JEMAT 2023, Vol 11, No 2, 64-71
https://doi.org/10.54987/jemat.v11i2.894

HIBISCUS PUBLISHER

JOURNAL OF ENVIRONMENTAL MICROBIOLOGY AND TOXICOLOGY

Website: http://journal.hibiscuspublisher.com/index.php/JEMAT/index

A Two-Level Factorial Design for Screening Factors that Influence the Decolorization of Congo Red by *Serratia marcescens* strain Neni-1

Rusnam[1]*, Rahman, M.F.A.[2], Mohd Ezuan Khayat[2,3], Mohd Badrin Hanizam Abdul Rahim[2,3], Motharasan Manogaran[4] and Hafeez Mohd Yakasai[5]

[1]Department of Agricultural Engineering, Faculty of Agricultural Technology, Andalas University, Padang, 25163, Indonesia.
[2]Department of Biochemistry, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
[3]Agribiotechnology Group, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, UPM 43400 Serdang, Selangor, Malaysia.
[4]Malaysia Genome and Vaccine Institute (MGVI) National Institute of Biotechnology Malaysia (NIBM) Jalan Bangi, 43000 Kajang, Selangor, Malaysia.
[5]Department of Biochemistry, Faculty of Basic Medical Sciences, College of Health Science, Bayero University Kano, P. M. B 3011, Kano State-Nigeria.

*Corresponding author:
Prof Dr Rusnam,
Department of Agricultural Engineering,
Faculty of Agricultural Technology,
Andalas University,
Padang, 25163,
Indonesia.
Email: rusnam_ms@yahoo.com

HISTORY

Received: 15th Oct 2023
Received in revised form: 24th Dec 2023
Accepted: 30th Dec 2023

KEYWORDS

Two-Level Factorial Design
Screening Factors
Decolorization
Congo Red
Serratia marcescens

ABSTRACT

The toxicity and persistence of Congo Red pollution make it a major threat to both the environment and human health. Biodegradation, which involves the breakdown of pollutants by microbes like bacteria, fungi, and algae, is an attractive and long-term solution to the problem of Congo Red pollution. Optimization techniques such as Response Surface Methodology can further improve biodegradation processes' efficiency, but initial screening is often needed. Using a two-level factorial design, this study successfully screened five parameters (temperature, pH, incubation time, concentration of Congo Red, and sucrose) that affect Congo red's decolorization by Serratia marcescens strain Neni-1. Additionally, it identified three significant parameters that contribute to optimized decolorization of Congo Red: pH, incubation time, and the concentration of Congo Red. Various diagnostic plots were used to analyze the important contributing factors or parameters, including ANOVA, Pareto's chart, and perturbations plot. The two-level factorial conclusion was supported by diagnostic plots such as half-normal, residual vs runs, Cook's distance, Box-Cox, leverage vs runs, DFBETAS, and DFFITS. Consistent with trends in the published literature, this study found that the majority of Congo Red-degrading microorganisms thrive in nearly neutral environments. In future works, RSM can be used to further optimize the parameters that contributed to the decolorization of this bacterium on Congo Red.

INTRODUCTION

Congo Red, a diazo dye, was first synthesized in 1883 by Paul Böttiger, and it quickly became one of the most prominent dyes used in the textile industry due to its bright red color and affinity for cellulose fibers. The chemical structure of Congo Red consists of two azo groups (-N=N-) and a biphenyl group, making it a complex and highly stable molecule. Despite its industrial advantages, Congo Red has garnered significant attention for its environmental and health impacts [1–5]. The widespread use of Congo Red in the textile industry has led to its release into the environment, primarily through wastewater discharge. Textile effluents containing Congo Red are often inadequately treated before being released into water bodies, leading to significant pollution. The presence of Congo Red in water bodies can cause severe environmental problems, including inhibiting photosynthesis in aquatic plants by blocking sunlight penetration and altering the aquatic ecosystem [6–8].

Congo Red's high solubility in water and resistance to photolytic and chemical degradation make it persistent in the environment. Its bright color is visually unappealing, and even at low concentrations, it can significantly affect the aesthetic quality of water bodies. Furthermore, the adsorption of Congo Red onto

sediment and its subsequent release can lead to long-term contamination [9]. The toxicity of Congo Red is well-documented, posing risks to both aquatic life and human health. The dye is known to cause various toxic effects, including mutagenicity and carcinogenicity. Studies have shown that Congo Red can induce chromosomal aberrations and DNA damage in human lymphocytes, highlighting its genotoxic potential [6–8].

In aquatic environments, Congo Red is toxic to a wide range of organisms, including fish, algae, and microorganisms. The dye can cause respiratory distress in fish by impairing gill function and disrupting the oxygen-carrying capacity of the blood [7]. In addition to direct toxicity, the breakdown products of Congo Red, formed under anaerobic conditions, are often more toxic than the parent compound. For example, aromatic amines released during the reductive cleavage of azo bonds are known carcinogens. Given the environmental persistence and toxicity of Congo Red, its removal from wastewater is critical. Traditional wastewater treatment methods, including physical and chemical processes, often fall short in effectively degrading azo dyes like Congo Red. These methods can be costly and may produce secondary pollutants. As a result, there is growing interest in biological treatment methods, particularly biodegradation, as a sustainable and effective solution for Congo Red removal [6–8].

Biodegradation involves the use of microorganisms to break down complex organic pollutants into simpler, non-toxic compounds. This process can occur under aerobic or anaerobic conditions, with different microorganisms exhibiting varying degradation capabilities. Aerobic biodegradation of Congo Red typically involves bacteria, fungi, and algae. One of the most studied bacteria for Congo Red degradation is Pseudomonas sp. This bacterial species can utilize Congo Red as a carbon and energy source, breaking it down into simpler compounds through enzymatic actions [6–8]. Anaerobic degradation of Congo Red often involves a consortium of microorganisms, including bacteria and archaea, that work synergistically to reduce azo bonds. This process typically results in the formation of aromatic amines, which can then be further degraded under aerobic conditions. Anaerobic-aerobic sequential treatment systems have been developed to enhance the complete mineralization of Congo Red [7].

