An Expedient Approach Towards Statistical Analysis of Formalin Business Policy Using Design of Experiment in a Petrocomplex Plant in India

Anupam Mukherjee
ARCL Organics Ltd.
Kunal Roy
Hindcon Chemicals Ltd.
Dipak Kumar Jana (dipakjana@gmail.com)
Indian Institute of Engineering Science and Technology,Shibpur https://orcid.org/0000-0003-2297-6576
Pijus Khatua
Haldia Institute of Technology

Research Article

Keywords: Formalin, Process system engineering, Design of experiment, Response surface methodology, Statistical optimization

Posted Date: October 21st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-901581/v1

An expedient approach towards statistical analysis of formalin business policy using design of experiment in a petrocomplex plant in India

Anupam Mukherjee[a,#], Kunal Roy[b,#], Dipak Kumar Jana[c,*], Pijus Kanti Khatua[d,*]

[a]Formaldehyde and Liquid Resins Unit, ARCL Organics Ltd., Kolkata-700141, India

[b]Production & Sales Unit, Hindcon Chemicals Ltd., Vasudha- 62B, Braunfeld Row, Kolkata-700027, India

[c]School of Applied Sciences, Department of Mathematics, Haldia Institute of Technology, Haldia-721657, India

[d]School of Applied Sciences, Department of Chemistry, Haldia Institute of Technology, Haldia-721657, India

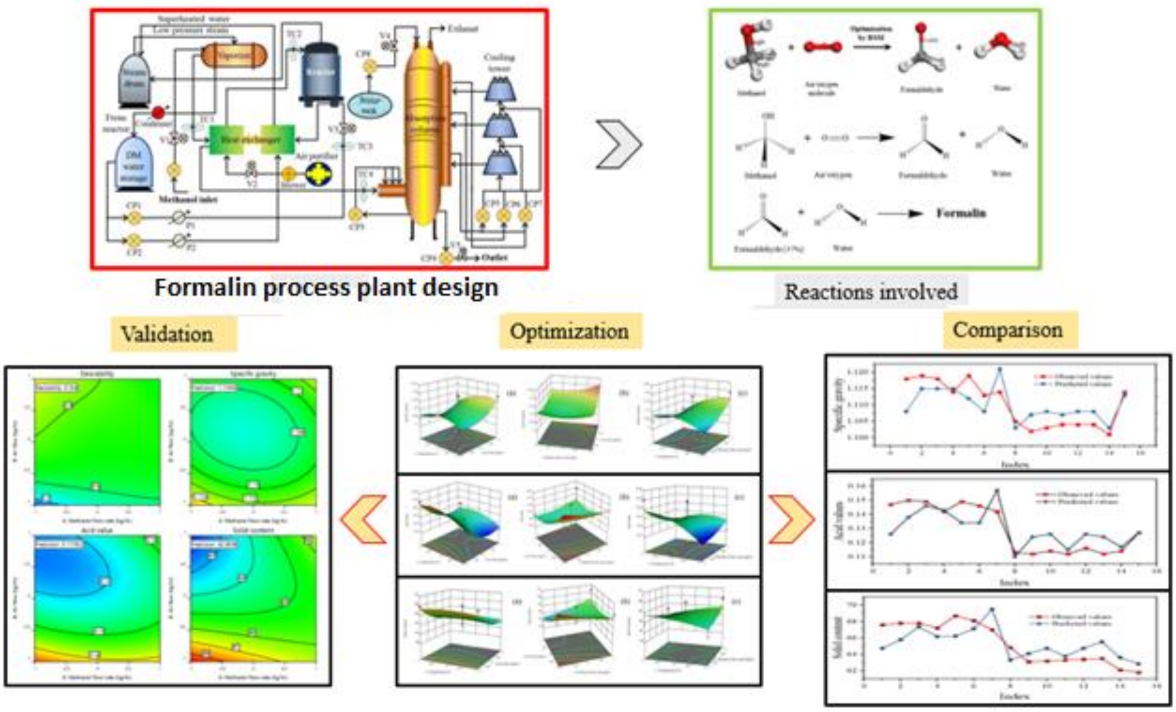
# Contributed Equally

Abstract: Due to civilization, solvent-based paints are abundantly used for painting. Typical solvents include raw Methanol, Ethanol, cellosolve, Amylacetate, and Xylene. After painting, these huge raw solvents are emitted into the atmosphere, which continuously pollutes our environment. Global environment consciousness induced scientists to use aqua-based paints as it never emits harmful material in the atmosphere. Formaldehyde is one of the major components used to produce aqua-based thermosetting resin adhesives, used worldwide in the paints and panel industries. Perceiving the current state of formaldehyde production, development, applications in industrial sectors and demand in the trading industry, a new approach has been envisaged to revitalize the quality of formalin/formaldehyde in petrocomplex plants by the unique design of experiment model based on the collected data. The superiority of formalin depends on some primary constraints such as specific gravity, acid value and solid content of the product. The parameters which control the primary quality measuring constraints are methanol flow-rate, air-supply, and temperature during the reaction process. Based on these three inputs and three output parameters a statistical optimization analysis has been explored with the help of Box-Behnken design by exploring the robustness

of soft computing tool (RSM) from an industrial engineering perspective with the overall desirability of 0.744.

Keywords: Formalin; Process system engineering; Design of experiment; Response surface methodology; Statistical optimization.

Graphical abstract



* Corresponding authors:Dipak Jana (dipakjana@gmail.com, Tel. +91 9474056163), PijusKhatua (pijuskhatua@gmail.com, Tel. +91 8617323409)

Nomenclature

VOC	Volatile organic compound
WHO	World health organization
RSM	Response surface methodology
ANOVA	Analysis of variance
DOF	Design of Experiment
HCHO	Formaldehyde
V_2O_5	Vanadium pentoxide
CH_3OH	Methanol
O_2	Oxygen
CO_2	Carbon dioxide
CO	Carbon mono-oxide
$C_2H_4O_2$	Methyl formate
CCD	Central composite design
BBD	Box-Behnken design
PRESS	Predicted residual error sum of square.
LOF	Lack of fit
R^2	Regression coefficient

1. Introduction

Formaldehyde basically comes under the category of volatile organic compounds (VOCs) having high vapor pressure, low boiling point, as well ashigh reactivity which is well known in the medical sector due to its very good disinfectant and biocidic properties (Bellat, et al., 2015; Musee, et al., 2008). Many of these VOCs are classed as airborne contaminants which may cause skin irritation and cancer etc (Kim, et al., 2011; Wang, et al., 2013; Zhang, et al., 2017) and also sick building syndrome that has become a grave environmental concern nowadays(Jeffrey & Lim, 2003; Shin & Song, 2011). Formaldehyde was first discovered by a Russian chemist, Alexander Butlerov in 1859 and ultimately was identified by German chemist August Hofmann in 1869 and its manufacture was started at the beginning of the twentieth century (Fair, 1980). It is well known that VOCs are emitted not only by industries but also from materials in homes and everyday life(Na, et al., 2018). Formaldehyde is one of the most common volatile organic pollutants, emits from various building materials including furniture and household products apart from the process industries(Wang, et al., 2019). It is now admitted by all the medical authorities that the exposure of animals and humans to formaldehyde can lead to the cancer (IARC, 2006; Liu, et al., 2019; Zou, et al., 2019; Gong, et al., 2018; Salthammer, et al., 2017) and also it causes sneezing and coughing, and leads to acute poisoning, dermal allergies and allergic asthma(Shinohara, et al., 2019); therefore, the WHO (World Health Organization) has recommended a short-term guideline of 0.1 mg/m^3 for a 30 min exposure to prevent sensory irritation in the general population (WHO, 2010; Li, et al., 2016).This chemical is also used in many other industrial applications. For example, formaldehyde is a common precursor for the synthesis of various resins(Liu, et al., 2018; Girods, et al., 2008; Marsal, et al., 2017; Lee, 2012) used in the textile industry, the automobile sector and more extensively the wood industry for the manufacture of wood-composites as plywood or chipboard (Bellat, et al., 2015; Jeong, et al., 2019).

Application of formaldehyde is based on its quality, which is tuned by its properties, like specific gravity, solid content, and acid value. To find out the quality of formalin (37% formaldehyde+water)(Cheung & Lam, 2017) by tuning the quality parameters of formaldehyde, a statistical analysis has been approached to design the parameters using response surface model.The quality of the formalin solution depends on the materials and applied conditions used to synthesize it that is considered to be the input variables and the characteristics of the final product that are considered to be the output variables. The classical method of optimization shows an inability to understand complex interactions between the variables and the response (Hamsaveni, et al., 2001; Soo, et al., 2004). Response surface methodology is one of the most predominantly used statistical tools touted for the optimization of several unpredictable influential interaction parameters simultaneously at a time (Ahmad, et al., 2019; Mirzaei, et al., 2018; Vebber, et al., 2019; Mohammadi, et al., 2019; Jaafari &Yaghmaeian, 2019). The main objective of using this particular analytical methodology is to optimize the system response based on the parameters influencing the process/system (Kim, et al., 2019; Montgomery, 1997; Myers & Montgomery, 2002; Abdulgader, et al., 2019; Gong, et al., 2019). The major advantages of this methodology are: (i) Experimental period minimization instead of a full experimental design at equivalent level(Samarbaf, et al., 2019); (ii) It allows the interaction effects of a factor atvariouslevels; (iii) It facilitates to acquire the surface outline that provides a good prediction for envisioning the interaction (Cochran & Cox, 1992; Jana, et al., 2018). The Design-Expert software package was used to develop the experimental plan for RSM. This software is also used to analyzethe data collected by performing an analysis of variance (ANOVA). During the simulation, if the model looks well fitted, then the three-dimensional surface and contour regions can be plotted for the interpretation of interaction effects while a good model must be significant and simultaneously the lack-of-fit must be insignificant(Jana, et al., 2018).

