# FORECASTING THE ADSORPTION CAPACITY OF ORGANIC DYE BY USING ZIRCONIUM-BASED METAL-ORGANIC FRAMEWORK (MOF): COMPARISON STUDIES BETWEEN RESPONSE SURFACE AND NEURAL NETWORK MODELS

**VESHMEN A/L POOPATHI**

**UNIVERSITI SAINS MALAYSIA**

**2021**

# FORECASTING THE ADSORPTION CAPACITY OF ORGANIC DYE BY USING ZIRCONIUM-BASED METAL-ORGANIC FRAMEWORK (MOF): COMPARISON STUDIES BETWEEN RESPONSE SURFACE AND NEURAL NETWORK MODELS

**by**

**VESHMEN A/L POOPATHI**

**Thesis submitted in fulfilment of the requirements**
**for the degree of**
**Chemical Engineering**

**JULY 2021**

# ACKNOWLEDGEMENT

I would like to express my gratitude to the almighty for giving me this opportunity of carrying out my very own final year project and giving me all the strength in completing this thesis. First and foremost, my genuine gratefulness to the Dean and all the staffs in School of Chemical Engineering for making sure I am technically equipped and providing me technical related suggestions at any given time. I would like to express my heartfelt appreciation to my FYP coordinator, Prof Dr. Mohd Roslee Othman, for his excellent advice and never-ending encouragement throughout this project. This final year project will not be finished on time without his assistance and direction. A very special gratitude goes out to my supervisor, Dr. Irvan Dahlan for guiding and assisting me throughout my thesis regardless of time and place. Thanks to my supervisor for his endless encouragement and all the sincere knowledge shared to me without any limitations. Thanks to my parents, for always comforting me and being my cheerleader at any low times during the completion of my thesis. Thank you, dear friends, for all the love and tremendous support given to me throughout the completion of my final year project.

## TABLE OF CONTENTS

## LIST OF TABLES

**Page**

## LIST OF FIGURES

# LIST OF SYMBOLS

| | |
|---|---|
| $i$ | Component Index |
| $k$ | Number of Variables |
| $m$ | Total Number of Levels in the Design |
| $n_i$ | Number of Observations - |
| $p$ | Number of Parameter of Model |
| $R$ | Coefficient of Correlation |
| $R^2$ | Coefficient of Determination |
| $SS_{lof}$ | Sum of Square due to Lack of Fit |
| $SS_{pe}$ | Sum of Square due to Pure Error |
| $SS_{reg}$ | Sum of Square due to Regression |
| $SS_{res}$ | Sum of Square due to Residual |
| $SS_{tot}$ | Total Sum of Square |
| $x_i$ | Variables |
| $x_n$ | Normalized Experimental Factor |
| $\varepsilon$ | Residual Associated to the Experiments |
| $\bar{y}$ | Overall Media |
| $\hat{y}_i$ | Estimated Value by the Model for Level i |
| $y_{ij}$ | Replicates Performed in Each Individual Levels |
| $\beta_{ij}$ | Coefficients of Interaction Parameters |
| $\beta_o$ | Constant Term |
| $\beta_i$ | Coefficients of Linear Parameters |

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ANN | Artificial Neural Network |
| ANOVA | Analysis of Variance |
| BBD | Box-Behnken Design |
| BOD | Biochemical Oxygen Demand |
| BP | Back-Propagation |
| CCD | Central Composite Design |
| COD | Chemical Oxygen Demand |
| CTAB | Cetyl Trimethyl Ammonium Bromide |
| CV | Coefficient of Variance |
| DC | Direct Current |
| DMF | N, N dimethylformamide |
| DOE | Department of Environment |
| DR2B | Direct red 2B dye |
| HCl | Hydrochloric acid |
| LMA | Levenberg-Marquardt Algorithm |
| MOF | Metal Organic Framework |
| MS | Media of Square |
| MSE | Mean Squared Error |
| NaOH | Sodium hydroxide |
| PANI | Copper (II) Chloride doped Polyaniline |
| RMSE | Root Mean Squared Error |
| RO | Reverse Osmosis |
| RPM | Rotation Per Minute |
| RSM | Research Surface Methodology |
| SS | Sum of Square |
| UF | Ultrafiltration |
| UV-Vis | Ultraviolet Visible |

## LIST OF APPENDICES

# RAMALAN KAPASITI PENJERAPAN PEWARNA ORGANIK DENGAN MENGGUNAKAN KERANGKA LOGAM-ORGANIK (MOF) BERASASASKAN ZIRCONIUM : KAJIAN PERBANDINGAN ANTARA KAEDAH PERMUKAAN SAMBUTAN DAN RANGKAIAN NEURAL

## ABSTRAK

Faktor-faktor yang mempengaruhi kapasiti penerapan Zirkonium MOF telah dianalisiskan termasuklah pH, masa sentuhan, jisim adsorben dan kepekatan pewarna awal. Eksperimen dijalankan berdasarkan reka bentuk komposit berpusat (CCD) dalam kaedah permukaan sambutan (RSM). Hasil eksperimen digunakan untuk menyiasat kesan faktor input pada kapasiti penjerapan Zirkonium MOF dan untuk mengembangkan model untuk meramalkan prestasi sistem. Menurut plot permukaan sambutan, kapasiti penjerapan Zirkonium MOF yang tinggi dapat dicapai dengan adsorben yang kurang dan kepekatan pewarna yang tinggi. RSM digunakan untuk membuat model matematik, dan prestasi model dinilai menggunakan analisis varians (ANOVA). Model rangkaian neural dibina dengan menggunakan peralatan rangkaian neural dalam Matlab, dan operasi rangkaian (net operation) dan fungsi peramal (predictor function) dalam Mathematica. Kapasiti penjerapan Zirkonium MOF diramalkan menggunakan model rangkaian neural dan Matematica. Oleh kerana kekurangan data eksperimen untuk latihan rangkaian neural, model matematik yang dihasilkan dalam RSM mempunyai ketepatan yang tinggi dalam meramalkan respon output, dengan $R^2$ 0.97 dan RMSE 2.87. RSM melakukan pengoptimuman berangka untuk kapasiti penerapan Zirkonium MOF untuk menentukan syarat kendalian terbaik. Kapasiti penjerapan maksimum Zirkonium MOF (46.75 mg/g) didapati pada pH 7, masa sentuhan 70 min, jisim adsorben 10 mg, dan kepekatan pewarna awal 44.99 ppm.