Experiment planning in fundamental research is often done with a "intuitive" mindset. The "one factor at a time" (OFAT) approach has long been used in biological experiments. Here, we isolate the item under study and examine its output while holding all other variables and factors constant. Due to component interactions, erroneous interpretation can be produced. The complexity of the process necessitates controlling numerous input factors for best results. There may be a lot of interesting data coming in, but the experiment's results may carry large background noise. In such a case, statistically based experimental design allows for the optimization of data point selection to maximize the amount of relevant information obtained, potentially leading to much more intriguing data.

One well-known screening method for finding important components early in the experimentation phase, when thorough knowledge of the system is usually lacking, is the Plackett-Burman (PB) experimental design, named after its developers- statisticians J.P. Burman and Robin L. Plackett, which created it to find important variables with as few experiments as possible. Two-factor interactions can easily confuse major effects when utilizing a Plackett-Burman design. Such designs are appropriate in situations where there is little chance of a two-way exchange of information. Despite its usefulness in two-level multi-factor experiments involving more than four factors, the Plackett-Burman design (PB) does not confirm whether the effect of one factor depends on another. Additionally, insufficient data has been collected to determine the nature of these effects due to its small size. In the screening step, the two-level factorial design outperforms the PB method because it considers the interplay between the distinct components. By calculating the interconnections between important cultural factors, this method yields a more precise estimate of the ideal condition. Using a two-level factorial design has been beneficial for many screening processes in the literature [10–16]. Here we describe the use of a two-level factorial design to screen for significant factors that influence the decolorization of *Serratia marcescens* strain Neni-1 on Congo Red.

MATERIALS AND METHODS

Growth and maintenance of Congo Red-degrading bacterium

Serratia marcescens strain Neni-1 was previously isolated as a Mo-reducer [17]. From an overnight pure culture of Serratia marcescens strain Neni-1 in nutrient broth, 0.1 mL was added into 45 mL of Congo Red enrichment medium in a 100 mL volumetric flask and the culture was incubated at 25 ℃ on an incubator shaker (Certomat R, USA) at 150 rpm for 48 h. The ingredients of the decolorization media (% w/v) were as follows: Glucose (1%), $(NH_4)_2.SO_4$ (0.3%), $MgSO_4.7H_2O$ (0.05%), yeast extract (0.05%), NaCl (0.5%), Na_2HPO_4 (0.705% or 50 mM). The media was adjusted to pH 7.0. Under neutral to basic conditions (pH > 5), Congo Red turns red and to study the effects of pH, the pH was varied from pH 5.8 to 7.8 taking into account the pKa of phosphate buffer. Decolorization was monitored 490 nm to cover maximum absorption values for specific dyes as this wavelength is available in the BioRad 680 microplate reader.

About 180 µL of the deolorization media containing the ingredients and conditions, including various Congo Red concentrations to be studied, was sterically pipetted into each well of a sterile microplate. Then 20 µL of the bacterium from a 48 h culture of the bacterium grown in the media above was then added to each well to initiate Mo-blue production. A sterile sealing tape that allows gas exchange (Corning® microplate) was used to seal the tape. The microplate was incubated statically at room temperature. At defined times absorbance was taken. The lambda maximum for Congo red is 498 nm [18]. Furthermore, these predetermined wavelengths are used because, typically, water soluble dyes have shallow maximum absorption spectra and a discrepancy of 20 nm from that peak does not result in a significant drop in absorbance. A percentage of decolorization was determined by subtracting the initial absorbance values from the final measurements taken after 48 hours of incubation.

Screening of significant parameters using two level factorial design

The primary objective of utilizing the two-level factorial design was to determine the relative significance of various factors that had an impact, even amid complex interconnections. We conducted a 2-factorial design using a total of five components. The code -1 denoted the lesser value, whereas the code 1 represented the bigger value. The result observed or response was the decolorization percentage. The tests were executed in accordance with the order outlined in **Table 1**. The experiment was conducted twice, and the findings of both trials are presented here, along with their respective means. The data were analyzed using Design Expert 7.0, a software developed by Stat-Ease, Inc. (trial version) to identify the parameters that have a much greater impact than others.

Table 1. Coded and actual range of values for Two-Level Factorial Design.

Factor	Name	Units	Minimum	Maximum	Coded Low	Coded High	Mean	Std. Dev.
A	pH		5.80	7.80	-1 ↔ 5.80	+1 ↔ 7.80	6.80	1.03
B	Temperature	°C	20.0	40.0	-1 ↔ 20.00	+1 ↔ 40.0	30.0	10.33
C	Congo Red	g/L	0.10	1.0	-1 ↔ 0.10	+1 ↔ 1.00	0.55	0.4648
D	Sucrose	g/L	1.0	10.0	-1 ↔ 1.00	+1 ↔ 10.0	5.5	4.65
E	Incubation time	Days	1	4	-1 ↔ 1.00	+1 ↔ 4	2.5	1.55

Statistical Analysis

The values represent the mean plus or minus the standard deviation. Experiments were carried out three times. The comparison between groups was conducted using either one-way analysis of variance with post hoc analysis using Tukey's test or Student's t-test. A p-value less than 0.05 was deemed significant. Values will be reduced to three decimal points where necessary.

RESULTS

A basic method used in experimental design, two-level factorial designs allow for the simultaneous examination of numerous components' impacts. They shine in industrial experiments where understanding the interplay of several components is crucial for process optimization. Each factor in a two-level factorial design has two levels, typically denoted as +1 and -1, that reflect the high and low settings of the factor, respectively [19]. Power analysis is an essential element in experimental design, as it assesses the test's capacity to detect a specific effect size with a defined level of confidence.