Producing a qualified product with nominal hazardousness and higher safety is always the pursuit of any chemical process industry by maintaining the quality measuring parameters of the product. Hence, obtaining finer product economization by providing proper parameter tuning is imperative and precious during the processing and designing of the plant (Jia, 2016). In current periods of research, traditional methodologies using deterministic or stochastic techniques(Mukherjee, et al., 2019; Roy, et al., 2019) have been extensively involved to recognize the finest/optimum outcomes from the developed model in several industrial processes (Enriquez, et al., 2011).But during the interpretation of interaction effects between independent and dependent variables, the relations do not follow explicit formulae in the practical field in most of the cases (Zhu, et al., 2015). Therefore, univariate investigations are frequently introduced for the establishment of each parameter's influence on process outcomes. The present study deals with the response surface methodology to utilize it in the evaluation of interaction effects of independent parameters on response parameters by combining experimental design with statistical analysis qualitatively (Khuri & Mukhopadhyay, 2010).

The primary aim of this analysis is to conduct a comprehensive quality study that would lead ultimately to optimum design, in a chemical engineering point of view, of a plant producing formalin with a specified capacity. This research will take into consideration features including the entire process flow of plant set up with basic manufacturing steps, reaction processes and safety precautions due to its high hazardous nature. The main product, formaldehyde (HCHO) is basically an organic compound in the category of aldehydes at its simplest form that can act as a baseline for the synthesis of various polymeric resins like urea-formaldehyde, melamine-formaldehyde, phenol-formaldehyde resins, etc. But the most extensively produced grade is formalin solution i.e. 37 wt. % of formaldehyde in water. Following aspects are taken as the main objectives of the study:

- A new model for formalin production policy in industrial scale has been developed for the first time using Box-Behnken design.
- Effective parameters controlling/tuning the formalin quality have been recognized.
- Conformational analysis of the quality determining parameters has been employed using response surface model.

2. Plant set up and production in brief

The total annual capacity of manufactured formaldehyde in 1998 was approximately 11 billion pounds which were expanded globally in an exponential way reaching a world's production of approx. 32 million metric tons by 2012. Compare to the other industrial graded manufactured materials, formaldehyde is relatively low cost, and high purity and therefore, it is considered the most widely demanded chemical worldwide.

Methanol from storage is pumped to a vaporizer where the liquid gets converted to vapor with the exchange of low-pressure steam from the steam drum. Then the vapor is passed to a heat exchanger where along with fresh air is also passed and with the steam-generating from the reactor end, it is heated up. The much heated air-methanol vapor is then passed into the reactor for reaction with molybdenum as a catalyst, present in the reactor bed, the reaction being an exothermic reaction cooled water is jacketed outside the reactor for controlled reaction to take place. The reaction takes place in the reactor and the gas generated is then passed into another heat exchanger to cool down the temperature. As the temperature gets decreased it is then passed into the absorption column where a counter-current absorption takes place between the formaldehyde gas and the solvent water. The absorption phenomenon takes place and the final product formalin is generated. **Fig.1.** represents the process flow sheet of a formaldehyde production unit.

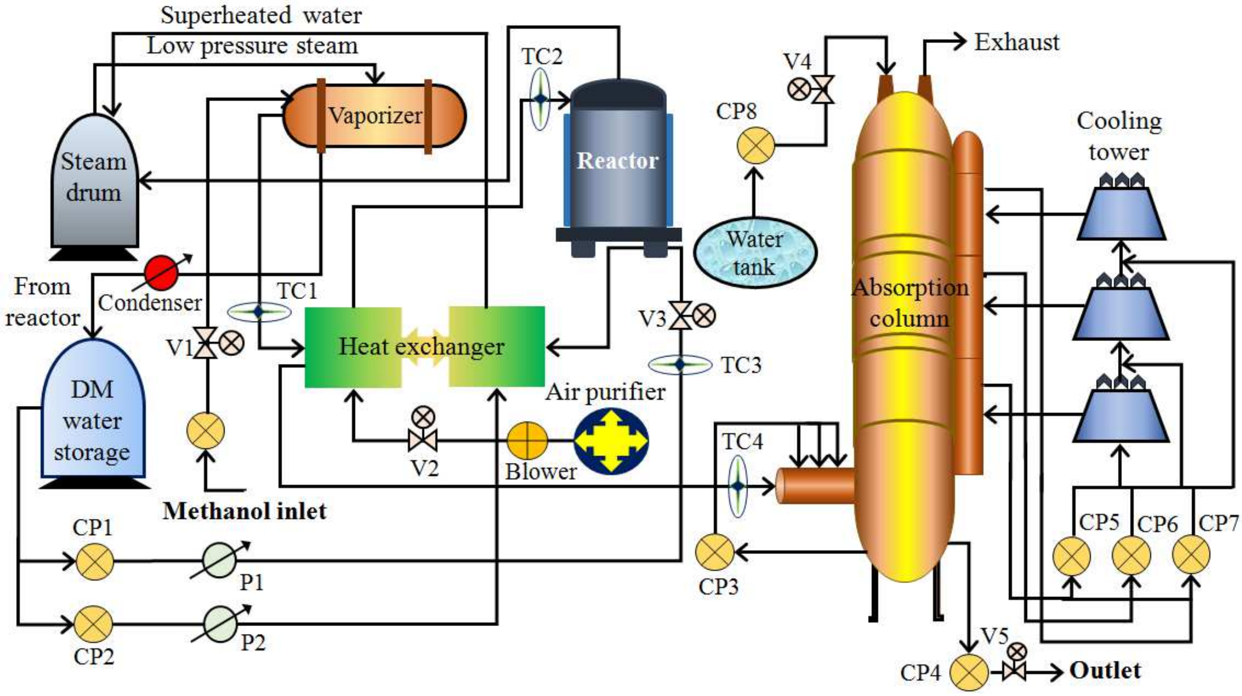


Fig. 1. Typical process flow diagram of formaldehyde production plant indicating following symbols: V1-V5: valve 1 to 5; CP1-CP7: centrifugal pump 1 to 7; P1-P2: pressure gauge 1 to 2; TC1-TC4: temperature controller 1 to 4.

2.1. Catalytic process

Apart from the catalytic vapour-phase oxidation reaction between methanol and air (oxygen), catalytic oxidation (Liu, et al., 2019; Zou, et al., 2019)way of synthesis reaction is also followed in which vanadium pentoxide (V_2O_5) was first introduced as catalyst during the formation of formaldehyde from methanol. Further research on this catalysis materials turned into the development of metal oxide catalysts like ion-molybdenum oxide, silver-based oxide catalyst, etc in the large scale production of formaldehyde with very high conversion yield. During the process, vaporized methanol and air are mixed together entering the reactor. Inside the jacketed heat exchanger reactor, feed is passed through the catalyst introduced tubes. The composition of formaldehyde in theabsorber outlet is controlled by the amount of water addition.

2.2. Reactions dynamics information

Formaldehyde is formed due to the reaction between two main reactants methanol and air/oxygen. In the reactor, this reaction is carried out through the presence of a catalyst which is followed by the oxidation of hydrocarbon i.e. methanol at its vapor phase and its geometrical representation is shown in **Fig. 2.**, while water is produced as a by-product.

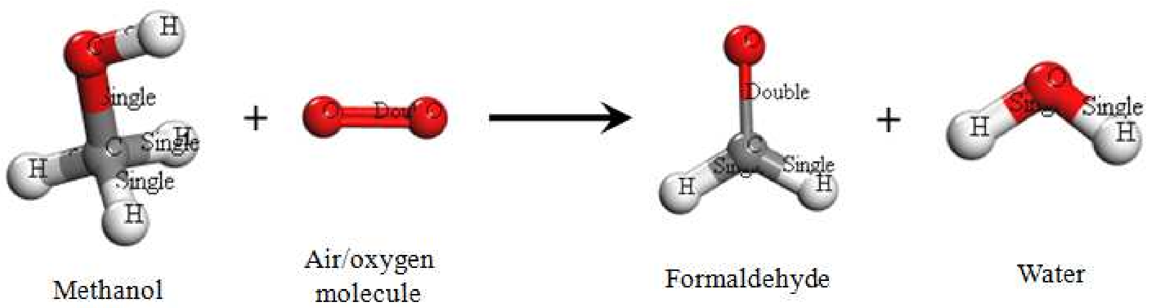


Fig. 2. Representation of reacting molecules during formaldehyde production reaction.

The molecular interaction between methanol and oxygen of the above method (Fig. 2.) has been represented in **Fig. 3**. more precisely through its structural concept and ultimate formalin production from formaldehyde.