**FORECASTING THE ADSORPTION CAPACITY OF ORGANIC DYE BY USING ZIRCONIUM-BASED METAL-ORGANIC FRAMEWORK (MOF): COMPARISON STUDIES BETWEEN RESPONSE SURFACE AND NEURAL NETWORK MODELS**

## ABSTRACT

The factors affecting the adsorption capacity of Zirconium Metal Organic Framework were analyzed which includes the pH, contact time, amount of adsorbent and initial dye concentration. The experiment was run based on central composite design (CCD) in response surface methodology (RSM). The experimental results were used to investigate the effect of input factors on the adsorption capacity of Zirconium MOF and to develop a model to predict system performance. According to the response surface plot, higher adsorption capacity of Zirconium MOF can be achieved with less adsorbent and a higher dye concentration. RSM was used to create a mathematical model, and the model's performance was evaluated using analysis of variance (ANOVA). Another neural network model was created using MATLAB's neural network toolbox and Mathematica's net operation and predictor function. The adsorption capacity of Zirconium MOF was predicted using a mathematical and neural network model. Due to a shortage of experimental data for neural network training, the mathematical model generated in RSM had a higher accuracy in predicting the output response, with an $R^2$ of 0.97 and an RMSE of 2.87. RSM performed numerical optimization for the adsorption capacity of Zirconium MOF to determine the best operating conditions. The maximum adsorption capacity of Zirconium MOF (46.75 mg/g) was found to be at pH 7, contact time of 70 min, adsorbent amount of 10 mg, and initial dye concentration of 44.99 mg.

# CHAPTER 1

# INTRODUCTION

## 1.1 Background Study

Dye is a colouring material which bonds to the substrate chemically. Generally, dyes are used in aqueous solution. Dyes is known for its existence since 1849 where picric acid once utilized as silk dye (Booth et al., 2000). Dyes commonly assumed to be pigment, but it is not true because dyes chemically bonds to the substrate while pigments cannot. Dyes can be produced by natural and synthetic mean. Natural dyes commonly derived from plants or animals, but synthetic dyes come from petrochemical made by man. Dyes are usually used in textile, leather, paper, and plastic industries (Drumond Chequer et al., 2013). Since colour is a main feature for customers to buy a product, dyeing process caught the attention of textile industry mainly. However, textile dyeing industry produces massive water pollution problems to the environment. The textile industry uses more than 3,600 textile dye which involves more than 8,000 chemicals. Most of these chemicals are hazardous and pollution-causing substance (Kant, 2012).

Wastewater effluent from textile industries can cause lots of environmental problems once released to the environment. The natural water bodies turn into dark color when the wastewater effluent combines with it. This increases the turbidity of water and reduces the sunlight penetration underwater. This means photosynthesis process got hampered and a severe depletion of dissolved oxygen occurs. A depletion in dissolved oxygen can cause major damage to the aquatic creatures living in the eco system (Kant, 2012). Besides that, heavy metals such as chromium which have higher possibilities to accumulate in the food chain. The accumulation of heavy metal in food chain can cause lots of cumulative effect to human being which have high chances to be transferred from generation to generation (Wang et al., 2011).

According to the recent global fast fashion market report, the apparel market is expected to grow up to $38.21 billion in 2023 despite the recent Covid-19 outbreak. The major waste producing activity from textile industry which originates from the wet finishing process use up 200 litre of water per kilogram fibre formed (Globenewswire, 2020). This means as the demand for clothes keep on increasing as predicted in future, more wastewaters will be generated which can harm the nature of our environment.

Wastewater from textile industry can be characterized in terms of pH, colour, suspended solid, chemical oxygen demand (COD), biochemical oxygen demand (BOD), metals, temperature and salts (Yaseen and Scholz, 2019). In Malaysia, the Environmental Quality Act 1974 corresponding to Environmental Quality (Sewage and Industrial Effluent) Regulation 1979 plays an important role in determining the acceptable condition of industrial discharge. The Department of Environment (DOE) has set some ground rules for the discharge of industrial waste into the environment. Table 1.1 shows a simple comparison of parameters between a typical untreated textile wastewater (Ghaly et al., 2014) and the limit of parameters for industrial effluent discharge in Malaysia (DOE, 2015). Parameters such as pH, temperature, BOD, COD, and total suspended solids are cross-checked with Malaysian industrial effluent discharge limit. All the parameters stated has exceeded the standard limit especially BOD, COD and total suspended solid.

Table 1.1 Comparison of parameters between typical untreated textile wastewater and Malaysia industrial effluent discharge limit

| Parameters | Typical untreated textile wastewater (Ghaly et al., 2014) | Malaysia industrial effluent discharge limit (DOE, 2015) |
|---|---|---|
| pH | 6 - 10 | 5.5 - 9 |
| Temperature ($^oC$) | 35 - 45 | 40 |
| BOD (mg/L) | 80 - 6,000 | 50 |
| COD (mg/L) | 150 - 12,000 | 100 |
| Total Suspended Solids (mg/L) | 15 - 8.000 | 100 |

Therefore, treatment of wastewaters from textile industries is a must before discharging it into any natural water bodies. There are some effective ways to treat these wastewaters. It can be divided into 3 sections consist of physical, chemical and biological methods. Table 1.2 shows the processes which falls under each method of treatment (Pang and Abdullah, 2013). Out of all the methods, adsorption seems to be the reasonable, cheapest and simplest method to treat textile wastewater (Kandisa et al., 2016). Hence, this adsorption process had the necessity to be explored in terms of treating textile wastewater.

