between 8 and 12, which reduces the sensitivity of ANN modeling and improves prediction capabilities. The model's forecasting ability, on the other hand, has only occasionally been observed due to the limited number of variables chosen. Therefore, neither PCA nor partial least squares (PLS) have been connected to the ANN modeling. However, as [52] point out, external validation has been carried out between the actual output and the data sets obtained by ANN. Fig. 11 clearly depicts the external validation points, demonstrating that the ANN-rendered model does not exhibit overfitting and that it is statically more appropriate.

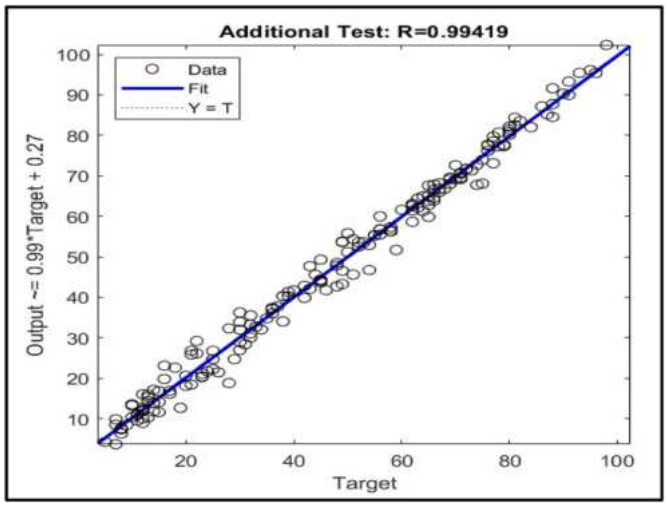

Fig. 10. Additional Test regression plot.

9.10. Inter-facial phenomena

Adsorption Mechanisms: This include physical adsorption, where dye molecules are held onto the surface by weak van der Waals forces, as well as chemical adsorption, involving stronger chemical bonds between the dye molecules and surface functional groups.

Surface Chemistry: We analyzed the surface chemistry of the nanocomposite and its interaction with the dye molecules. This involved considering factors such as surface charge, surface energy, and the presence of specific functional groups that can facilitate or hinder dye adsorption.

Interfacial Forces: The interfacial forces acting at the solid-liquid interface during the adsorption process. This may include considerations of electrostatic interactions, hydrogen bonding, steric effects, and the role of solvent molecules in mediating the interaction between the dye molecules and the nanocomposite surface.

Overall, by elucidating the experimental steps and interfacial

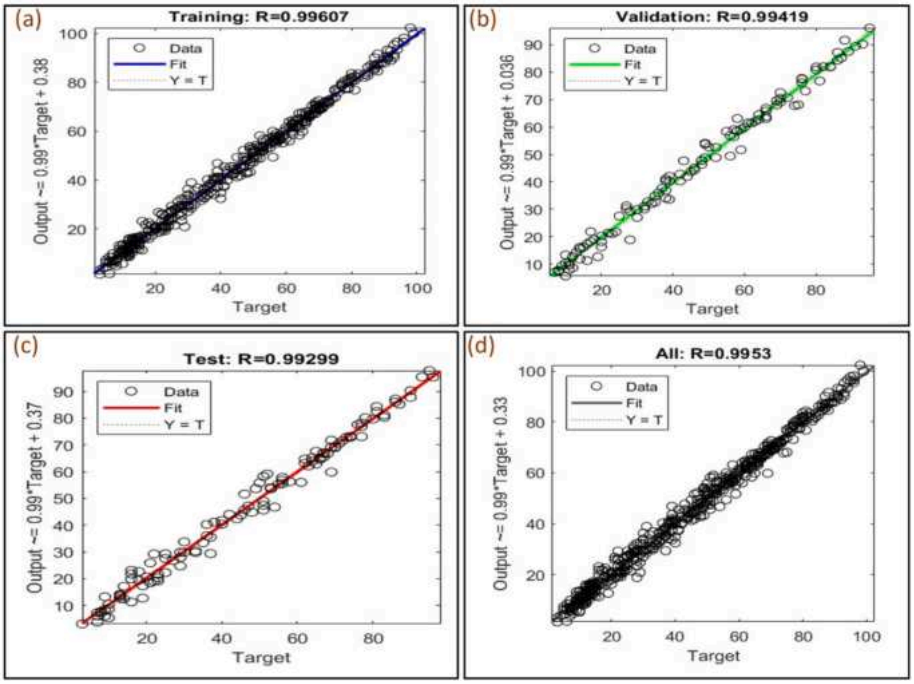

Fig. 9. Comparison between ANN derived and experimentally measured values of BPB dye adsorption for a) training b) validation c) testing and d) overall datasets.