The ideal power of a test, defined as the likelihood of correctly rejecting the null hypothesis when the alternative hypothesis is true, should be 80% or higher. This means there is an 80% chance of discovering a true effect if it exists. Power analysis is used in two-level factorial designs to determine the probability of detecting significant main effects and interactions between factors. This is especially useful when considering a specific signal-to-noise ratio (SNR), which measures the strength of the effect compared to the variability in the data [20]. The alpha level (α) is a key component that influences power. It is usually set at 5% and represents a Type I error probability, which occurs when a genuine null hypothesis is incorrectly rejected. Reducing alpha levels decreases the probability of Type I error but can also diminish statistical power [19].

The signal-to-noise ratio (SNR), defined as the ratio of the effect amplitude to the error term's standard deviation, indicates a more distinct signal in the presence of noise, enhancing the test's power [20]. Increasing the sample size of an experiment often enhances its statistical power by yielding more data and reducing the standard error of the estimations. The power of a study is influenced by the magnitude of the effect under investigation, as larger effects are more readily detectable (Montgomery, 2017). The number of components and their levels in a factorial design directly impact the number of experimental runs. Increasing the number of factors or levels leads to a higher level of complexity and requires a larger sample size to ensure sufficient statistical power [20].

A design is deemed appropriately powered to detect the required effect sizes if the computed power is 80% or above. If the sample size is smaller, it may be essential to increase it or reevaluate the parameters of the experimental design (Cohen, 1988). Power analysis for regular two-level factorial designs ensures that experiments have the ability to detect significant effects, while also balancing the risks of Type I and Type II errors. This analysis provides a strong framework for drawing reliable conclusions from experimental data [19,20]. The results of this investigation showed a power of above 80%, with a specific value of 94.9%, indicating that the study had sufficient power (**Table 2**).

Table 2. Regular Two-Level Factorial Design design power.

Name	Units	Delta (Signal)	Sigma (Noise)	Signal/Noise	Power for A	Power for B	Power for C	Power for D	Power for E
Decolorization %	%	2	1	2	94.9%	94.9%	94.9%	94.9%	94.9%

Two-level factorial design for screening the operational factors

The five operational parameters considered for the factor screening investigation were temperature, incubation period, concentration of Congo Red, and sucrose concentrations. The study used a regular two-level factorial design. The bacterial decolorization rate varied between 77 and 97%, covering the whole range of values that were considered. The experimental values, projected response values, and actual values of the experiment's variables are all shown in **Table 3**, which also illustrates the design plan.

Table 3 displays the results of an analysis of variance (ANOVA), the F-test, and the P-value for a selected factor. With these checks, we can see whether the model is statistically significant. The results demonstrated that the model is very significant, as indicated by the low P value of less than 0.0001 and the F value of 18.22. A low P value for the model makes this very evident. You can check the model's reliability by calculating the correlation coefficient (R2: 0.9011, closer to unity) and the adjusted correlation coefficient (Adj R2: 0.8516), which means that 85.16 percent of the total variance in the response data is accounted for. A finding of 12.738, for the adequacy accuracy means that the model has a suitable signal to use in exploring the design space. In this case, A (pH), C (Congo red concentration), E (incubation time), and the interaction AE were significant model terms, as confirmed by P-values <0.05. Predicted bacterial decolorization as a response can be obtained and expressed in terms of coded and actual factors equation by applying the two-factor interactive method (**Table 4**).

Table 3. A two-level factorial analysis using analysis of variance (ANOVA).

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Model	420.58	5	84.12	18.22	< 0.0001	significant
A-pH	44.36	1	44.36	9.61	0.0113	
B-Temperature	0.3186	1	0.3186	0.0690	0.7981	
C-Congo Red	86.05	1	86.05	18.63	0.0015	
D-Sucrose	0.0050	1	0.0050	0.0011	0.9743	
E-Incubation time	289.84	1	289.84	62.77	< 0.0001	
Residual	46.18	10	4.62			
Cor Total	466.76	15				
Std. Dev.	2.15	R^2	0.9011			
Mean	86.60	Adjusted R^2	0.8516			
C.V. %	2.48	Predicted R^2	0.7467			
		Adeq Precision	12.7386			

Table 4. Bacterial decolorization prediction based on equations involving coded and actual factors.

Coded		Actual	
Decolorization	Factor	Decoloriza-tion	Factor
+86.60		+70.61810	
+1.67	A	+1.66509	pH
+0.1411	B	+0.014112	Temperature
-2.32	C	-5.15354	Congo Red
-0.0177	D	-0.003941	Sucrose
+4.26	E	+2.83747	Incubation time

Table 5 lists the investigated components' estimated coefficients along with their corresponding standard errors, confidence limits, and variance inflation factors (VIF). Of the components that were chosen, the only ones with positive coefficients are incubation time and pH; of these, incubation time has a higher positive value. Accordingly, it appears that both parameters positively affect the growth of this bacterium on Congo Red, but the incubation duration has a stronger positive impact. Conversely, a negative value for the coefficient estimate of the Congo Red concentration indicates that fed more than the ideal concentration of Congo Red hinders the growth of this bacterium.

How much non-orthogonality in the design increases the variance of the model's coefficients is measured by the variance inflation factor (VIF). The VIF design has a higher standard error for a model coefficient than the orthogonal design by a factor equal to the square root of the VIF. An optimal VIF is 1, which suggests that the coefficient is orthogonal to the other model components and has a correlation coefficient of 0. However, VIFs over 10 may be suspicious. VIFs larger than 100 show multicollinearity-related coefficient errors, and VIFs greater than 1,000 indicate severe collinearity. Variance inflation factor (VIF) = 1, indicating multicollinearity in regression analysis [21–23]. Based on the result obtained, out of five screened parameters, only three forms major influential factors as obtained through two-level factor analysis.