Table 1.2 The type of wastewater treatment for textile industry

| Treatment method | Treatment type |
|---|---|
| Physical method | a) Adsorption<br>b) Ion exchange<br>c) Membrane filtration |
| Chemical method | a) Fenton oxidation<br>b) Ozonation<br>c) Radiolysis<br>d) Photocatalytic<br>e) Chemical coagulation & flocculation |
| Biological method | a) Aerobic process<br>b) Anaerobic process |

## 1.2 Problem Statement

Adsorption process is one of the reasonable options to treat wastewater effluent from textile industry because it is the cheapest, simplest, and economical way. Therefore, to carry out adsorption process, a suitable adsorbent with high adsorption capacity is needed. Usually, adsorbent such as silica, polymers, activated carbon, activated alumina and zeolite are famous in the market. However, zirconium based MOF are well known for its great flexibility in pore size, pore shape, functionality, thermal resistance and high surface area (Russo et al., 2020).

Even though Zirconium MOF is a good adsorbent, an optimum operating condition to achieve maximum adsorption need to be found out. To get a better understanding and analysation on the adsorption capacity, Response Surface Methodology (RSM) and Artificial Neural Network (ANN) modelling techniques are used. Based on experimental data from Hasanzadeh et al. (2019), a mathematical model needs to be developed to estimate the effect of pH, contact time, adsorbent amount and initial dye concentration for both RSM and ANN model.

Most researcher narrow their research to one modelling technique and how various operating condition affect the adsorption capacity. There are researchers comparing more than one modelling technique to predict the adsorption capacity but very less studies based on this topic are reported. To be more specific there is no study on the comparison of modelling techniques such as RSM and ANN on the DR2B dye adsorption using Zirconium MOF adsorbent. This means no studies are based on comparison of ANN and RSM model and how these models can optimise the adsorption capacity of Zirconium MOF and the interaction between variables.

## 1.3 Objective

This research has three main objectives. Those objectives are:

a) To study the effect of pH, contact time, adsorbent amount, and initial dye concentration on highly water stable microporous Zirconium based metal organic framework (MOF)

b) To compare the adsorption capacity of Zirconium MOF using Response Surface Methodology (RSM) and Artificial Neural Network (ANN).

c) To determine the optimum operating condition for adsorption capacity

## 1.4 Scope of Study

The adsorption capacity of Zirconium MOF was analysed and modelled in this study using statistical tools (RSM and ANN), with experimental data obtained from previous study (Hasanzadeh et al., 2019). The pH, contact time, adsorbent amount, and initial dye concentration were then modelled. The quadratic model was used for modelling, and the feedforward backpropagation network was used for the ANN. In this study, two ANN tools were used which are MATLAB ANN toolbox and Wolfram Mathematica. To compare the predicted value from the mathematical model and the experimental data, a comparison study was conducted. Statistical parameters such as $R^2$ and RMSE values are used to assess the model's fitness. The final stage of the research was optimization. The optimization was carried out to obtain the best operating conditions for maximum adsorption capacity.

## 1.5 Thesis Organization

This thesis is divided into five chapters: introduction, literature review, materials and methodology, results and discussion, and conclusion and recommendations. Chapter 1 described the research's background and problem statements. The objectives of this study were defined, which summarised the study's goals. Chapter 2 documented the literature review to provide related studies to gain a better knowledge of current work. Chapter 3 described the materials and methodology used in this study. The experimental setup for dye adsorption was described, followed by the development and optimization of RSM and ANN models. Chapter 4 covered the results and discussion of the research, which included raw data analysis using RSM, model performance developed by the software, a comparison study, and process optimization. And at last, in Chapter 5, the study's findings were summarised, and some recommendations for future research were made.

## 1.6 Sustainability of dye adsorption

The Sustainable Development Goals (SDGs) are a compilation of 17 global goals. The goal of these SDGs is to function as a guideline for achieving a better future in terms of sustainable development for all. These goals are designed to reduce poverty and other existing deprivations by collaborating on sustainable strategies that would mitigate health problems, education, alleviate inequality, boost economic growth, reduce global warming, and prioritise the preservation of natural resources such as lands, oceans, and forests. The 17 Sustainable Development Goals are goals are (SDG1) No Poverty, (SDG2) Zero Hunger, (SDG3) Good Health and Well-being, (SDG4) Quality Education, (SDG5) Gender Equality, (SDG6) Clean Water and Sanitation, (SDG7) Affordable and Clean Energy, (SDG8) Decent Work and Economic Growth, (SDG9) Industry, Innovation and Infrastructure, (SDG10) Reducing Inequality, (SDG11) Sustainable Cities and Communities, (SDG12) Responsible Consumption and Production, (SDG13) Climate Action, (SDG14) Life Below Water, (SDG15) Life On Land, (SDG16) Peace, Justice and Strong Institutions and (SDG17) Partnership For The Goals.

Dyes which are used in textile industry must undergo wastewater treatment before releasing into the water bodies. One of the important aspect of sustainable goal development for this textile industry will be SDG6 Clean water and sanitation. In order to achieve this sustainable goal, industry must have a water treatment plant and a water recycling system. Another consideration in this regard is the reuse of dye liquor for the next dyeing to reduce the need for freshwater. Finally, industry should have a zero discharge wastewater treatment plant to avoid polluting nearby water bodies. This can be achieved by proper adsorption of the dye using the specified adsorbent. Adsorbent of maximum capacity of adsorption can make this possible to achieve Goal 6.