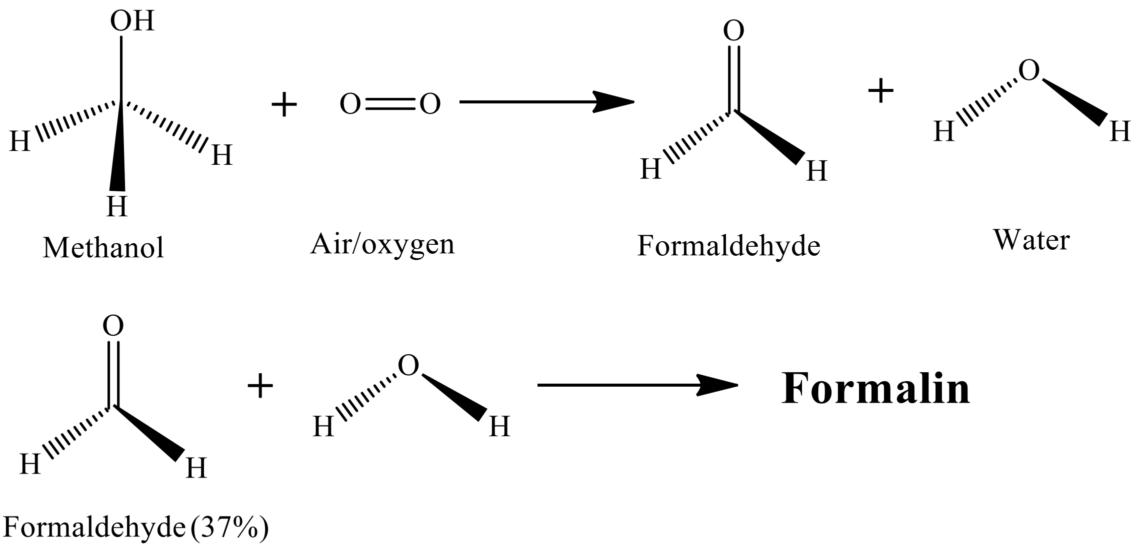


Fig. 3. Molecular interaction between methanol and air/oxygen during formaldehyde production.

The main reactions involved in the formation of formaldehyde from methanol and oxygen and it's major by-products formations are given in following Eqs. (1) and (3) through kinetic expressions:

$$CH_3OH + \frac{1}{2}O_2 \xrightarrow{k_1} HCHO + H_2O \tag{1}$$

Corresponding rate expression is in Eq. (2):

$$-r_1 = k_1 \frac{C_{HCHO} \cdot C_{H_2O}}{C_{CH_3OH} \cdot {}^1/_2 C_{O_2}} \tag{2}$$

$$CH_3OH \xrightarrow{k_2} HCHO + H_2 \uparrow \tag{3}$$

Corresponding rate expression is in Eq. (4):

$$-r_2 = k_2 \frac{C_{HCHO} \cdot C_{H_2}}{C_{CH_3OH}} \tag{4}$$

where k_1 and k_2 are the rate constants and r_1 and r_2 are the reaction rate values C_{HCHO}, C_{CH_3OH}, C_{H_2O}, and C_{O_2} are the concentration of formaldehyde, methanol, water, and oxygen respectively.Excepting the above-mentioned reactions, some undesired by-products are also generated during the reaction including carbon monoxide (CO), carbon dioxide (CO_2), methylformate ($C_2H_4O_2$).

2.3. Safety and environmental precautions

It is very well known that formaldehyde is a highly toxic material which can cause fatal accidents and carcinogenic effect on the entire human body due to ingestion up to 30 ml. It can range from being toxic, carcinogenic as well as allergenic(Hoque, et al., 2018; Hodkovicova, et al., 2019; Payani, et al., 2019).Mainly the occupational hazardousness and side effects depend upon the composition and phase of the material i.e. formaldehyde and these hazards include headache, sore throat, watery eyes, breathing problem, and often cancerous in extreme conditions (Munro, et al., 1999; Salk, et al., 1954; UNC, n.d.).

3. Some design factors used in statistical analysis by response surface methodology

3.1. Experimental domain

Before going to the in-depth discussions on the applicability and feasibility of response surface methodology in analytical optimization methods, it would be pertinent to know about some basic ground factors of this tool. The experimental domain is basically the boundary of the analytical field i.e. upper and lower limits/region of the experimental data variables to be studied.

3.2. Experimental design

It is mainly an explicit set of experimental matrix consists of the various interaction combinations of the studying variables. During carrying out the experiments, several types of available designs can be followed. Some of the popular experimental designs are: (a) central composite design (CCD), (b) Box-Behnken design (BBD), (c) full three-level factorial design, which are discussed later. Essentially, they differ from each other concerning their experimental runs and selectivity of experimental points.

3.3. Independent variables

Independent variables are nothing but those experimental variables thatcan be manipulated or altered with time irrespective of any other factors/parameters. The response or outcome of a system can be affected by a huge number of independent variables at a time during the experiment which is never possible to be included at certain time due to some economic or screening problems and therefore it is necessary to recognize the parameters having major effects on the system response(Wongkaew, et al., 2016).

3.4. Dependent variables

It is sometimes also termed as response variables i.e. output or outcome of the experimental system. It is actually the measured values from experiments depending on which independent variables can be tuned or controlled to optimize the entire process.

4. Response surface methodology and robust design

In the present context, a robust experimental design for the prediction of optimum process parameters has been developed from the plant data. Therefore selection methodologies of the independent and response variables for the model are also explored in this section.

4.1. Parameters selection for the statistical analysis using RSM

Formalin i.e. mainly formaldehyde production process comprises a lot of factors at a time among which the major parameters having the significant effects on the ultimate quality of the produced formalin are methanol flow, air-flow, and temperature of the reactor. Hence, these three major parameters are selected as independent variables of the model. Simultaneously, the grade of the formaldehyde regulates the quality of the formalin i.e. mixture of 37% formaldehyde and water while the quality of this formaldehyde depends on the measurement of the acid value, solid content, and specific gravity. Therefore, these three are selected as response variables.

4.1.1. Methanol flow

Methanol is the most imperative component in the production of formaldehyde and formalin in pilot-scale along withan industrial scale. The principle concept behind the production of formaldehyde is the oxidation of methanol through dehydrogenation. Hence, philosophically the production of formaldehyde is also possible by the oxidation of methane but this method is not industrially viable due to its very low reactivity compare to methanol. During the process, flow-rate of methanol has to be maintained to control its vapor phase reaction and hence control the corresponding acid value and solid content. Since methanol helps to prevent the polymerization of the final product and therefore also inhibits to convert into paraformaldehyde precipitation. In the present study, methanol flow-rate has been varied over the range of 49000 kg/hr to 51000 kg/hr.

4.1.2. Air-flow

Methanol reacts with oxygen supplied through air-flow at vapor phase at the presence of metal oxide catalysts and forms formaldehyde which gets dissolved in water to produce ultimate formalin at a particular concentration. Generally, methanol reacts with air between 300 to 400 ^{o}C in the presence of a catalyst. By controlling the inflow of air within the reactor, the conversion rate can be controlled and also excess air supply can be standardized to achieve the solid content, specific gravity and acid value at their optimum level which will lead to the production of high-quality formalin. Our present process has been carried out with the air-flow of 710 to 1200 kg/hr.

4.1.3. Temperature

In the production of formaldehyde from methanol and air, the temperature needs to be maintained in the range of 100 to 120 ^{o}C to vaporize the entered methanol to reach the activation energy before reacting with oxygen in presence of a catalyst. The temperature range should not be crossed the 120 ^{o}C because methanol is quite volatile liquid and flammable and hence at a higher temperature, it may start to degrade with fire.

4.1.4. Specific gravity

Specific gravity basically defines the ratio of the density of a material to the density of reference material at a standard temperature of 25 ^{o}C. The specific gravity of formaldehyde in formalinsolution typically lies in between 1.100 to 1.150 at standard condition. In our present investigation, it ranges in between 1.100 to 1.200. Therefore it can be considered that the material studied in this process is of proper concentration which can be controlled by the methanol flow, air-flow and providing temperature during the processing.

4.1.5. Acid value

Acid value is a major parameter used for the implication of quality assurance of formalin. This parameter does not consider the pH of the solution rather than quantifying the acidic index i.e. the quantity of contaminating agent of formic acid in the formalin solution which should be below 1.000 as per the plant guidelines. Since in formalin solution, formaldehyde always starts to break down into formic acid which is not desired quality of the product, it would have no chance of getting neutral or basic nature and therefore the contamination level by formic acid in the solution is determined. Here, the acid value of the product solution ranges in between 0.110 to 0.115 which is highly acceptable for further applications.

4.1.6. Solid content

Solid content in any suspension epitomizes the proportion of non-volatile material contained in the suspension. It is basically the constituents left after the volatile solvent which serves as a carrier or vehicle for the solid content, has been vaporized. Typically, the solid content appears due to rapid polymerization of formalin solution after production to form paraformaldehyde precipitation which is controlled by the addition of methanol. Characteristically, the solid content value should be varied in between 60 to 75 based on the grade of the solution as per the area of application.

4.2. Formulation of experimental designs

4.2.1. Full three-level factorial design

A full three-level factorial design is quite less applicable design matrix due to its requirement of a higher number of experimental numbers. In general, it is used where factor number higher than two. Experimental number is calculated using the following equation(Duan, et al., 2013):

$$N = 3^k \tag{5}$$

where, N is the number of experiments, and k is the number of factors used during the experiment. Due to the requirement of a higher number of experimental numbers, its modeling efficiency gets reduced in quadratic functions (Bezerra, et al., 2008). Therefore, other design matrices such as the Box-Behnken design, central composite design, which require a much smaller number of experimental points to represent the model, are often used in common practice economically (Morris, 2000).

4.2.2. Central composite design

It was first developed by Box and Wilson in 1951 (Box & Wilson, 1951). Central composite design mainly comprises of the following: (a) a complete factorial design, (b) a surplus design, often termed as star design which contains the experimental points at a particular distance from the centre, and (c) a centre point. It includes following characteristics:

- Required experimental number is calculated by following equation:

  $$N = (2^k + 2k + 1) + C_o \qquad (6)$$

  where,N is experimental number, k is factor number, and C_o is number of central point.