Table 5. Coefficient estimate obtained during ANOVA for two-level factorial design.

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	86.60	1	0.5372	85.40	87.80	
A-pH	1.67	1	0.5372	0.4681	2.86	1.0000
B-Tempera-ture	0.1411	1	0.5372	-1.06	1.34	1.0000
C-Congo Red	-2.32	1	0.5372	-3.52	-1.12	1.0000
D-Sucrose	-0.0177	1	0.5372	-1.21	1.18	1.0000
E-Incubation time	4.26	1	0.5372	3.06	5.45	1.0000

Fig. 1 displays the Pareto charts that were made to examine the statistical significance of each response coefficient. The two limit lines utilized in the Pareto chart to classify the t-value of the effect (2.228) are the Bonferroni limit line (t-value of effect: 3.827) and the t-limit line. When classifying coefficient importance, there are three separate groups to consider. The significance level of a coefficient is determined by how quickly its t-value of effect rises above the Bonferroni threshold. Coefficients two and three, with t-values of effect that lie between the Bonferroni line and the t-limit line, are considered to have a high likelihood of being statistically significant; coefficients one and two are not to be included in the analysis because they are not statistically significant. The t-values for both of these coefficients fall in the middle of the two lines representing the Bonferroni and t-limit tests. According to a Pareto chart, the variables that have the most positive impact are the incubation period and pH, whereas the concentration of Congo Red has a significant but negative impact. The results obtained using the coefficient estimate are comparable to these.

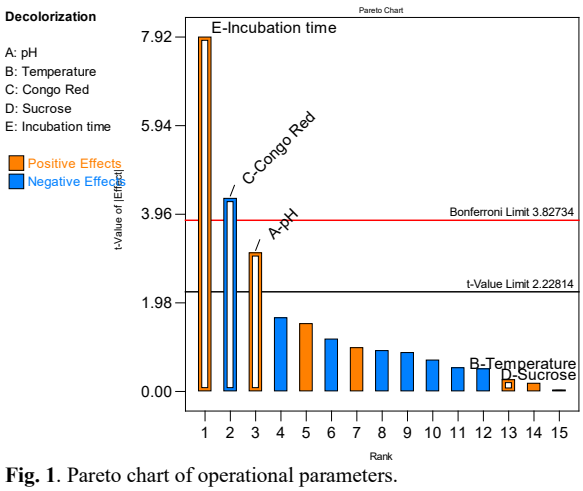

Fig. 1. Pareto chart of operational parameters.

In order for this bacterium to grow on Congo Red, the most important factors were the concentration of Congo Red, the pH of the medium, and the amount of time it was left to incubate. These features have been identified in various OFAT-based methods as critical for the high microbial decolorization on Congo Red. The concentrations of Congo Red used in this study were well within the range that most microorganisms that can degrade Congo Red are known to tolerate. The level of harm that Congo Red causes to microbes depends on the microbes themselves and their environment. The amounts of Congo Red that are toxic to different bacteria have been the subject of several studies. Studies have shown that Congo Red can inhibit bacterial growth and activity at concentrations as low as 10-50 mg/L. For example, it has been noted that at these doses, the growth of some species of *Bacillus* and *Pseudomonas* is hindered [6]. Unlike bacteria, fungi like *Phanerochaete chrysosporium* can withstand larger quantities of Congo Red. Nonetheless, levels beyond 100 mg/L can still have harmful effects and hinder the growth of fungi and enzyme activity [24]. Certain types of algae are particularly vulnerable to the effects of Congo Red.

Photosynthesis and development in different species of algae can be negatively impacted by concentrations as low as 1-10 mg/L [8]. Tolerance might vary greatly in mixed microbial cultures because various species interact with each other. According to [7], certain consortia can efficiently break down Congo Red up to 200 mg/L, but as the concentrations rise, the microbes' efficiency and health suffer a sharp reduction. In order to examine the statistical importance of each response coefficient, a Pareto chart was generated (**Fig. 1**). The Bonferroni limit line (t-value of effect: 3.728) and the t-limit line (t-value of effect: 2.201), which are two distinct sets of t-values, are shown in the Pareto graphic. In general, there are three main contexts in which coefficients are useful. For a coefficient to be considered significant, its t-value of effect must be higher than the Bonferroni threshold, and the parameters E and C are the two parameters to do so. pH,

with t-values of effect that lie between the Bonferroni line and the t-limit line, is considered to have a high likelihood of being statistically significant; the rests are not to be included in the analysis because they are not statistically significant.

The important contributing parameters in the decolorization of this bacterium on Congo Red were the incubation time and pH giving positive effects and Congo Red concentration, which gave a negative effect. These are factors that have been identified in several OFAT-based approaches as being important in contributing high decolorization of microorganisms on Congo Red [25–33]. This investigation was conducted out using Congo Red concentrations that fell well within the range that previous research has shown to be tolerated by the majority of bacteria capable of Congo Red degradation.

The perturbation plot (**Fig. 2**) offers valuable insights into the impact of different factors on the decolorization of Congo Red by bacteria. The plot demonstrates that both pH and incubation time significantly influence the decolorization process. pH shows a prominent slope, while incubation time also has a substantial effect. In contrast, the plot shows that temperature, Congo Red concentration, and sucrose concentration have relatively stable or unchanging values. This suggests that although changes in pH and incubation time significantly impact the bacterium's ability to remove color, the process remains relatively consistent when it comes to different temperatures, dye concentrations, and sucrose levels.

Therefore, it is crucial to optimize the pH and prolong the incubation time to improve decolorization efficiency. However, controlling temperature, Congo Red concentration, and sucrose concentration within the specified range is not as crucial. This comprehension allows for modifications in industrial procedures to attain enhanced outcomes in the decolorization of Congo Red. The perturbation plot clearly shows an interaction between the incubation period and pH, which would be ignored by the Plackett-Burman screening method [34–37].