Water pollution caused by the dyes released to water bodies can be closely related to Goal 14 Life Below Water. This SDG aims to ensure the sustainable use of seas, oceans, and marine resources. The direct release of impurities into inland waters may have a significant impact on the lives of those who live there. As a result, dyes should be adequately treated to protect the biodiversity and lives of aquatic animals, as well as the quality of the water in which they live.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Dyes and its classification

Dyes must have four main properties to be functional. A dye should have colour and have solubility properties in water. A dye should be able to be absorbed or retained by fibre or to be chemically combined with it. Dyes should be able to withstand washing, dry cleaning, and light exposure. In detail, dyes are organic compound with three main groups which contains chromophore, auxochrome and the matrix. Chromophore is made up of a group of atoms which are nitro (─NO2), azo (─N═N─), nitroso (─N═O), thiocarbonyl (─C═S), carbonyl (─C═O), and the alkenes (─C═C─) (Benkhaya et al., 2020).

The dye has a colour due to the presence of chromophore. Chromophore plays as the active site of the dye which can absorb the light energy. The absorption of light energy takes place in the chromophore because of the excited electrons in the molecule. The molecules with excited electron become chromogenic where it has the ability of dyeing through the addition of other groups of atoms called auxochrome (Benkhaya et al., 2020). Auxochrome can fix and modify the colour of the dye. They are also capable of enhancing the dye solubility in water and its affinity towards fibre. The polar auxochrome makes the dye water-soluble and binds the dye to the fabric using oppositely charged groups of fabric structure. Example of auxochrome is amine, hydroxyl, carboxyl and sulfonic radical (Malik and Grohmann, 2011). Dyes can be classified based on their chemical structure or mode of application. So, different classes of dyes with their mode of application and main chromophore are shown in Table 2.1 (Booth et al., 2000).

Table 2.1 Industrially important dyes, their classes, and applications (Booth et al., 2000).

| Class | Properties | Chromophores | Application |
| --- | --- | --- | --- |
| Acid | - Water-soluble<br>- Anionic | - Azo (incl. premetallized chromium complexes)<br>- Anthraquinone<br>- Triarylmethane | Nylon, polyamide, wool, silk, paper, inks, and leather. |
| Azoic | - Water-insoluble azo dye formed on fibre | - Azo | Cotton, rayon, cellulose acetate, and polyester. |
| Basic | - Water-soluble,<br>- Cationic | - Triphenylmethane<br>- Methine<br>- Modified azo<br>- Anthraquinone<br>- Cyanine<br>- Oxazine<br>- Azine<br>- Triarylmethane<br>- Acridine<br>- Xanthene | Paper, polyacrylonitrile modified nylon, polyester and inks. |
| Direct | - Water-soluble,<br>- Anionic | - Azo (incl. copper complexes)<br>- Stilbene<br>- Phthalocyanine<br>- Nitro<br>- Stryl<br>- Anthraquinone | Cotton, rayon, paper, leather, and nylon. |

Table 2.1 Continued

| | | | |
|---|---|---|---|
| Disperse | - Water-insoluble<br>- Non-ionic | - Azo<br>- Anthraquinone<br>- Nitro<br>- Stryl | Polyester, polyamide, acrylic, acetate, and plastics. |
| Mordant | - Water-soluble<br>- Anionic,<br>- Chromium complex formed on fibre | - Azo | Wool |
| Reactive | - Water-soluble<br>- Anionic,<br>- Forms covalent bond with substrate | - Azo (incl. premetallized)<br>- Anthraquinone<br>- Phthalocyanine<br>- Formazan<br>- Oxazine | Cotton, wool, silk, and nylon. |
| Sulphur | - Temporarily solubilized with alkali<br>sulphide | - Sulfur | Cotton and rayon |
| Vat | - Temporarily solubilized as leuco ester<br>with alkaline sodium hydrosulphite | - Anthraquinone<br>- Indigoid | Cotton, rayon, and wool. |
| Pigment | - Water-insoluble<br>- Non-ionic | - Azo<br>- Anthraquinone<br>- Phthalocyanine<br>- Quinacridone | Paints, inks, plastics, and textiles. |

## 2.2 Direct dye and its treatment

Direct dyes are anionic dyes soluble in water and hold groups of sulphonic acid in the structure. They are taken up by the fibres directly. In 1883, Walter Geigy found the direct synthetic dye Direct Yellow 11. Figure 2.1 shows an example chemical structure of a direct red 2B dye. Dyes are supported by hydrogen bond and van der Waal force on the fibre. Even though direct dye's light fastness is good, the wash quality declines a significant amount. It exhibits low wet fastness because the dye particles are small and water soluble. These dyes are preferable to others in terms of cost, improved colour intensity, ease of operation, shorter dye duration, cheaper auxiliary cost, lower overall water usage, and decreased effluent salt level (Shiela, 2020).