- Coded α-values are calculated using following equation based on number of factors:

  $$\alpha = 2^{(k-p)/4} \qquad (7)$$

  For example, if factor number is three; then corresponding α-value will be 1.68.

- Here all the factors are analysed in five levels including –α, -1, 0, +1, +α.

4.2.3. Statistical analysis through Box-Behnken design

Box-Behnken design was first introduced by George E. P. Box and Donald Behnken as an experimental designer.The basic fundamentals concepts, advantages and short-comings were first illustrated by Ferreir et al. (2007)(Ferreira, et al., 2007; Jana, et al., 2018). The major advantage of this Box-Behnken design matrix is its designing proficiency of all factors simultaneously at their maximum or minimum levels with involving any combinational approach and therefore it can dodge the experimental limitation under extreme conditions (Kazemia, et al., 2010; Kazemzadeh, et al., 2019). Box-Behnken design includes the following features (Ferreira, et al., 2007):

- Here, the experimental points are generally sited at hypersphere equidistant from the central point of the model.
- Required experimental number is calculated by following equation(Latchubugata, et al., 2018; Chen, et al., 2013):

$$N = (k^2 + 2k + 1) + C_o \quad (8)$$

where, N is number experiments, k is factor numbers, and C_o is the number central points.

- The response surface design is developed with all combinations of the factors at three levels(high, +1, 0, and low, -1levels) (Mujtaba, et al., 2014).

Though Box-Behnken design has been applied in the field of analytical chemistry or at industrial scale at very low quantity compare to the central composite design, but in present context this Box-Behnken design(Daraei, et al., 2019) has been considered as our basic platform of experimental modelling at its large scale application to investigate its robustness as a statistical optimization tool due to its simplicity of exploration in experimental model designing in analytical chemistry research.

Table 1. Coded variables formulation in Box-Behnken design.

Model variables	Symbols	Coded levels		
		Low	Medium	High
Methanol flow (kg/hr)	X_1	49000 (-1)	50500 (0)	52000 (+1)
Air flow (kg/hr)	X_2	700 (-1)	950 (0)	1200 (+1)
Temperature (°C)	X_3	110 (-1)	115 (0)	120 (+1)

4.3. Codification of the levels of parameters

Codification of the parameter levels are of important concern because of its compatibility at different orders of magnitude excepting higher influencing the evaluation of the design which considers the transformation of each studied real assessment into the coordinates with dimensionless systems having the proportionality with experimental space (see Table 1.). To transform the real assessment into coded value by codification as per the determinate design, following equation is followed:

$$x_i = \left(\frac{z_i - z_i^0}{\Delta z_i}\right)\delta_d \qquad (9)$$

where, Δz_i is the gap between the real value at the central point and at superior or inferior level, δ_d is primary coded limit value within the matrix for each parameter, and z_i^0 is real value at the centre of the design.

4.4. Model adequacy check and analysis of fitted design

After procuring the model based on the experimental points of the preferred design, fitting of the model with corresponding mathematical correlation is obligatory to exemplify the response performance of fitted model at its studied level. It has been discovered that the mathematical model fitted on the function often may not be a silver bullet to emphasize the required domain and therefore it would be exorbitantly reliable to illustrate the model adequacy in terms of the application of analysis of variance which imperatively relies on comparing treated model variation with the variation for random errors of generated responses (Bezerra, et al., 2008). In this model adequacy verification an important deterministic parameter is sum of the squares error or residuals $(\mathrm{S}_{\mathrm{Error}})$which has been expressed in following Eq (10):

$$S_{Error} = \sum_{i=1}^{n}\left(y_{pred_i} - y_{obs_i}\right)^2 \qquad (10)$$

where, y_{pred_i} indicates the predicted value by the model at point i and y_{obs_i} indicates the observed value by the model at point i.e, S_{Total} (sum of square of the total) is another important parameter in the analysis of fitting accuracy which is expressed in Eq (11):

$$S_{Total} = \sum_{i=1}^{n}\left(y_{obs_i}\right)^2 \qquad (11)$$

PRESS or prediction residual error sum of the squares quantify the fitted design by measuring how the model fits each point in the design which is shown in Eq. (12):

$$PRESS = \sum_{i=1}^{n}\left(y_{pred_i} - y_{obs_i}\right)^2 \qquad (12)$$

Usually, a wide variation between the residual error and PRESS residual designates a pinch point where the model gets well-fitted.

The model adequacy can also be verified by lack of fit test (LOF) because it can analyse the model failure percentage by the interpretation of data points in experimental domain through the comparison between residual error and pure error which should be insignificant. If the model does not fit the data properly, then it will be significant (Nair, et al., 2014). LOF (F_{LOF}) test can be expressed by Eq. (13):

$$F_{LOF} = \frac{S_{LOF}/(f-p)}{S_{Pure\ error}/(n-p)} \qquad (13)$$

where, F_{LOF} indicates the sum of squares for LOF, $S_{Pure\ error}$ indicates sum of squares for pure error, f indicates no. of specifically different parameter interactions, n indicates experimental number in the set and p indicates factor number in the set.

Apart from the above mentioned parameters, one more essential parameter is coefficient of determination or regression (R^2) which would be discussed later in result and discussion section.

4.5. Optimization of multiple responses

During the optimization analysis of multiple responses, the considered parameters should meet the desirable criteria to reach optimum conditions. **Fig. 4.** exemplifies the schematic algorithm of process flow for design of experiment approach through response surface methodology.In this type of optimization analysis, generally multicriteria method of analysis is followed which is known as desirability function approach. Each desirability function made up of conversion of each response varying over the range of 0 to 1.

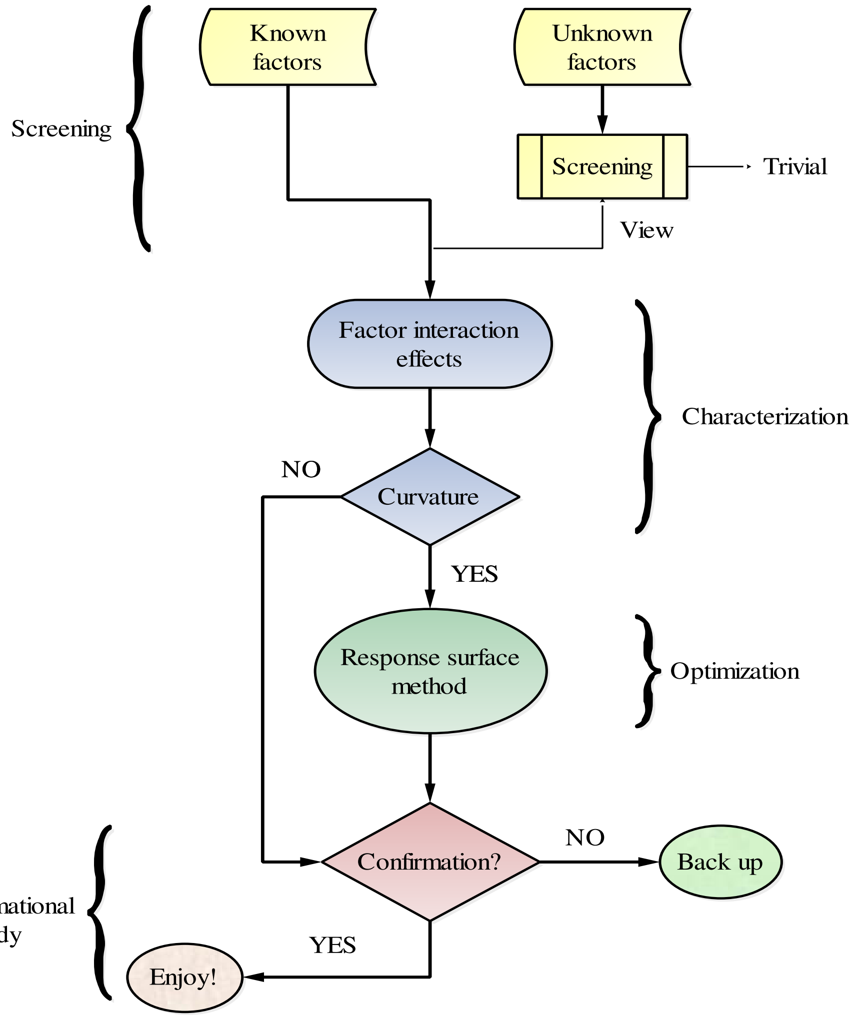


Fig. 4. Philosophical concept of statistical model algorithm through response surface methodology.

When the response comes to its target, desirability function becomes 1 and vice-versa (Montgomery, 1997; Myers & Montgomery, 2002). The one-sided desirability can be evaluated from the Eq. (14):

$$D_n = \int_1^0 \left[\frac{O_n - O_{n-min}}{O_{n-max} - O_{n-min}}\right]^r; \qquad \text{if,} O_n \leq O_{n-min}$$

$$\text{if,} O_{n-min} < O_n < O_{n-max} \qquad (14)$$

$$\text{if, } O_n \geq O_{n-max}$$

where, O_n is the response value, O_{n-min} & O_{n-max} are the minimum and maximum acceptable values of response n, and r is the positive constant (weight) used to find the desirability. If the overall desirability is D', then it follows Eq. (15):

$$D' = (D_1 \times D_2 \times D_3 \times \cdot\cdot\cdot\cdot\cdot)^{1/k} \quad (15)$$

where, k is the number of responses and $0 \leq D \leq 1$.