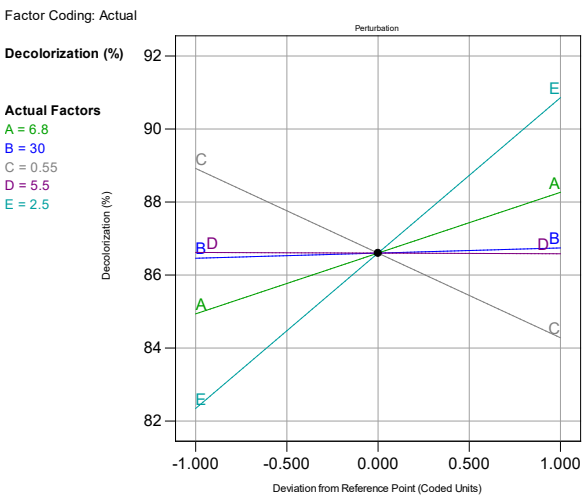

Fig. 2. Operating parameters obtained through a standard two-factor design and shown in a perturbation plot.

To confirm the assumption of normality, we created and analyzed a half-normal probability plot of the residuals **(Fig. 3)**.

This plot aims to ascertain if the disparities between the anticipated and observed values conform to a normal distribution. Assessing the alignment of the residuals with a line, which represents the expected distribution pattern, is essential for this evaluation. The majority of data points in the plot are in close proximity to this line, suggesting that the residuals generally conform to a normal distribution. Nevertheless, there is one remarkable anomaly, namely Run 6, situated on the left side, which diverges considerably from this pattern. This discrepancy indicates that the residual linked to Run 6 deviates significantly from what would be anticipated assuming normality.

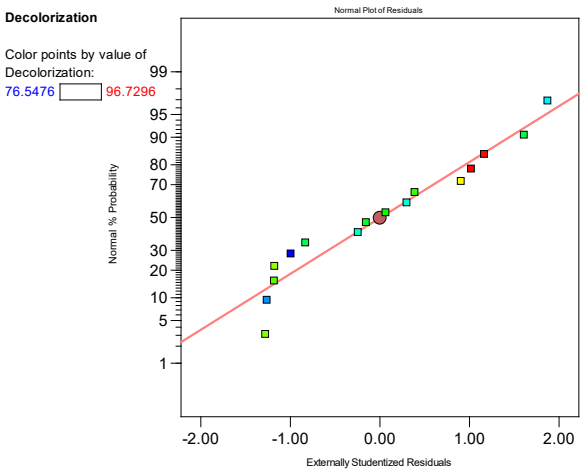

Fig. 3. For the two-level factorial optimization experiments, diagnostics were plotted using a normal plot of residuals.

Fig. 4 shows a good match between the model's predicted values and the actual experimental results. For determining which power law transformation is most suitable given the value of lambda, the Box-Cox plot (**Fig. 5**) provides useful information. Avoid further transforming the observed response to fit the model; the 95% confidence interval already has a value of 1, which is the value that was designed into the model.

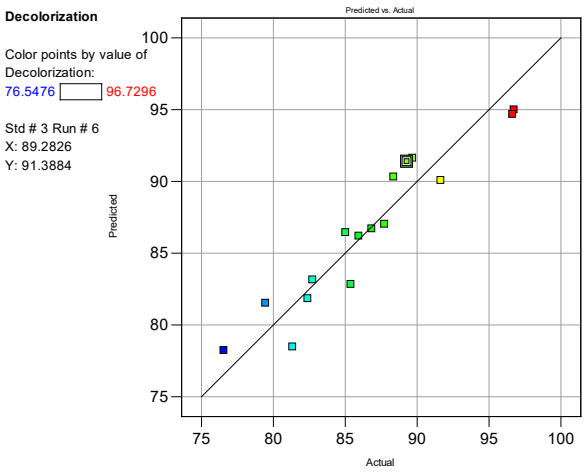

Fig. 4. Diagnostic's plot for the two-level factorial optimization studies, showing the relationship between predicted and actual results.

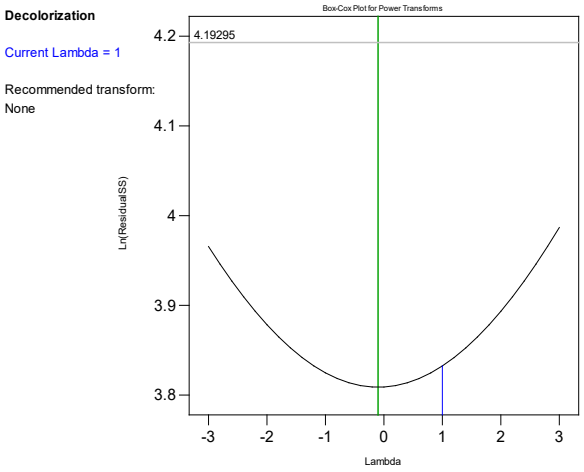

Fig. 5. Diagnostic's plot in the form of Box-Cox plot for the two-level factorial optimization studies.

Fig. 6 displays a graph showing the relationship between leverages and runs. The graph illustrates that all the numerical values obtained are within the standard range of 0 to 1. This range indicates that no specific design point excessively affects the model's accuracy. If a data point is flawed, a high leverage value greater than one is considered problematic because it would significantly impact the model. The leverage plot demonstrates that no data points surpass the average leverage, suggesting the absence of any problematic data. In addition, Cook's distances can be used to quantify response outliers, similar to those observed in an experimental trial (**Fig. 7**).

A higher value of Cook's distance, which is always positive, indicates a more influential observation. Researchers commonly employ a threshold of three times the average value of Cook's D to ascertain the significance of an observation. In this instance, there are no outliers identified, as all Cook's distance values fall within the acceptable threshold of 1. Finally, **Fig. 8** illustrates the residuals plotted against the run data. The lack of discernible serial correlation in this plot indicates that the data is random, thereby bolstering the assumption of randomness in the dataset [38–42].