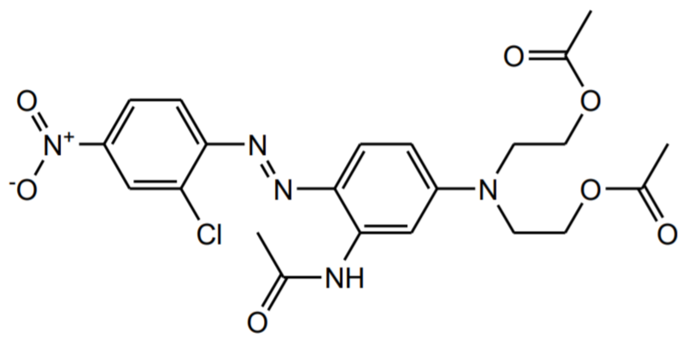


Figure 2.1 Chemical structure of DR2B dye

Substantive dyes are also called direct dyes. These dyes have direct affinity with the fibres and non-ionic powers bind to the fibre. The dyeing method is also cheaper. Dyes are broken up into dye anion and sodium cation as shown in equation 2.1. Direct dyes are mainly used to dye cellulosic fabric, paper, and leather (Benkhaya et al., 2020).

$$D - SO_3Na \rightarrow DSO_3^- + Na^+ \qquad (2.1)$$

The commonly used direct dyes in clothing, leather, paper, and ink industries provides human carcinogenic potential and high environmental toxicity. Direct dyes are proven to be carcinogenic to a variety of mammalian species, including humans. In tests on laboratory animals, two direct dyes, Direct Blue 6, and Direct Black 38, have been reported to be such potent carcinogens that hepatocellular carcinomas and neoplastic liver nodules occurred in rats after only 13 weeks of exposure (Robens et al., 1980). The dyes disintegrate into carcinogenic under anaerobic conditions (Foguel et al., 2015). The aromatic amines and their elimination in bodies of water have the possibility of allergic dermatitis, inflammation of the skin, mutations, and cancer (Lellis et al., 2019). There are also evidence of kidney, liver, and urinary bladder cancers on workers after prolonged exposure to direct dyes (Morikawa et al., 1997).

As the dyes from textile effluent can cause this much of water pollution, removal of direct dyes from wastewater are explored. There are studies on removal of direct dyes using adsorption, coagulation, membrane filtration, foam separation, biodegradation, and DC diaphragm discharge. Adsorption is a method that is typically used in large-scale industry because of it is cheaper, high performance and simpler design. Adsorption treatment happens when the dissolved molecules (dye) are attached to the surface of an adsorbent by attractive forces such as Van der Waals forces. For instance, adsorbents can be zeolites, charcoal, clays, and other waste sources such as coconut shell, rice husk, fly ash, chitosan, and many others (Kandisa et al., 2016)

Coagulation is another method to remove direct dyes from wastewater textile effluent. The coagulant should be able to transform water constituents into forms that can be separated out physically. Natural coagulants such as Moringa seeds has a high molecular weight water-soluble protein in it which are positively charged. The protein creates positive charges that behave like magnets and attract the mainly negatively charged particles when the crushed seeds are introduced to the water. These bound particulates grow under proper agitation to form the flocs, which can be extracted by filtration or left to settle by gravity (Gelebo and Ahmed, 2019).

Membrane filtration is also one of the ways to treat direct dyes from wastewater textile effluents. Nanofiltration appears to be the best for dye treatment as it has been recognized having the properties in between UF and reverse osmosis (RO). Nanofiltration have some significant advantages such as lower osmotic pressure difference, higher permeate flux, higher retention of multivalent salts and molecular weight compounds (>300), relatively low investment and low operation and maintenance costs (Han et al., 2010).

Foam separation is also another alternative for removal of dyes from wastewater treatment. However, foam separation should not be used directly to extract dyes from wastewater, since when aerated, textile wastewater alone cannot create stable foam. Direct dyes are known for its anionic dye property so, they always carry negative charges in aqueous solutions due to the existence of sulfonate group. To generate stable foam and facilitate the formation of dye-surfactant complexes through electrostatic interactions, cetyl trimethyl ammonium bromide (CTAB) was chosen as a surfactant. Complexes of the dye and surfactant adsorb on the surface of bubble that rise from the liquid. Gravity drains out the interstitial water between bubbles in such a way that dye-surfactant complexes are concentrated, and the remaining liquid is clarified. In the foamate, precipitation of the dye-surfactant complexes occurs because the dye-complex concentration in the foamate increases dramatically while their solubility decreases (Lu et al., 2010).

Biodegradation treatment is a well-known method when it comes to treat dyes from textile wastewater effluent. The bacteria/ fungi should be able to degrade a wide variety of recalcitrant compounds, including complex mixtures of pollutants because the enzymes are extracellular with diffusion limitation of substrates into the cell. There are multiple of bacteria with a specific strain to decolourise the dye and purify wastewater effluent (Pakshirajan and Singh, 2010). Biotreatment is the microbial clean-up method that, using microorganisms, contributes to biomineralization and biotransformation of toxic chemicals to less harmful types. In addition to being cost-effective, with minimal sludge production, the most satisfactory feature of biotreatment is being environmentally sustainable (Sharma et al., 2019)

Another unique way to remove direct dyes from textile wastewater effluents are using diaphragm discharge. Wastewater effluent containing direct dyes will be treated using high voltage direct current (DC) provided by diaphragm discharge with an input power of 160W. High voltage applied on electrodes is separated by a dielectric barrier (diaphragm) with a small pinhole where the discharge is ignited in the orifice. Opposite electrode polarity on both sides of the diaphragm attracts the respective charges to separate the dyes from the water solution. Besides separating the dye, this method also provides extra advantage to change the solution properties such as pH and solution conductivity of the solution (Kozáková et al., 2010).

A literature survey was carried out among the treatment method. Table 2.2 shows the literature survey for different type of treatment/method to remove direct dye from textile wastewater effluent.