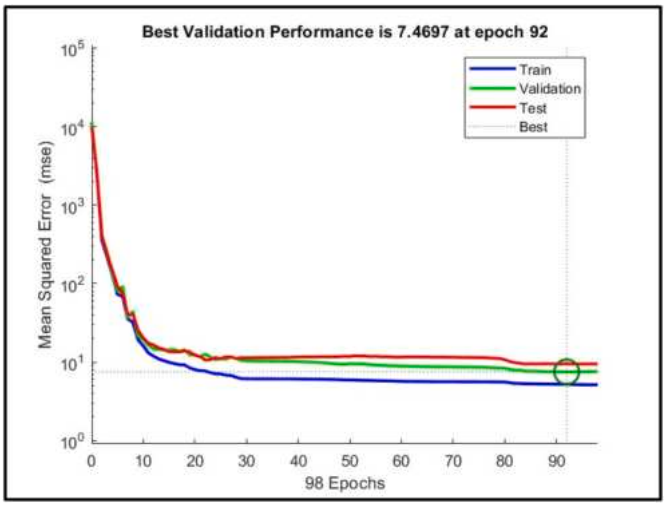

Fig. 11. MSE versus the number of neurons in hidden layer of the ANN model.

phenomena involved in the adsorption study, we have provided valuable insights into the mechanisms underlying dye adsorption onto the CoO/NiO/MnO_2 nanocomposite, facilitating a deeper understanding of its potential applications in various fields (Table 8).

10. Conclusion

The statements or assumptions that are made for the data that are obtained from experiments and need to be tested in order to accept or reject them are known as hypotheses. A method for proving these assumptions is hypothesis testing. In order to determine the optimal pH value for the maximum removal of BPB dye adsorption by CoO/NiO/MnO_2, evaluate the experiment's success, and infer that a higher adsorbent dose contributed to a higher percentage of BPB dye removal, we made assumptions, which were expressed as the null hypothesis in this study. Within a confidence level of 5 %, the problems were tested with the *t*-test, paired *t*-test, and Chi-square test. Our calculated values were found to be within the acceptance range of the probability charts, according to the findings.

The adsorption isotherms of Langmuir and Freundlich were used to further examine the experiment data. The favorable isotherms and the linear fit of the data into the isotherm equations were demonstrated by the outcomes. The Langmuir and Freundlich adsorption isotherms' respective Karl Pearson's correlation coefficients of 0.9693 and 0.9993 indicate a stronger correlation between the two variables. Future research should use the F-test and analysis of variance (ANOVA) to compare the means of multiple samples simultaneously.

Using two models, RSM and ANN, the current study has focused on carefully optimizing BPB dye adsorption. The constructed model has produced enhanced BPB dye adsorption with success, proving to be an accurate predictive tool and producing a more significant model. The research would undoubtedly persuade the scientific community to extensively utilize ANN's key advantages over conventional RSM analyses.

Funding

This research was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R1), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors are also thankful to Department of Science and Technology, India funded this research, (Sanction No. DST/CURIE-PG/2022/35(C)).

Table 8
Comparative study of BPB dye adsorption by various adsorbents.

Sr. No.	Adsorbent	Adsorption efficiency (mg/g)	Reference
1	Copper Oxide nanoparticles	49.85	Rashad, M. [41]
2	NiO NPs	93.46	Al-Aoh, Hatem A [42]
3	C_{14}-4-C_{14}im-SiNSs	230.77	Xiang, Yang [43]
4	Fe_3O_4/MIL-88A	167.2	Liu, Yi [44]
5	PAN-CD nanofibers crosslinked with citric acid	23.529	Mandla B. Chabalala [45]
6	Solanum tuberosum peel–nanoparticle hybrid (STpe-AgNP)	9.604	Akpomie, Kovo G [46]
7	ZnONPs	3.099	Akpomie, Kovo [47]
8	Activated carbon-Fe_3O_4–CuO composite	88.6	Alorabi, Ali Q [48]
9	CoO/NiO/MnO_2 Nanocomposite	140	Present study

Availability of data and materials

Not Applicable

CRediT authorship contribution statement

Thamraa Alshahrani: Funding acquisition. **Ganesh Jethave:** Funding acquisition, Investigation, Writing – original draft, Writing – review & editing. **Anil Nemade:** Supervision. **Yogesh Khairnar:** Resources. **Umesh Fegade:** Visualization, Writing – review & editing. **Monali Khachane:** Software, Writing – original draft. **Amir Al-Ahmed:** Data curation, Writing – review & editing. **Firoz Khan:** Formal analysis, Writing – review & editing.