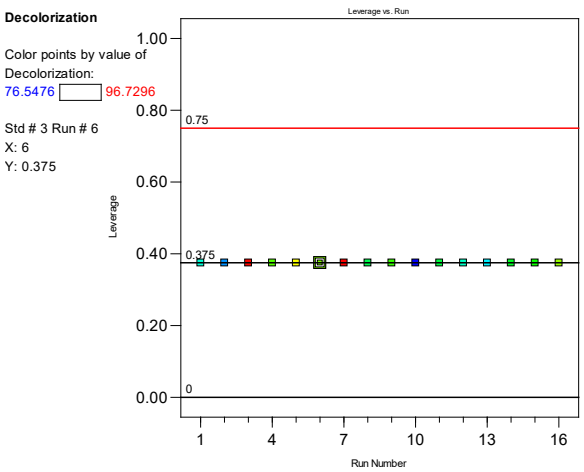

Fig. 6. The two-level factorial optimization studies' diagnostic plot based on leverage versus runs.

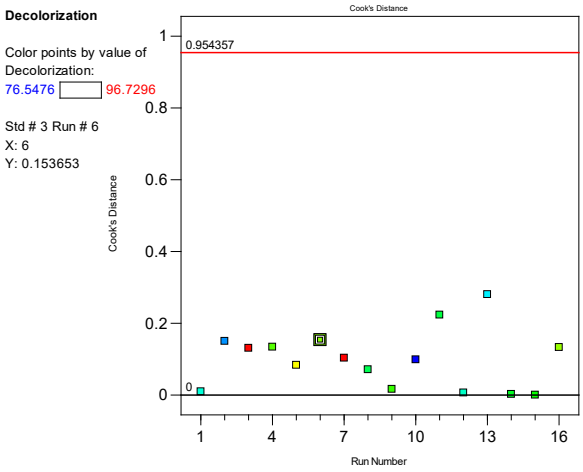

Fig. 7. The two-level factorial optimization studies' diagnostic plot based on Cook's distance versus runs.

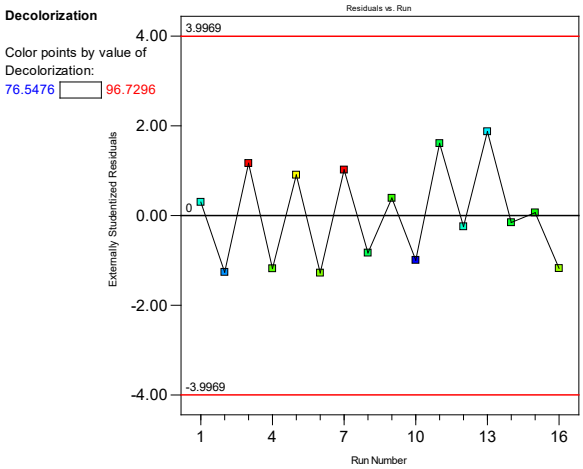

Fig. 8. The diagnostic plot for the two-level factorial optimization studies is shown as residuals vs runs.

It is crucial to acknowledge and consider significant observations that have a strong impact to uphold a study's credibility and reliability. To prevent coefficient estimations and model interpretations from being skewed, it is necessary to eliminate outliers or data entry errors that lead to excessive influence measures. Integrating these essential data points during the process of model fitting can result in erroneous conclusions. In the past, the process of identifying outliers involved using histograms and scatterplots, which were performed prior to conducting linear regression analysis. Nevertheless, these approaches were based on personal opinions and did not measure the influence of specific extreme values on the ultimate outcomes. In order to tackle this issue, quantitative measurements such as DFFIT and DFBETA were created. The DFFITS algorithm assesses the influence of each example on the predicted value. DFFITS values have the ability to be either positive or negative, unlike Cook's distances which are limited to being positive only. A DFFITS value of 0 indicates that the data point is exactly on the regression line.

The possibility arises from the concept of leverage, which elucidates the connection between an observed value, its anticipated value, and an unobserved value. DFFITS is a mathematical measure that is calculated by multiplying the externally studentized residual (ti) by the leverage. High leverage points have a positive impact on the DFFITS value, whereas low leverage points have a negative impact on it. This facilitates a more intricate comprehension of how individual data points impact the model, allowing for more accurate identification and management of influential observations [37,43,44]. The plots show the DFBETAS values (**Fig. 9**) were within the size-adjusted threshold acceptable range while the DFFITS values were within the cut-off values (**Fig. 10**).

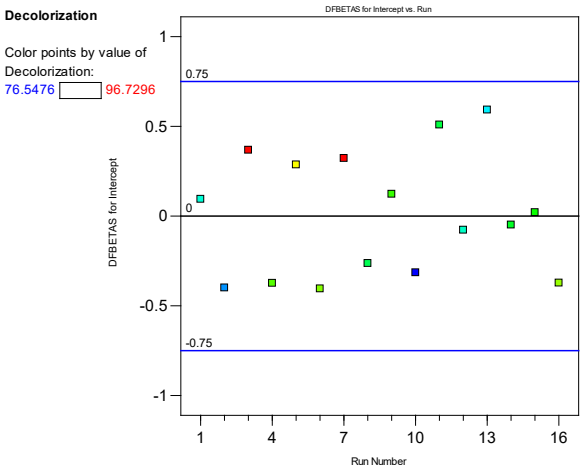

Fig. 9. The diagnostic plot for the two-level factorial optimization studies is shown as DFBETAS, which stands for intercept versus runs.