Table 2.2 Literature survey on the method used to treat direct dyes wastewater.

| Method used | Categories | Optimum operating conditions | | | Remarks | Reference |
|---|---|---|---|---|---|---|
| | | Temperature ($^{o}C$) | Concentration | Time (min) | | |
| Adsorption | Physical method | 21 | 25 mg/L | 30 | - Chitosan–zinc oxide nanoparticle used as adsorbent<br>- Have high potential of the adsorption of dyes<br>- Chitosan is an environmentally benign, non- toxic, antibacterial, and biodegradable. | (Salehi et al., 2010) |
| Adsorption | Physical method | 550 | 4 M | 60 | - $ZnCl_2$ immersed sludge used as adsorbent<br>- 80% of COD value reduction<br>- $ZnCl_2$ immersed sludge has a capacity of removing dye of 71% of granular activated carbon. | (Chen and Zhuang, 2009) |
| Adsorption | Physical method | 21 | 50 mg/L | 1440 | - Soy meal hull is used as adsorbent<br>- Adsorption capacity of the direct dye is 178.57 mg/g.<br>- Maximum desorption ($\geq$ 99.8%) occurs at pH of 10 | (Arami et al., 2006) |
| Coagulation | Physical method | 30 | 50g/L | 45 | - Seeds of Moringa trees used as coagulant.<br>- Efficiency of colour removal was 94.5%.<br>- The colour removal and turbidity removal were 85.8% with method (A) simple extract and 53% with method (B) saline extract respectively at optimized point. | (Gelebo and Ahmed, 2019) |
| Membrane filtration | Physical method | 20 | 1 g/L | 30 | - Copoly (phthalazine biphenyl ether sulfone) is used as nanofiltration membrane.<br>- Rejection of direct dye was 100%.<br>- Membrane flux was higher than 102 L/$m^2$.hr<br>- PPBES UF membrane had good thermal stability and could be used at elevated temperature as high as 95$^{o}$C. | (Han et al., 2010) |

Table 2.2 Continued

| | | | | | | |
|---|---|---|---|---|---|---|
| Foam separation | Physical method | 20 | N/A | N/A | - Cetyl trimethyl ammonium bromide (CTAB) used as surfactant.<br>- Removal efficiency reached 88.9%.<br>- The colour of wastewater decreased from 150 to 17 Pt‑Co<br>- Wastewater COD decreased about 50%, from 728 to 365 mg/L. | (Lu et al., 2010) |
| Biodegradation | Biological method | 37 | 0.1 g/L | 2880 | - *Alcaligenes* sp. strain TEX S6 was used to decolour the wastewater.<br>- The strain removed 86.7% of direct dye from the feed wastewater. | (Sharma et al., 2019) |
| Biodegradation | Biological method | N/A | 50 mg/L | 1440 | - Microbes used are immobilised fungus to produce lignin peroxidase and manganese peroxidase.<br>- Reaction carried out in batch-operated rotating biological contactor reactor.<br>- Decolorization efficiency of 100% can be obtained | (Pakshirajan and Singh, 2010) |
| Biodegradation | Biological method | 30 | 25 mg/L | 2880 | - *Enterococcus gallinarum* strain was used to decolorise the direct dye.<br>- Decolourization was tested and the dye removal was found to be 71–85%. | (Bafana et al., 2008) |
| Diaphragm discharge | Physical method | N/A | 20 mg/L | 40 | - Input power of 160W is used to create high voltage DC.<br>- Decolouration efficiency of 60% was observed.<br>- Solution pH dropped from the initial neutral value to acidic conditions (up to 3) in the anode part and to alkaline conditions (up to 10) in the cathode part. | (Kozáková et al., 2010) |

## 2.3 Statistical Analysis for Adsorption process

### 2.3.1 Response Surface Methodology (RSM)

RSM consists of a set of mathematical and statistical methods focused on the adaptation of empirical models to experimental data collected in relation to experimental design. Linear or square polynomial functions are used to define the method studied for this objective and, accordingly, to explore modelling and displacing of the experimental conditions before optimization. In RSM technique, the basic of theory of the experimental condition need to be studied so that specific optimisation on minimising and maximising of experimental input and response can be done.

The first method will be selection of independent variables of significant effects and delimitation of experimental area made on the system via screening studies. Next will be the selection of the experimental design and the execution of the experiment. Step 3 will be fitting of a polynomial function based on the obtained experimental data. Followed by evaluation of the model fitness. Then, the optimisation of each variable is carried out using numerical or graphical method. Lastly will be getting the optimal values for each variable (Bezerra et al., 2008).

The screening studies mainly targets on those variables giving large or significant impact on the output response. This is reasoning for screening studies because there will be lots of variables affecting the response of the system, but it might just deal a small amount of contribution to the output response. Thus, it is important to carry out screening designs to decide which of the many experimental variables have more critical effects by their interactions to the response output (Bangdiwala, 2018).

For linear function, it is important to choose equation 2.2 as the model as the response will well fit the equation. Equation 2.2 represent the first order degree model where the $k$ is the number of variables, $\beta_o$ is the constant term, $\beta_i$ represents the coefficients of the linear parameters, $x_i$ represents the variables, and $\varepsilon$ is the residual associated to the experiments. For curved functions, second order degree model need to be used. For first order degree model, two-level factorial, Plackett–Burman, and simplex designs are used to see the effect of first-order estimation. Unfortunately, these designs are not enough to estimate second order degree model due to the additional effect in the curvature function. So, a central point in two-level factorial designs can be used to estimate the effects of second order. Examples of second order designs are 3k factorial, central composite, and the Box–Behnken designs. For polynomial model that have additional terms, which means the correlation between different experimental variables. Equation 2.3 shows the second order interaction model, where $\beta_{ij}$ represents the coefficients of the interaction parameters (Uyanık and Güler, 2013).

$$y = \beta_o \sum_{i=1}^{k} \beta_i x_i + \varepsilon \tag{2.2}$$

$$y = \beta_o + \sum_{i=1}^{k} \beta_i x_i + \sum_{1 \le i \le j}^{k} \beta_{ij} x_i x_j + \varepsilon \tag{2.3}$$