5. Results and discussions

After exemplification of the model, its validation has to be patterned and this model validation has been studied through overall efficiency of the model by the evaluation of deterministic coefficient or regression coefficient R^2 which is expressed in Eq. (17-18). Here a non-linear polynomial functionhas been used to design the experimental system is as follows (Verma & Sarkar, 2017) in Eq. (16):

$$y = \beta_0 + \sum_{i=1}^{3} \beta_i x_i + \sum_{i=1}^{3} \beta_{ii} x_i^2 + \sum_{i=1}^{3} \beta_{ij} x_i x_j + \epsilon \quad (16)$$

where, y is the predicted response, x_i indicates the independent variables, β_0 indicates the constant term, β_i indicates the linear coefficient, β_{ii} indicates the squared coefficient, β_{ij} indicates the interaction coefficient and ϵ is the error term.

The predictive efficiency of the proposed BBD designwas assessed by the test data in the trained data and comparing the predicted and observed values.In addition, the statistical parameters such as the deterministiccoefficient(R^2) in Eq. (17-18), adjusted regression coefficient $\left(R_{adj}^2\right)$ in Eq. (19-20), predicted regression coefficient $\left(R_{pred}^2\right)$ in Eq. (21) were used to compare predicted and measured values of flexible modulus:

$$R^2 = 1 - \frac{S_{Error}}{S_{Total}} \quad (17)$$

$$\text{Or,} \quad R^2 = \frac{\sum_{i=1}^{n}\left(y_{pred_i} - y_{obs_i}\right)^2}{\sum_{i=1}^{n}\left(y_{obs_i}\right)^2} \quad (18)$$

$$R^2_{adj} = 1 - \frac{S_{Error}/(n-p)}{S_{Total}/(n-1)} \quad (19)$$

Or, $$R^2_{adj} = 1 - \left(\frac{n-p}{n-1}\right)(1 - R^2) \quad (20)$$

$$R^2_{pred} = 1 - \frac{PRESS}{S_{Total}} \quad (21)$$

$$F_{R^2} = \frac{Mean\ square\ of\ the\ developed\ model}{Mean\ square\ of\ the\ residual\ error} \quad (22)$$

where, all the terms for Eq. (17-21) are already discussed in the section 4.4. In Eq. (22) F_{R^2} indicates the significance of the regression coefficient.

Table 2. Experimental design matrix with respect to the Box-Behnken design factors.

Air (Kg/Hr)	Methanol (Kg/hr)	Temperature (°C)	Specific Gravity (Kg/m^3)		Acid Value (mg KOH/g)		Solid Content (mg/L)	
			Observed	Predicted	Observed	Predicted	Observed	Predicted
0	0	0	1.118	1.108	0.147	0.126	67.6	64.733
-1	0	-1	1.119	1.115	0.15	0.138	67.8	65.775
0	-1	-1	1.118	1.115	0.149	0.146	67.8	67.338
+1	-1	0	1.114	1.115	0.142	0.143	67.2	66.163
+1	+1	0	1.119	1.112	0.149	0.134	68.7	66.213
0	-1	+1	1.113	1.108	0.146	0.134	68.1	67.113
-1	-1	0	1.114	1.121	0.142	0.157	67	69.488
-1	0	+1	1.105	1.103	0.113	0.110	64.8	63.300
0	+1	-1	1.102	1.107	0.112	0.124	63.1	64.088
0	0	0	1.103	1.108	0.114	0.126	63.2	64.733
0	+1	+1	1.104	1.107	0.112	0.115	63.3	63.763
0	0	0	1.104	1.108	0.116	0.126	63.4	64.733
+1	0	+1	1.104	1.108	0.112	0.124	63.5	65.525
+1	0	-1	1.101	1.103	0.114	0.117	62.1	63.600
-1	+1	0	1.114	1.113	0.127	0.127	61.8	62.838

5.1. Conformational study of regression model

In the field of statistical analysis, regression analysis characteristically provides the information about the correlation efficiency between the operating and response variables. The practicability of the simulation and its effects on the interaction/operating parameters sustained by the collected data has been well-fashioned in the proposed model. **Table 2** exemplifies the experimental design matrix and corresponding predicted values of response variables given by the process simulation. After configuring the quadratic model with linear regression fit, the independent variables have been found with good adequacy and the corresponding consequential regression model has been optimized which has been represented in the given **Table 3** and the model equations of response variables are also described with their significant terms in **Table 3**.

Table 3. Regression summary of predicted responses specific gravity, acid value, and solid content.

Responses	Specific gravity (Kg/m^3)	Acid value (mg KOH/g)	Solid content (mg/L)
SD	0.0027	0.0050	0.8398
Mean	1.02	0.1298	64.85
C.V. %	0.2434	3.8400	1.29
R^2	0.9864	0.9540	0.8845
R^2_{adj}	0.9232	0.9077	0.8429
R^2_{pred}	0.8406	0.9235	0.8063
Adeq. Prec.	18.4756	20.5429	24.8145
PRESS	0.0009	0.0026	111.28

Model equations with significant terms:
Sp. Gr. = $1.11 + 0.0049X_2 + 0.0063X_3 - 0.0028X_2^2$.
Acid value = $0.127 - 0.0132X_1 + 0.0099X_2 - 0.0067X_2^2$.
Solid content = $63.64 - 3.9X_1 + 1.05X_2$.

R^2_{adj} = adjusted R^2; R^2 = regression coefficient; R^2_{pred} = predicted R^2; C.V = coefficient of variation; SD = standard deviation; Adeq. Prec. = adequate precision; PRESS = predicted residual error sum of square.

Statistically, the accuracy of any model is typically evaluated by the regression coefficient or determination of coefficient (R^2) (Singh, et al., 2010; Latchubugata, et al., 2018). From Table 3. it can be demonstrated that the regression model has been well fitted in the optimization of specific gravity and acid value due to > 95% of their regression coefficient (0.98 and 0.95 respectively) and in case of solid content prediction, the model cannot well explain the model adequacy and goodness of fit to low value of R^2 (< 90%). Usually, p-value < 0.05 defines the significant terms of a model at its 95% confidence level within probability limit. As per the system of statistical analysis, typically p-values are considered as the smallest values as level of significance which can be determined using p-value = 1 – level of significant. It designates the mode terms are as follows: (i) highly substantial (i.e. p < 0.01); (ii) substantial (0.01 < p < 0.025); (iii) ordinary (0.025 < p < 0.05); and (iv) weak (0.05 < p < 0.1) and the values higher than the 0.1 are considered as insignificant model terms. Herein, present context the p-values < 0.05 i.e. significant, have mainly come into the linear terms of the model. In case of sp. gr. p-values are significant only for model term, X_2, and X_3 while in case of acid value and solid content, significant p-terms are include for model, X_1, X_2, X_2^2 and model, X_1, X_2 respectively. The F-values of 29.13, 44.92, and 57.03 for sp. gr., acid value and solid content respectivelyand corresponding standardized errors of 0.0006, 0.0012, and 0.1945 are also indicating that the model terms are highly significant as well as accurately fitted as shown in Table 4.

Table 4. ANOVA analysis of response surface functions based on Box-Behnken design model.

Responses	Source	Sum of squares	Mean square	F-value	p-value
Specific gravity	Model	0.0019	0.0002	29.13	<0.0001
-	X_1	0.0000	0.0000	1.52	0.2216
-	X_2	0.0002	0.0002	31.17	<0.0001
-	X_3	0.0001	0.0001	8.80	0.0039
-	X_1X_2	8.71E-06	8.71E-06	1.19	0.278
-	X_1X_3	4.33E-06	4.33E-06	0.5929	0.4434
-	X_2X_3	2.89E-06	2.89E-06	0.3929	0.5313
-	X_1^2	0.0000	0.0000	1.67	0.1996

-	X_2^2	0.0000	0.0000	5.21	0.025
-	X_3^2	0.0000	0.0000	1.72	0.1936
-	Residual	0.0006	7.30E-06	-	-
-	Cor total	0.0025	-	-	-
Acid value	Model	0.0100	0.0011	44.92	<0.0001
-	X_1	0.0006	0.0006	24.98	<0.0001
-	X_2	0.0009	0.0009	36.66	<0.0001
-	X_3	0.0000	0.0000	0.4726	0.4937
-	X_1X_2	2.80E-07	2.80E-07	0.0113	0.9156
-	X_1X_3	4.93E-07	4.93E-07	0.0199	0.8882
-	X_2X_3	8.86E-08	8.86E-08	0.0036	0.9525
-	X_1^2	0.0000	0.0000	1.81	0.1815
-	X_2^2	0.0002	0.0002	8.55	0.0044
-	X_3^2	0.0000	0.0000	0.4104	0.5235
-	Residual	0.0021	0.0000	-	-
-	Cor total	0.0121	-	-	-
Solid content	Model	361.97	40.2200	57.03	<0.0001
-	X_1	53.64	53.6400	76.06	<0.0001
-	X_2	10.29	10.2900	14.59	0.0003
-	X_3	2.21	2.21	3.13	0.0803
-	X_1X_2	1.32	1.32	1.87	0.1749
-	X_1X_3	1.18	1.18	1.68	0.1986
-	X_2X_3	0.4923	0.4923	0.698	0.4058
-	X_1^2	1.69	1.69	2.4	0.1249
-	X_2^2	0.0105	0.0105	0.0149	0.903
-	X_3^2	0.0175	0.0175	0.0248	0.8753
-	Residual	59.95	0.7053	-	-
-	Cor total	421.92	-	-	-
Lack of fit = 3 Pure error = 2	Which can be recommended as valid LOF test.				