Declaration of competing interest

The authors declare that there is no conflict of interest regarding the publication of this research paper, "Application of ANN, Hypothesis Testing, and Statistics for Adsorptive Removal of Toxic Dye by Nanocomposite." This research was driven solely by academic and scientific objectives, with a focus on advancing knowledge in the field of environmental science and nanotechnology. We confirm that this research was conducted with integrity and transparency, adhering to the highest ethical standards of scientific inquiry.

Data availability

The authors do not have permission to share data.

Acknowledgement

Authors are thankful to Graduate Students Miss. Karuna Patil, Miss. Supriya Nerkar and Miss. Namrata Sonar for their help. Authors are also thankful to Principal, Dr. Annasaheb G. D. Bendale Mahila mahavidyalaya, Jalgaon for her support, guidance and encouragement for this project.

References

[1] S.M. Bamforth, I. Singleton, Bioremediation of polycyclic aromatic hydrocarbons: current knowledge and future directions, J. Chem. Technol. Biotechnol.: International Research in Process, Environmental & Clean Technology 80 (7) (2005) 723–736.

[2] Rasha H. Khudhur, Nisreen S. Ali, Eman H. Khader, Noor S. Abbood, Issam K. Salih, Talib M. Albayati, Adsorption of anionic azo dye from aqeous wastewater using zeolite NaX as an efficient adsorbent, Desalination Water Treat. 306 (2023) 245–252.