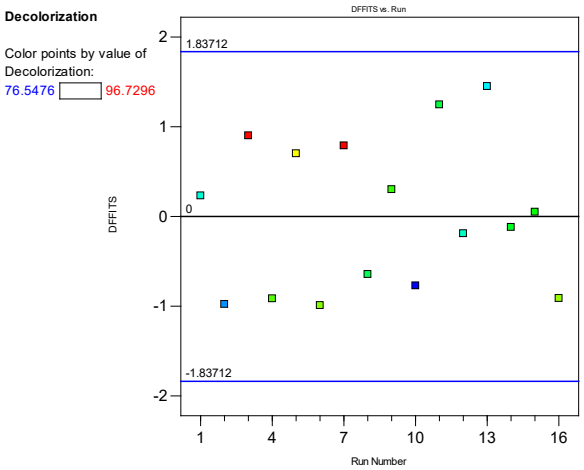

Fig. 10. The diagnostic plot for the two-level factorial optimization studies is presented as DFFITS vs runs.

CONCLUSION

The five independent factors influencing the decolorization of *Serratia marcescens* strain Neni-1 on Congo Red were screened using a two-level factorial design. The concentration of sucrose, pH, temperature, incubation time, and Congo Red are all relevant variables. This bacterium successfully decolorized Congo Red using a two-factor factorial design, which successfully identified the concentration of Congo Red, pH, and incubation time as critical significant parameters. These parameters can be further optimized in future works using RSM. Analytical tools such as analysis of variance (ANOVA), pertubation plot, Pareto chart and other diagnostic plots were used to examine the significant contributing factors or parameters. Box-Cox, DFBETAS, DFFITS, Cook's distance, half-normal, and other diagnostic plots all corroborated the two-level factorial conclusion. Many microorganisms that degrade Congo Red have an incubation time ranging from two to five days for optimized decolorization and are inhibited by high concentrations of Congo red, so this is a predicted result. A longer incubation time allows for more decolorization. The results obtained in this study align with the trends in the published literature regarding the optimal growth conditions for most microorganisms that degrade Congo Red.

REFERENCES

1. Abu Talha M, Goswami M, Giri BS, Sharma A, Rai BN, Singh RS. Bioremediation of Congo red dye in immobilized batch and continuous packed bed bioreactor by *Brevibacillus parabrevis* using coconut shell bio-char. Bioresour Technol. 2018 Mar 1;252:37–43.
2. Mondal NK, Kar S. Potentiality of banana peel for removal of Congo red dye from aqueous solution: isotherm, kinetics and thermodynamics studies. Appl Water Sci. 2018 Sep 18;8(6):157.
3. Litefti K, Freire MS, Stitou M, González-Álvarez J. Adsorption of an anionic dye (Congo red) from aqueous solutions by pine bark. Sci Rep. 2019 Nov 11;9(1):16530.
4. Igwegbe CA, Ighalo JO, Onyechi KK, Onukwuli OD. Adsorption of Congo red and malachite green using H3PO4 and NaCl-modified activated carbon from rubber (Hevea brasiliensis) seed shells. Sustain Water Resour Manag. 2021 Jul 28;7(4):63.
5. Philip I, Sarojini S, Biswas S, Jayaram S. Exploring the Potential of Bacillus velezensis, an Endophytic Bacteria Isolated from Alternanthera philoxeroides for Plant Growth Promotion and Bioremediation Properties. J Pure Appl Microbiol. 2023 Sep 2;17.
6. Pearce CI, Lloyd JR, Guthrie JT. The removal of colour from textile wastewater using whole bacterial cells: a review. Dyes Pigments. 2003 Sep 1;58(3):179–96.
7. Pandey A, Singh P, Iyengar L. Bacterial decolorization and degradation of azo dyes. Int Biodeterior Biodegrad. 2007 Mar 1;59(2):73–84.
8. Vandevivere P, Biachi R, Verstraete W. ChemInform Abstract: Treatment and Reuse of Wastewater from the Textile Wet-Processing Industry: Review of Emerging Technologies. Cheminform. 2010 Oct 27;29.
9. Forgacs E, Cserháti T, Oros G. Removal of synthetic dyes from wastewaters: a review. Environ Int. 2004 Sep 1;30(7):953–71.
10. Liao CT. Two-level factorial designs for searching dispersion factors and estimating location main effects. J Stat Plan Inference. 2006;136(11):4071–87.
11. Hedayat AS, Pesotan H. Tools for constructing optimal two-level factorial designs for a linear model containing main effects and one two-factor interaction. J Stat Plan Inference. 2007;137(4):1452–63.
12. Bezerra MA, Santelli RE, Oliveira EP, Villar LS, Escaleira LA. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. Talanta. 2008 Sep 15;76(5):965–77.
13. Saravanan P, Pakshirajan K, Saha P. Kinetics of phenol and m-cresol biodegradation by an indigenous mixed microbial culture isolated from a sewage treatment plant. J Environ Sci. 2008;20(12):1508–13.
14. Gottipati R, Mishra S. Process optimization of adsorption of Cr(VI) on activated carbons prepared from plant precursors by a two-level full factorial design. Chem Eng J. 2010;160(1):99–107.
15. Ting CP. D-optimal two-level fractional factorial designs of resolution III with two dispersion factors. Commun Stat - Theory Methods. 2010;39(14):2515–32.
16. Chang FK, Ting CP. Optimal two-level fractional factorial designs for location main effects with dispersion factors. Commun Stat - Theory Methods. 2011;40(11):2035–43.
17. Gusmanizar N, Halmi MIE, Rusnam M, Rahman MFA, Shukor MS, Azmi NS, et al. Isolation and characterization of a molybdenum-reducing and azo-dye decolorizing *Serratia marcescens* strain Neni-1 from Indonesian soil. J Urban Environ Eng. 2016;10(1):113–23.
18. Sladewski TE, Shafer AM, Hoag CM. The effect of ionic strength on the UV–vis spectrum of congo red in aqueous solution. Spectrochim Acta A Mol Biomol Spectrosc. 2006 Nov 1;65(3):985–7.