5.2. The influence of interacting parameters on response variables: a statistical approach

To get the confirmation about the model adequacy and goodness of fit in the present investigation, the specific gravity, acid value, and solid content offer a valuable estimation of the proposed real environment through the model regression equations described in Table 3., having very good and significant LOF of 3 (Table 4.), the simulated results are compared to the observed collected data and the results are illustrated in **Fig. 5(a-c).**

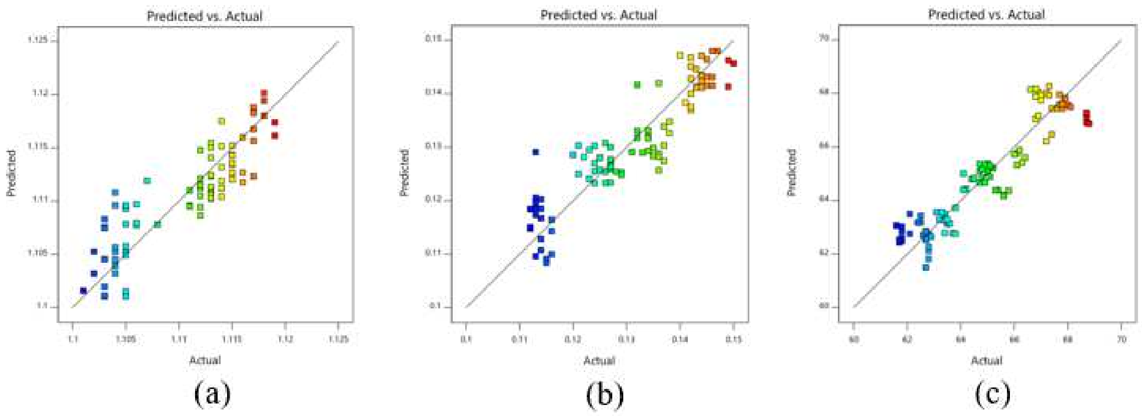


Fig. 5. Predicted values vs actual data for the design responses (a) Specific gravity, (b) acid value, and (c) solid content.

To envisage the influential effects of interaction parameters on the independent operating variables, a detail explanation has been explored through the sensitivity analysis of 3-dimensional surface plots considering the each response variables as a function of two independent variables which are displayed in **Figs. 6, 8 and 10**.

5.2.1. Specific gravity

Figs. 6 (a-c). represent the 3D surface plots of an important response specific gravity as a function of two essential variables air flow and methanol flow-rate, air flow and temperature, and also temperature and methanol flow-rate respectively while third parameter is considered to be zero value or at the centre location of the design. As displayed in the plots, the specific gravity has been increased gradually from 1.012 to 1.117 i.e. we can say from lower region to higher region with correspondingly increase in air flow and temperature slightly within the range of -1 to +1 and therefore, the p-value is significant at this place which is $p < 0.05$ (Table 4.). The reason behind this might be the quadratic effect of air flow with a p-value of < 0.0001. Again at the lower values of methanol flow-rate, specific gravity becomes significant comparatively due to its lower p-value of 0.221which is not significant.

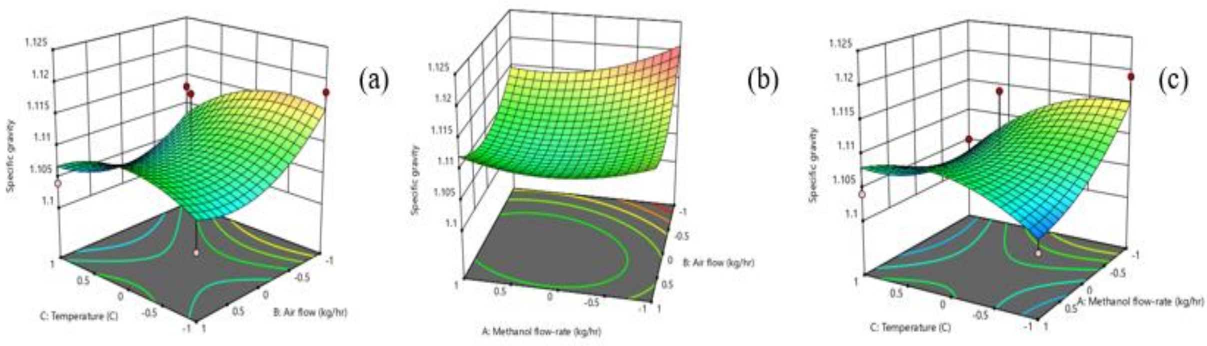


Fig. 6. Response surface characterization model displaying specific gravity as a function of two parameters while third one remains at its centre position: (a) temperature (X_3) and air flow (X_2); (b) methanol flow-rate (X_1) and air flow (X_2); and (c) temperature (X_3) and methanol flow-rate (X_1).

Figs. 7, 9 and 11. represent the comparative briefs on the respective responses of the system specific gravity, acid value and solid content based on their corresponding predicted and observed data collected from the design of experiment.

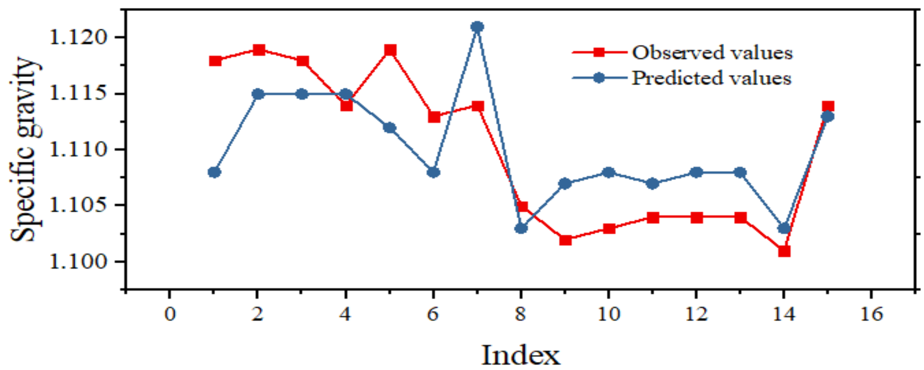


Fig. 7. Comparison between predicted and observed values of specific gravity based on the simulation outcomes.

5.2.2. Acid value

Herein, **Figs. 8(a-c).** represent the 3D surface plots of an important response acid value as a function of two essential variables air flow and methanol flow-rate, air flow and temperature, and also temperature and methanol flow-rate respectively while third parameter is considered to be zero value or at the centre location of the design as similar as specific gravity. As shown in the figure, the acid value has been increased gradually from 0.110 to 0.115 due to some significance terms present in the model development. The acid value has been increased with the respected decrement in temperature and increment in air flow because the p-valuesare significant at X_1 and X_2 i.e. $p < 0.05$ (Table 4.). This is may be because of non-linear quadratic effect of temperature in the system.

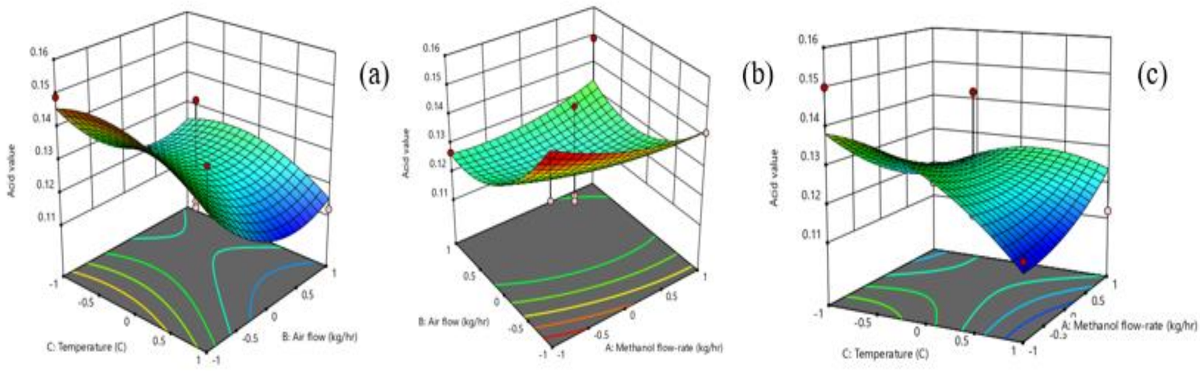


Fig. 8. Response surface characterization model displaying acid value as a function of two parameters while third one remains at its centre position: (a) temperature (X_3) and air flow (X_2); (b) methanol flow-rate (X_1) and air flow (X_2); and (c) temperature (X_3) and methanol flow-rate (X_1).

Fig. 9. displays the predicted vs observed data plot for the comparison with simulation results of acid value.

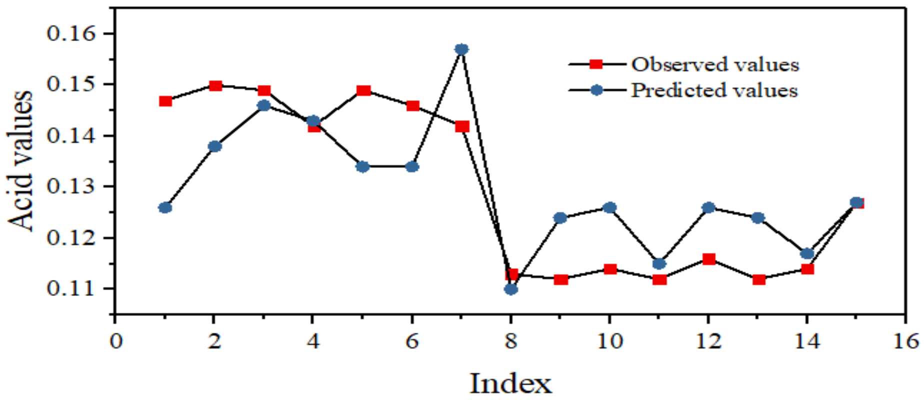


Fig. 9. Comparison between predicted and observed values of acid values based on the simulation outcomes.