[3] R. Pietzsch, S.R. Patchineelam, J.P. Torres, Polycyclic aromatic hydrocarbons in recent sediments from a subtropical estuary in Brazil, Mar. Chem. 118 (1–2) (2010) 56–66.
[4] T. Rengarajan, P. Rajendran, N. Nandakumar, B. Lokeshkumar, P. Rajendran, I. Nishigaki, Exposure to polycyclic aromatic hydrocarbons with special focus on cancer, Asian Pac. J. Trop. Biomed. 5 (3) (2015) 182–189.
[5] D.R. Dudhagara, R.K. Rajpara, J.K. Bhatt, H.B. Gosai, B.K. Sachaniya, B.P. Dave, Distribution, sources and ecological risk assessment of PAHs in historically contaminated surface sediments at Bhavnagar coast, Gujarat, India, Environ. Pollut. 213 (2016) 338–346.
[6] N.M. Gazzaz, M.K. Yusoff, A.Z. Aris, H. Juahir, M.F. Ramli, Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors, Mar. Pollut. Bull. 64 (11) (2012) 2409–2420.
[7] D. Bingöl, M. Hercan, S. Elevli, E. Kılıç, Comparison of the results of response surface methodology and artificial neural network for the biosorption of lead using black cumin, Bioresour. Technol. 112 (2012) 111–115.
[8] Y. Yasin, F.B.H. Ahmad, M. Ghaffari-Moghaddam, M. Khajeh, Application of a hybrid artificial neural network–genetic algorithm approach to optimize the lead ions removal from aqueous solutions using intercalated tartrate-Mg–Al layered double hydroxides, Environ. Nanotechnol. Monit. Manag. 1 (2014) 2–7.
[9] M. Buyukada, Co-combustion of peanut hull and coal blends: artificial neural networks modeling, particle swarm optimization and Monte Carlo simulation, Bioresour. Technol. 216 (2016) 280–286.
[10] R.M. Atlas, Handbook of Media for Environmental Microbiology, CRC press, 2005.
[11] J.C. Willison, Isolation and characterization of a novel sphingomonad capable of growth with chrysene as sole carbon and energy source, FEMS Microbiol. Lett. 241 (2) (2004) 143–150.
[12] J.K. Bhatt, C.M. Ghevariya, D.R. Dudhagara, R.K. Rajpara, B.P. Dave, Application of response surface methodology for rapid chrysene biodegradation by newly isolated marine-derived fungus Cochliobolus lunatus strain CHR4D, J. Microbiol. 52 (2014) 908–917.
[13] H. Kiyohara, K. Nagao, K. Yana, Rapid screen for bacteria degrading water-insoluble, solid hydrocarbons on agar plates, Appl. Environ. Microbiol. 43 (2) (1982) 454–457.
[14] C.M. Ghevariya, J.K. Bhatt, B.P. Dave, Enhanced chrysene degradation by halotolerant Achromobacter xylosoxidans using response surface methodology, Bioresour. Technol. 102 (20) (2011) 9668–9674.
[15] J. Ye, P. Zhang, E. Hoffmann, G. Zeng, Y. Tang, J. Dresely, Y. Liu, Comparison of response surface methodology and artificial neural network in optimization and prediction of acid activation of bauxsol for phosphorus adsorption, Water, Air, Soil Pollut. 225 (2014) 1–11.
[16] K.M. Desai, S.A. Survase, P.S. Saudagar, S.S. Lele, R.S. Singhal, Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: case study of fermentative production of scleroglucan, Biochem. Eng. J. 41 (2008) 266–273.
[17] A. Çelekli, F. Geyik, Artificial neural networks (ANN) approach for modeling of removal of Lanaset Red G on Chara contraria, Bioresour. Technol. 102 (10) (2011) 5634–5638.
[18] E.A. Dil, M. Ghaedi, A. Asfaram, The performance of nanorods material as adsorbent for removal of azo dyes and heavy metal ions: application of ultrasound wave, optimization and modeling, Ultrason. Sonochem. 34 (2017) 792–802.
[19] Y. Xu, M. Lu, Bioremediation of crude oil-contaminated soil: comparison of different biostimulation and bioaugmentation treatments, J. Hazard Mater. 183 (1–3) (2010) 395–401.
[20] M.H. Beale, M.T. Hagan, H.B. Demuth, Neural network toolbox, User's Guide, MathWorks 2 (2010) 77–81.
[21] A. Witek-Krowiak, K. Chojnacka, D. Podstawczyk, A. Dawiec, K. Pokomeda, Application of response surface methodology and artificial neural network methods in modelling and optimization of biosorption process, Bioresour. Technol. 160 (2014) 150–160.
[22] N.G. Turan, B. Mesci, O. Ozgonenel, The use of artificial neural networks (ANN) for modeling of adsorption of Cu (II) from industrial leachate by pumice, Chem. Eng. J. 171 (3) (2011) 1091–1097.
[23] L. Bruzzone, R. Cossu, G. Vernazza, Detection of land-cover transitions by combining multidate classifiers, Pattern Recogn. Lett. 25 (13) (2004) 1491–1500.
[24] K.P. Singh, A. Basant, A. Malik, G. Jain, Artificial neural network modeling of the river water quality—a case study, Ecol. Model. 220 (6) (2009) 888–895.
[25] T. Hill, L. Marquez, M. O'Connor, W. Remus, Artificial neural network models for forecasting and decision making, Int. J. Forecast. 10 (1) (1994) 5–15.
[26] P. Singh, S.S. Shera, J. Banik, R.M. Banik, Optimization of cultural conditions using response surface methodology versus artificial neural network and modeling of L-glutaminase production by Bacillus cereus MTCC 1305, Bioresour. Technol. 137 (2013) 261–269.
[27] P.A. Willumsen, U. Karlson, Effect of calcium on the surfactant tolerance of a fluoranthene degrading bacterium, Biodegradation 9 (5) (1998) 369–379.