19. Montgomery DC. Design and Analysis of Experiments. 10th edition. Hoboken, NJ: Wiley; 2020. 688 p.
20. Box GEP, Hunter JS, Hunter WG. Statistics for Experimenters: Design, Innovation, and Discovery, 2nd Edition. 2nd edition. Hoboken, N.J: Wiley-Interscience; 2005. 633 p.
21. Bonate PL. Linear Models and Regression. In: Bonate PL, editor. Pharmacokinetic-Pharmacodynamic Modeling and Simulation [Internet]. Boston, MA: Springer US; 2011 [cited 2022 Dec 24]. p. 61–100. Available from: https://doi.org/10.1007/978-1-4419-9485-1_2
22. Mundry R. From nonparametric tests to mixed models: A brief overview of statistical tools frequently used in comparative psychology. In: APA handbook of comparative psychology: Basic concepts, methods, neural substrate, and behavior, Vol 1. Washington, DC, US: American Psychological Association; 2017. p. 157–77. (APA handbooks in psychology®).
23. Fox J. Regression Diagnostics: An Introduction. SAGE Publications; 2019. 138 p.
24. Pointing S. Feasibility of Bioremediation by White-Rot Fungi. Appl Microbiol Biotechnol. 2001 Nov 1;57:20–33.
25. An SY, Min SK, Cha IH, Choi YL, Cho YS, Kim CH, et al. Decolorization of triphenylmethane and azo dyes by Citrobacter sp. Biotechnol Lett. 2002 Jun 1;24(12):1037–40.
26. Gill PK, Arora DS, Chander M. Biodecolourization of azo and triphenylmethane dyes by Dichomitus squalens and Phlebia spp. J Ind Microbiol Biotechnol. 2002 Apr;28(4):201–3.
27. Gopinath KP, Asan Meera Sahib H, Muthukumar K, Velan M. Improved biodegradation of Congored by using Bacillus sp. Bioresour Technol. 2009 Jan 1;100(2):670–5.
28. Gopinath KP, Murugesan S, Abraham J, Muthukumar K. Bacillus sp. mutant for improved biodegradation of Congo red: Random mutagenesis approach. Bioresour Technol. 2009 Dec 1;100(24):6295–300.
29. Chakraborty S, Basak B, Dutta S, Bhunia B, Dey A. Decolorization and biodegradation of congo red dye by a novel white rot fungus *Alternaria alternata* CMERI F6. Bioresour Technol. 2013;147:662–6.
30. Raja P, Chellaram C, Jebasingh SEJ, Maheshwari N, Chandrika M, Gladis C. Bio-degradation of harmful textile dyes by marine bacteria from Tuticorin coastal Waters Southeastern India. J Chem Pharm Res. 2013;5(7):146–51.
31. Ng IS, Chen T, Lin R, Zhang X, Ni C, Sun D. Decolorization of textile azo dye and Congo red by an isolated strain of the dissimilatory manganese-reducing bacterium *Shewanella xiamenensis* BC01. Appl Microbiol Biotechnol. 2014;98(5):2297–308.
32. Ning XA, Yang C, Wang Y, Yang Z, Wang J, Li R. Decolorization and biodegradation of the azo dye Congo red by an isolated *Acinetobacter baumannii* YNWH 226. Biotechnol Bioprocess Eng. 2014;19(4):687–95.
33. Yusuf F, Yakasai HM, Usman S, Muhammad JB, Yaú M, Jagaba AH, et al. Dyes-decolorizing potential of fungi strain BUK_BCH_BTE1 locally isolated from textile industry effluents: Characterization and LC-MS analysis of the metabolites. Case Stud Chem Environ Eng. 2023 Dec 1;8:100453.
34. Dey A, Suen CY, Das A. Asymmetric fractional factorial plans optimal for main effects and specified two-factor interactions. Stat Sin. 2005;15(3):751–65.
35. Bailey RA, Łacka A. Nested row-column designs for near-factorial experiments with two treatment factors and one control treatment. J Stat Plan Inference. 2015;165:63–77.
36. Lawson J. Comparison of conditional main effects analysis to the analysis of follow-up experiments for separating confounded two-factor interaction effects in 2 IVk−p fractional factorial experiments. Qual Reliab Eng Int. 2020;36(4):1454–72.
37. Chobisa D. Design of Experiments for the Development of Injectable Drug Products. In: Beg S, editor. Design of Experiments for Pharmaceutical Product Development: Volume II : Applications and Practical Case studies [Internet]. Singapore: Springer; 2021 [cited 2022 Dec 24]. p. 69–96. Available from: https://doi.org/10.1007/978-981-33-4351-1_5
38. Draper NR, Stoneman DM. Factor Changes and Linear Trends in Eight-Run Two-Level Factorial Designs. Technometrics. 1968;10(2):301–11.
39. Hedayat AS, Pesotan H. Designs for two-level factorial experiments with linear models containing main effects and selected two-factor interactions. J Stat Plan Inference. 1997;64(1):109–24.
40. Leitnaker MG, Mee RW. Analytic use of two-level factorials in incomplete blocks to examine the stability of factor effects. Qual Eng. 2001;14(1):49-58+xii.
41. Wei J, Carroll RJ, Harden KK, Wu G. Comparisons of treatment means when factors do not interact in two-factorial studies. Amino Acids. 2012;42(5):2031–5.
42. Parui S, Parsad R, Mandal BN. Efficient block designs for incomplete factorial experiments for two factors with unequal block sizes. Commun Stat - Theory Methods. 2021;50(11):2531–45.
43. Cornell JA. Classical and Modern Regression With Applications. Technometrics. 1987 Aug 1;29(3):377–8.
44. Aydin D, Alma ÖG. Diagnostics of influential observations in partially linear models. Erzincan Univ J Sci Technol. 2015 Jan 2;8(1):51–68.