Analysis of variance (ANNOVA) is a more reliable way to determine the quality of the fitted model. This is because the mathematical model found after fitting the function to the data can sometimes not satisfactorily describe the experimental domain studied (Bezerra et al., 2008). ANOVA's central concept is to compare the variation due to the treatment (change in the combination of variable levels) with the variation due to random errors inherent in the measurements of the responses produced. This comparison helps a lot in evaluating the importance of the regression used to predict responses considering the sources of experimental variance. The sum of the square is considered the total sum of the square ($SS_{tot}$) for all observation deviations in relation to the media. Total sum of square ($SS_{tot}$) is the summation of sum of square due to regression ($SS_{reg}$), and sum of square due to residual generated by the model ($SS_{res}$) as shown in equation 2.4. Sum of square due to residual generated by the model ($SS_{res}$) is the summation of sum of square due to pure error ($SS_{pe}$) and sum of square due to lack of fit ($SS_{lof}$) as shown in equation 2.5. Media of square (MS) can be calculated by dividing the sum of square with degree of freedom. Equations related to the source of variations for the calculation of $SS_s$ and $MS_s$ are also presented in Table 2.3.

$$SS_{tot} = SS_{reg} + SS_{res} \tag{2.4}$$

$$SS_{res} = SS_{pe} + SS_{lof} \tag{2.5}$$

Table 2.3 Analysis of variance for fitted mathematical model to an experimental data set using multiple regressions (Bezerra et al., 2008).

| Variation Sources | Sum of the square | Degree of freedom | Media of the square |
|---|---|---|---|
| Regression | $SS_{reg} = \sum_i^m \sum_i^{n_i} (\hat{y}_i - \bar{y})^2$ | $p$ - 1 | $MS_{reg} = \frac{SS_{reg}}{p-1}$ |
| Residual | $SS_{res} = \sum_i^m \sum_i^{n_i} (y_{ij} - \hat{y}_i)^2$ | $n$ - $p$ | $MS_{res} = \frac{SS_{res}}{n-p}$ |
| Lack of fit | $SS_{lof} = \sum_i^m \sum_i^{n_i} (\hat{y}_i - \bar{y}_1)^2$ | $m$ - $p$ | $MS_{lof} = \frac{SS_{lof}}{m-p}$ |
| Pure error | $SS_{pe} = \sum_i^m \sum_i^{n_i} (y_{ij} - \bar{y}_1)^2$ | $n$ - $m$ | $MS_{pe} = \frac{SS_{pe}}{n-m}$ |
| Total | $SS_{tot} = \sum_i^m \sum_i^{n_i} (y_{ij} - \bar{y})^2$ | $n$ - $1$ | $MS_{tot} = \frac{SS_{tot}}{n-1}$ |

*$n_i$, number of observations; m, total number of levels in the design; p, number of parameters of model; $\hat{y}_i$, estimated value by the model for the level i; $\bar{y}$, overall media; $y_{ij}$, replicates performed in each individual level; $\bar{y}_i$, media of replicates performed in the same set of experimental conditions.*

### 2.3.2 Artificial Neural Network (ANN)

An ANN is made up of a group of simple processors connected by weighted connections. The processing nodes may be called' neurons' by analogy. The output of each node depends only on the information available locally at the node, either stored internally or accessed through the weighted connections. Each unit receives inputs that propagate its output to yet another node from several other nodes. A single processing element is, by itself, not very efficient. It contains a single numerical value scalar output, which is a simple nonlinear function of its inputs. ANN chosen only if the requirement of certain criteria is met. ANN works the best if there is plenty of data and the problem is poorly understood to derive an approximate model. When one does not have data covering a large portion of operating conditions, or if it is noisy, then it is obvious that neural network technology is not the correct solution (Dongare et al., 2012).

ANN can be composed of different number of neurons. It is possible to bring all the neurons in ANN into one layer or to form two, three or even more layers of neurons. The neurons are drawn as circles and the weights as lines connecting the circles (nodes) in different layers as shown in Figure 2.2. In Figure 2.2, there are four neurons (sometimes called nodes) in a one-layer network, each with four weights. In total, there are 16 weight in one layered network. Both input signals plus the additional input from the bias, which is always equal to one, are acknowledged by every four neurons. Referring to Figure 2.2 (b), a two-layer ANN model with six neutrons (two in the first layer and four in the second layer) which has 20 weights altogether. Perhaps the most used ANN model for chemical applications is the two-layer network. It can safely be assumed that in more than 75% of all ANN applications this type of network architecture is used with different numbers of nodes at each level (Zupan, 1994)

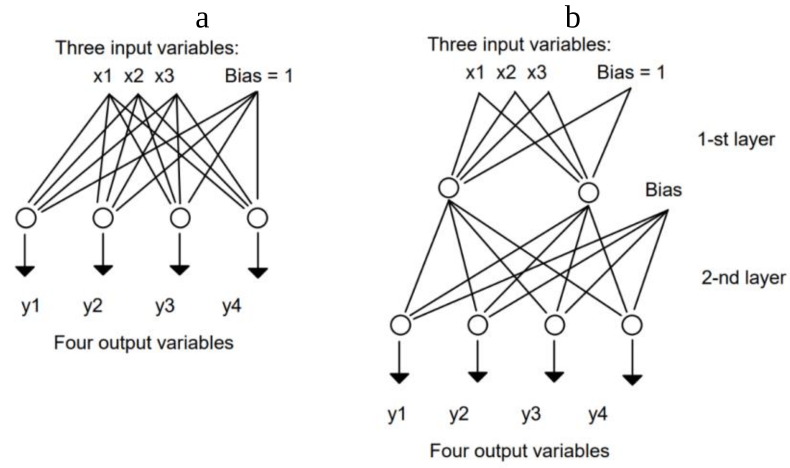


Figure 2.2 (a) One-layer and (b) two-layer ANN (Zupan, 1994).