5.2.3. Solid content

Here, **Figs. 10 (a-c)**. have explained the influential effects of methanol flow-rate, air flow and temperature due to interaction with the corresponding response solid content though the response surface analysis and the respective optimization analysis has been studied via numerical optimization and statistical desirability studies which has been discussed later section in depth. Figs. 10(a), 10(b), and 10(c) have well designed the response solid content as a function of temperature and air flow, air flow and methanol flow rate, temperature and methanol flow rate respectively. As displayed in the figure, response values are increasing slightly at the lower values of methanol flow-rate from 62 to 65 in average while the methanol flow lies near -1 level. From ANOVA result also it can be well explained that the p-values are highly significant in model terms, X_1, and X_2 with $p << 0.05$ which indicates very high goodness of fit also validating the result as well-fitted.

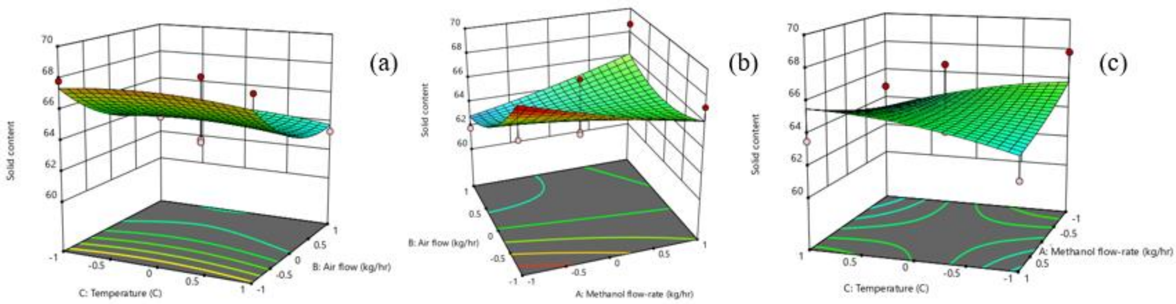


Fig. 10. Response surface characterization model displaying solid content as a function of two parameters while third one remains at its centre position: (a) temperature (X_3) and air flow (X_2); (b) methanol flow-rate (X_1) and air flow (X_2); and (c) temperature (X_3) and methanol flow-rate (X_1).

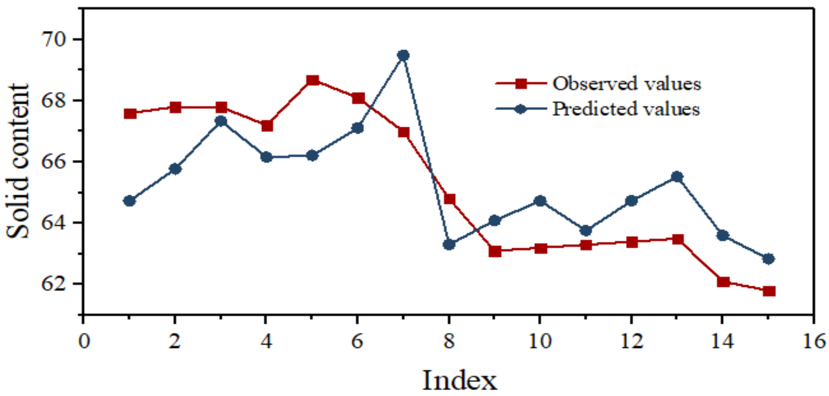


Fig. 11. Comparison between predicted and observed values of solid content based on the simulation outcomes.

Fig. 11. illustrates the comparative study between predicted and observed values of solid content as per the design of experiment analysis with average similarity.

5.3. Parameter optimization

In multiple-response optimization study, the optimum conditions should meet simultaneously the desired criteria of proposed model which can be recognised visually through the superimposed plot of response surface 3D plot and desirability contour plot. By considering the acid value and solid content simultaneously reach the minimum and specific gravity reach the targeted value under the influence of optimum state, the corresponding optimized outcomes have been achieved via desirability study and their analysis through response surface model.

Table 5. Simulation results of the responses after optimization via DOF.

Response variables	Observed values	Predicted values	Error prediction (%)
Specific gravity	1.113	1.1122	0.07
Acid value	0.127	0.126	0.8
Solid content	62.591	62.601	-0.016
Error (%)= [(Observed-Predicted)/Predicted] × 100			

From the optimization study through statistical numerical optimization it has been shown that the standard deviation between the observed and predicted values of the response functions are not high (see Table 5.) and therefore it can be considered that the model is fit to the optimal analysis rationally. **Fig. 12(a) and 12(b)**. represent the graphical desirability study confirming the optimum predicted point on the model which shows the viable response values in the range factors zone with the overall desirability of 0.744. The optimum points are identified by taking into account the model responses specific gravity, acid value, and solid content. The high quality of the product i.e. formalin has been achieved at a low flow-rate of methanol (towards -1 level), higher flow rate of air (towards +1 level) and medium range of applied temperature (see Fig. 12(a) and (b).) simultaneously. The corresponding optimum response are specific gravity of 1.110, acid value of 0.117, and solid content of 62.063.

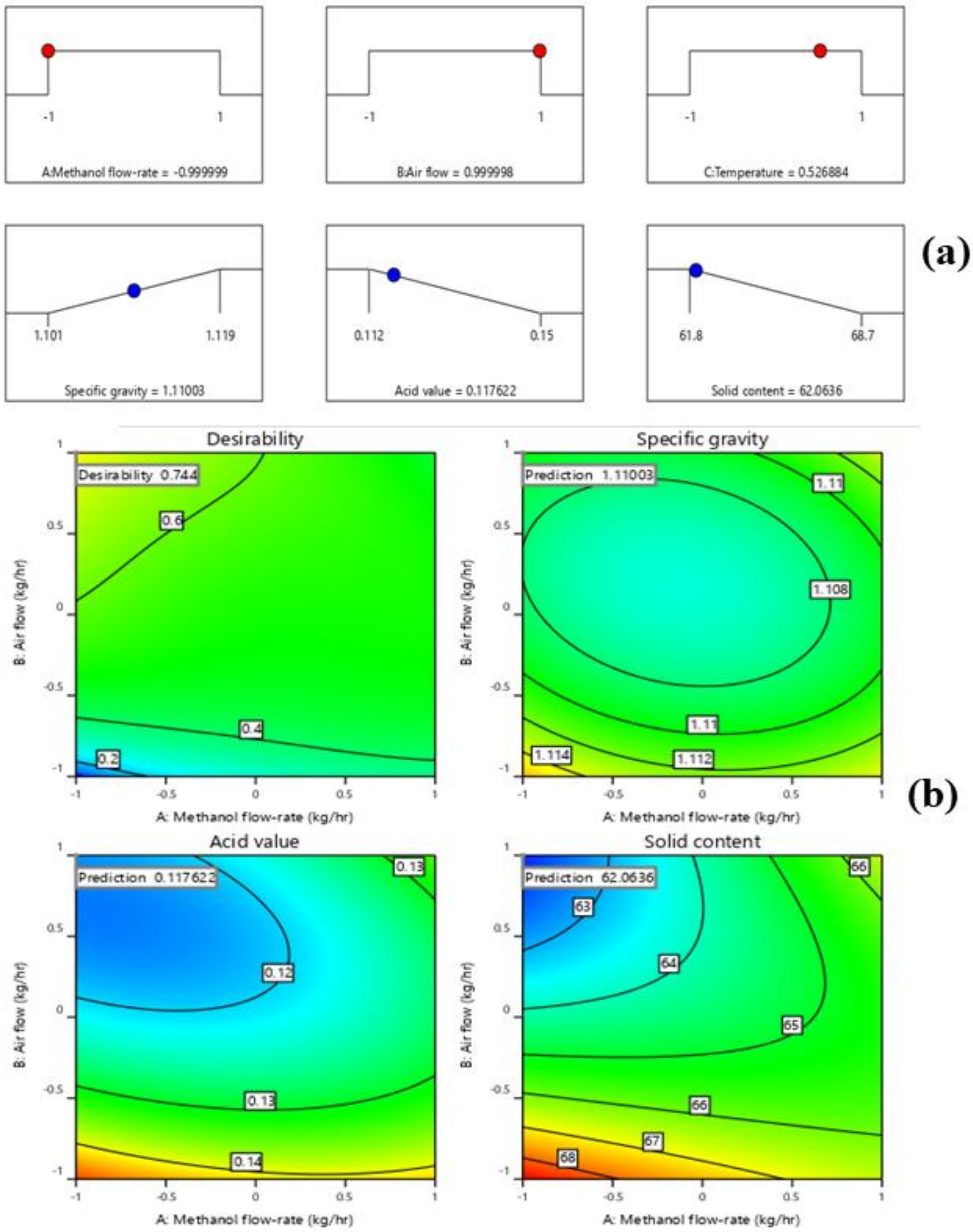


Fig. 12. (a) Optimization desirability study of the process parameters and **(b)** Optimum simulation predictions via design of experiment based on BBD.