[28] A. Gulbag, F. Temurtas, A study on quantitative classification of binary gas mixture using neural networks and adaptive neuro-fuzzy inference systems, Sensor. Actuator. B Chem. 115 (1) (2006) 252–262.
[29] D.J. Choi, H. Park, A hybrid artificial neural network as a software sensor for optimal control of a wastewater treatment process, Water Res. 35 (16) (2001) 3959–3967.
[30] A.E. Mahdi, N.S. Ali, H. Sh Majdi, Talib M. Albayatia, M.A. Abdulrahman, D. J. Jasim, K.R. Kalash, I.K. Salih, Effective adsorption of 2-nitroaniline from wastewater applying mesoporous material MCM-48: equilibrium, isotherm, and mechanism investigation, Desalination Water Treat. 300 (2023) 120–129.
[31] Jasim I. Humadi, Saad A. Jafar, Nisreen S. Ali, Mustafa A. Ahmed, Mohammed J. Mzeed, Raheem J. Al-Salhi, Noori M. Cata Saady, Hasan Sh Majdi, Sohrab Zendehboudi, Talib M. Albayati, Recovery of fuel from real waste oily sludge via a new eco-friendly surfactant material used in a digital baffle batch extraction unit, Sci. Rep. 13 (1) (2023) 9931.
[32] Nisreen S. Ali, Sh Majdi Hasan, Talib M. Albayati, Dheyaa J. Jasim, Adsorption of aniline from aqueous solutions onto a nanoporous material adsorbent: isotherms, kinetics, and mass transfer mechanisms, Water Pract. Technol. 18 (9) (2023) 2136–2150.
[33] A. Kaushal, S.K. Singh, Application of statistical tools and hypothesis testing of adsorption data obtained for removal of heavy metals from aqueous solutions, Int J Adv Res Innov 4 (2016) 82–84.
[34] A.L. Bowley, Elements of Statistics (No. 8), King, 1926.
[35] T.W. Anderson, S.D. Gupta, G.H.P. Styan, T.W. Anderson, An introduction to multivariate statistical analysis. Wiley, New York, Stat 10 (1958) 955–961.
[36] C.R. Kothari, Quantitative Techniques, second ed., Vikas Publishing House Pvt. Limited, New Delhi, 1984.
[37] C.R. Kothari, Research Methodology: Methods and Techniques, 2nd revised edn., New Age International Publishers, New Delhi, 2004.
[38] W.A. Chance, Statistical Methods for Decision Making, 1969.
[39] R.E. Treybal, Mass Transfer Operations International Edition, third ed., McGraw Hill Book Co., Singapore, 1981.
[40] A. Kaushal, S.K. Singh, Removal of Zn (II) from aqueous solutions using agro-based adsorbents, Imp J Interdiscip Res 2 (6) (2016) 1215–1218.
[41] M. Rashad, A. Al-Aoh Hatem, Promising adsorption studies of bromophenol blue using copper oxide nanoparticles, Desalination Water Treat. 139 (2019) 360–368.
[42] Hatem A. Al-Aoh, Adsorption performances of nickel oxide nanoparticles (NiO NPs) towards bromophenol blue dye (BB), Desalination Water Treat. 110 (2018) 229–238.
[43] Yang Xiang, Manglai Gao, Tao Shen, Gaili Cao, Bingbing Zhao, Shangxin Guo, Comparative study of three novel organo-clays modified with imidazolium-based gemini surfactant on adsorption for bromophenol blue, J. Mol. Liq. 286 (2019) 110928.
[44] Yi Liu, Yumin Huang, Aiping Xiao, Huajiao Qiu, Liangliang Liu, Preparation of magnetic Fe3O4/MIL-88A nanocomposite and its adsorption properties for bromophenol blue dye in aqueous solution, Nanomaterials 9 (1) (2019) 51.
[45] Mandla B. Chabalala, Mohammed Z. Al-Abri, Bhekie B. Mamba, Edward N. Nxumalo, Mechanistic aspects for the enhanced adsorption of bromophenol blue and atrazine over cyclodextrin modified polyacrylonitrile nanofiber membranes, Chem. Eng. Res. Des. 169 (2021) 19–32.
[46] Kovo G. Akpomie, Conradie Jeanet, Biogenic and chemically synthesized Solanum tuberosum peel–silver nanoparticle hybrid for the ultrasonic aided adsorption of bromophenol blue dye, Sci. Rep. 10 (1) (2020) 17094.
[47] Kovo G. Akpomie, Soumya Ghosh, Marieka Gryzenhout, Jeanet Conradie, One-pot synthesis of zinc oxide nanoparticles via chemical precipitation for bromophenol blue adsorption and the antifungal activity against filamentous fungi, Sci. Rep. 11 (1) (2021) 8305.
[48] Ali Q. Alorabi, M. Shamshi Hassan, Azizi Mohamed, Fe_3O_4-CuO-activated carbon composite as an efficient adsorbent for bromophenol blue dye removal from aqueous solutions, Arab. J. Chem. 13 (11) (2020) 8080–8091.
[49] S. Palani, S.Y. Liong, P. Tkalich, An ANN application for water quality forecasting, Mar. Pollut. Bull. 56 (9) (2008) 1586–1597.
[50] A.E. Mahdi, Nisreen S. Ali, K.R. Kalash, I.K. Salih, M.A. Abdulrahman, T. M. Albayati, Investigation of equilibrium, isotherm, and mechanism for the efficient removal of 3-nitroaniline dye from wastewater using mesoporous material MCM-48, Progress in Color, Colorants and Coatings 16 (4) (2023) 387–398.
[51] L. Mohajeri, H.A. Aziz, M.H. Isa, M.A. Zahed, A statistical experiment design approach for optimizing biodegradation of weathered crude oil in coastal sediments, Bioresour. Technol. 101 (3) (2010) 893–900.
[52] E. Heidari, M.A. Sobati, S. Movahedirad, Accurate prediction of nanofluid viscosity using a multilayer perceptron artificial neural network (MLP-ANN), Chemom. Intell. Lab. Syst. 155 (2016) 73–85.