Training set selection is the first thing to do when it comes to apply the learning method for ANN. Data sets are divided into 3 parts which consist of training det, control set and final test set. Each of these three sets of data should contain approximately one third of the data, with the training set preferring to be the smallest and the test set preferring to be the largest. The most frequently neglected control set is used as a provider of input information to correct the initially selected ANN architecture in cases where the training set performs well. The training process continues until desired output is reached. The error between the output of the network and the actual output is minimized by changing the weights and biases. When all the errors are below the necessary tolerance, the training phase is automatically stopped (Kalogirou, 2000).

In artificial neural networks, learning refers to the method of modifying the weights of connections between the nodes of a specified network (Zupan, 1994). Learning is the mechanism by which a neural network's random-valued parameters (weights and biases) are adapted through a continuous simulation process of the system in which the network is placed. The rate of learning is defined as the rate at which the network adapts. As one of the best learning algorithms, back-propagation (BP) algorithm is a gradient descent algorithm that can be used to learn these multilayer feed forward networks with varying transfer functions. Three transfer functions that are the most utilized transfer function for multilayer networks are shown in Table 2.4.

Table 2.4 The transfer functions that can be used to design multilayer network.

| Transfer function | Equations | Notes |
|---|---|---|
| Log-sigmoid | $y = logsig(x) = \frac{1}{(1+e^{-x})}$ | As the net input of the node goes from negative to positive infinity, the function log-sig generates outputs between 0 and 1. |
| Tan-sigmoid | $y = tansig(x) = \frac{2}{(1+e^{-2x})} - 1$ | For pattern recognition problems, sigmoid output nodes are also used. |
| Purelin | $y = purelin(x) = x$ | Linear (purelin) transfer function is applied to problems of function fitting |

To get a proper training ANN process, there are a total of six steps need to be obeyed. First, the data set must be collected from experiments and the variables that affect the performance of output response must be recognized. Second, the inputs and output variables must be identified. Third, the inputs and outputs data are normalised in the range between 0 and 1; between −1 and 1 or 0.1 and 0.9. Normalisation is important as it can increase the training rate of network. Equation 2.6 shows the normalised value ($x_n$) where $x_{max}$ and the $x_{min}$ are the maximum and minimum value of $x_i$ (Chowdhury and Saha, 2013). Fourthly, the data collected is split randomly into two data sets: the data set for training and the data set for testing. To produce the model output, the training data set is used in the ANN to learn. For evaluating the parameters of the qualified ANN, the testing data set is used. Approximately 70-80% of the randomly chosen data sets are appointed as training data sets and it is possible to apply the remaining data to evaluate the model. In fifth step, the ANN model parameters must be optimized for the construction and training of the ANN model to get accurate results. Lastly, the best ANN model selected based on performance parameters and extraction of results from the optimum qualified model (Ghaedi and Vafaei, 2017).

Values of statistical parameters such as coefficient of determination ($R^2$), root mean squared error (RMSE) and mean squared error (MSE) are evaluated by modifying the model parameters to determine the efficiency of ANN models. The $R^2$ depicts the fit of the predicted output variable approximation curve to the experimental data output variable curve. A model with better performance prediction is demonstrated by greater $R^2$ and lower MSE. The equations related to the statistical parameters are shown in equation 2.7, 2.8 and 2.9.

$$x_n = 0.8\left(\frac{x_i - x_{min}}{x_{\max} - x_{min}}\right) + 0.1 \qquad (2.6)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(|y_{prd,i} - y_{exp,i}|)\text{^}2 \qquad (2.7)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(|y_{prd,i} - y_{exp,i}|)^2}{\sum_{i=1}^{N}(|y_{prd,i} - y_m|)^2} \qquad (2.8)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(|y_{prd,i} - y_{exp,i}|)^2}{N}} \qquad (2.9)$$

### 2.4 Application of RSM and ANN in Adsorption process

#### 2.4.1 Response Surface Methodology (RSM)

RSM has been used to study the adsorption of Methylene Blue dye in aqueous solution using Ho-$CaWO_4$ nanoparticles (Igwegbe et al., 2019). The factors affecting the amount of dye adsorbed including initial dye concentration, initial solution pH, contact time and Ho-$CaWO_4$ nanoparticles dosage. Central composite design (CCD) method is used to evaluate the relationship between factors and response. Quadratic models were used to fit the experimental data. The maximum adsorption capacity predicted for methylene blue dye is 103.09 mg/g at pH of 2.03, contact time of 15.16 min, Ho-$CaWO_4$ nanoparticles dose of 1.91 g/L, and Methylene Blue concentration of 100.65 mg/L. The predicted results were acceptable, as from the ANOVA analysis, the $R^2$ value was close to one for quadratic model of Methylene Blue dye, which is 0.9995.

Another study using activated carbon adsorbent from *Thespesia populnea* pods was developed for the adsorption of Orange G dye from aqueous system (Arulkumar et al., 2011). Three-factor central composite design (CCD) used to maximize the removal of Orange G dye. There are three independent variables to be studied, which were agitation time, initial dye concentration and adsorbent dosage. The model obtained was used to predict the maximum adsorption of Orange G dye. The result showed that the maximum dye adsorbed on activated carbon happens under the optimum condition (concentration 17.6 mg/l; agitation time 4.03 hour; adsorbent dose 0.54g). At these optimum values, the maximum predicted adsorption capacity in terms of percentage of Orange G dye adsorption was 100%. These optimum values were checked experimentally which resulted in 99% of Orange G dye adsorption by activated carbon.