6. Conclusion and future prospects

Ever since proceeding decades, the development in formaldehyde processing has been involved significantly in research and development in the novel and innovative solution for chemical as well as polymer industry to revitalize the production quality and productrecovery through the control of quality assurance and hazardousness. Although the formaldehyde production in various chemical and pharmaceutical process industries has been popularized exorbitantly in this twentieth century due to its over-reached application in the field of biomedical, polymer, plywood, construction etc. Therefore, research investigations are also spurring towards the opening of newer and safer way of formaldehyde production and quality control at optimized level. Present context has well-represented the philosophical concept of response surface methodology as a unique approach for the design of experiment in formalin production plant. It emphasizes the robustness of design experiment tool for the optimization of produced formalin by controlling the quality maintaining parameters of formaldehyde through the help of statistical analysis in a well-organised fashion. By the implementation of proposed design, practically in processing plant, the project engineers will not have to face any intricacy to render the parameter setting as well as quality assertion determination of formalin like a trial and error process.The current prediction method and simulation of optimized parameters for the quality assertion of the product will not only assist the manufacturing sectors but also will be helpful for the end users to identify the product superiority in marketing industry.

Declarations:

Conflict of interest: Authors have declared no conflict of interest.

Ethical approval: The submitted work is original and has not been published elsewhere in any form or language. Furthermore, it has not been submitted to more than one journal for simultaneous consideration

Consent to participants: Not Applicable

Consent for publications: Not Applicable

Funding: Not Applicable

Availability of Data & Materials: Data that support the findings of the study are kept confidential and will be available upon request.

CRediT authorship contribution statement:

Anupam Mukherjee: Conceptualization, Methodology, Investigation, Writing-Original Draft, Visualization.

Kunal Roy: Data Curation, Formal Analysis, Software, Writing-Original Draft

Dipak Jana: Validation, Investigation, Supervision

Pijus Khatua: Resource, Formal Analysis, Supervision

References

Abdulgader, M. et al., 2019. Process analysis and optimization of single stage flexible fibre biofilm reactor treating milk processing industrial wastewater using response surface methodology (RSM). Chemical Engineering Research and Design, Volume 149, pp. 169-181.

Ahmad, F. et al., 2019. Optimization for silver remediation from aqueous solution by novelbacterial isolates using response surface methodology: Recovery andcharacterization of biogenic AgNPs. Journal of Hazardous Materials, Volume 380, p. 120906.

Bellat, J. -P. et al., 2015. Capture of formaldehyde by adsorption on nanoporous materials. Journal of Hazardous Materials, Volume 300, pp. 711-717.

Bezerra, M. A. et al., 2008. Response surface methodology (RSM) as a tool for optimization in analytical chemistry. Talanta, Volume 76, pp. 965-977.

Box, G. E. P. & Wilson, K. B., 1951. On the Experimental Attainment of Optimum Conditions. Journal of the Royal Statistical Society B, Volume 13, pp. 1-45.

Chen, Y. -D., Chen, W. -Q., Huang, B. & Huang, M. -J., 2013. Process optimization of K2C2O4-activated carbon from kenaf core using Box-Behnken design. Chemical Engineering Research and Design, Volume 91, pp. 1783-1789.

Cheung, A. F. P. & Lam, D. S. C., 2017. Formalin: a formaldehyde analogue. Canadian Journal of Ophthalmology, Volume 52, p. 229.

Cochran, W. G. & Cox, G. M., 1992. Experimental Designs. New York: John Wiley & Sons.

Daraei, P., Zereshki, S. & Shokri, A., 2019. Application of nontoxic green emulsion liquid membrane prepared by sunflower oil for water decolorization: Process optimization by response surface methodology. Journal of Industrial and Engineering Chemistry, Volume 77, pp. 215-222.

Duan, X. et al., 2013. Regenration of spent catalyst from vinyl acetate synthesis as porous carbon: Process optimization using RSM. Chemical Engineering Research and Design.

Enriquez, A. H., Tanco, M. & Kim, J. K., 2011. Simulation-based process design and integration for the sustainable retrofit of chemical processes. Industrial & Engineering Chemistry Research, 50(21), pp. 12067-12079.

Fair, J. R., 1980. Advanced Process Engineering. New York: American Institute of Chemical Engineers.

Ferreira, S. L. C. et al., 2007. Box-Behnken design: An alternative for the optimization of analytical methods. Analytica Chimica Acta, Volume 597, pp. 179-186.

Girods, P. et al., 2008. Thermal removal of nitrogen species from wood waste containing urea formaldehyde and melamine formaldehyde resins. Journal of Hazardous Materials, Volume 159, pp. 210-221.

Gong, H. et al., 2019. Structural optimization of a demulsification and dewatering device coupled with swirl centrifugal and high-voltage fields by response surface methodology combined with numerical simulation. Chemical Engineering Research and Design, Volume 148, pp. 361-374.

Gong, Y., Zhou, X., Ma, X. & Chen, J., 2018. Sustainable removal of formaldehyde using controllable water hyacinth. Journal of Cleaner Production, Volume 181, pp. 1-7.

Hamsaveni, D. R., Prapulla, S. G. & Divakar, S., 2001. Response surface methodological approach for the synthesis of isobutyl isobutyrate. Process Biochemistry, Volume 36, pp. 1103-1109.

Hodkovicova, N. et al., 2019. The effects of a therapeutic formalin bath on selected immunological and oxidative stress parameters in common carp (Cyprinus carpio). Science of the Total Environment, Volume 653, pp. 1120-1127.

Hoque, M. S. et al., 2018. Evaluation of artificially contaminated fish with formaldehyde under laboratory conditions and exposure assessment in freshwater fish in Southern Bangladesh. Chemosphere, Volume 195, p. 702.

IARC, 2006. Monographs on the Evaluation of Carcinogenic Risks to Humans 88, s.l.: s.n.

Jaafari, J. & Yaghmaeian, K., 2019. Optimization of heavy metal biosorption onto freshwater algae (Chlorella coloniales) using response surface methodology (RSM). Chemosphere, Volume 217, pp. 447-455.

Jana, D. K., Roy, K. & Dey, S., 2018. Comparative assessment on lead removal using micellar-enhanced ultrafiltration (MEUF) based on type-2 fuzzy logic and response surface methodology. Separation and Purification Technology, Volume 207, pp. 28-41.

Jeffrey, P. O. & Lim, S. F., 2003. Airborne concentrations of bacteria in a hospital environment in Singapore. Water, Air and Soil Pollution, Volume 144, pp. 333-341.

Jeong, B., Park, B. -D. & Causin, V., 2019. Influence of synthesis method and melamine content of urea-melamine-formaldehyde resins to their features in cohesion, interphase, and adhesion performance. Journal of Industrial and Engineering Chemistry.

Jia, B. -J., 2016. RSM-Based Analysis and optimization Approach for Chemicl Processes. Journal of Advance Research in Applied Science, Volume 3, pp. 70-81.

Kazemia, S. Y., Hamidi, A. S., Asanjaranib, N. & Zolgharnein, J., 2010. Optimization of a new polymeric chromium (III) membrane electrode based on methyl violet by using experimental design. Talanta, Volume 81, pp. 1681-1687.

Kazemzadeh, A., Ein-Mozaffari, F. & Lohi, A., 2019. Mixing of highly concentrated slurries of large particles: Applications of electrical resistance tomography (ERT) and response surface methodology (RSM). Chemical Engineering Research and Design, Volume 143, pp. 226-240.

Khuri, A. I. & Mukhopadhyay, S., 2010. Response surface methodology. Computational Statistics, Volume 2, pp. 128-149.

Kim, D. I. et al., 2011. Comparison of removal ability of indoor formaldehyde over different materials functionalized with various amine groups. Journal of Industrial and Engineering Chemistry, Volume 17, pp. 1-5.

Kim, Z. et al., 2019. Develoopment of NOx removal process for LNG evaporation system: Comparattive assessment between reponse surface methodology (RSM) and artificial neural network (ANN). Journal of Industrial and Engineering Chemistry, Volume 74, pp. 136-147.

Latchubugata, C. S. et al., 2018. Kinetics and optimization studies using Response Surface Methodology in biodiesel production using heterogeneous catalyst. Chemical Engineering Research and Design, Volume 135, pp. 129-139.

Lee, C. -T., 2012. Preparation of porous, nano-sized urea-formaldehyde polymer powder. Journal of Industrial and Engineering Chemistry, Volume 18, pp. 919-925.

Liu, L. et al., 2019. Formaldehyde adsorption in carbon nanopores- New insights from molecular simulation. Chemical Engineering Journal, Volume 370, pp. 866-874.

Liu, M., Wang, Y., Wu, Y. & Wan, H., 2018. Hydrolysis and recycling of urea formaldehyde resin residues. Journal of Hazardous Materials, Volume 355, pp. 96-103.

Li, Y. et al., 2016. Effects of formaldehyde exposure on anxiety-like and depression-like behaviour, cognition, central levels of glucocorticoid receptor and tyrosine hydroxylase in mice. Chemosphere, Volume 144, pp. 2004-2012.

Marsal, A. et al., 2017. Reduction of the formaldehyde content in leathers treated with formaldehyde resins by means of plant polyphenols. Journal of Cleaner Production, Volume 148, pp. 518-526.

Mirzaei, A., Yerushalmai, L., Chen, Z. & Haghighat, F., 2018. Photocatalytic degradation of sulfamethoxazole by hierarchial magnetic ZnO@g-C3N4: RSM optimization, kinetic study, reaction pathway and toxicity evaluation. Journal of Hazardous Materials, Volume 359, pp. 516-526.

Mohammadi, F. et al., 2019. Modeling and optimizing pyrene removal from the soil by phytoremediation using response surface methodology, artificial neural networks, and genetic algorithm. Chemosphere.

Montgomery, D. C., 1997. Design and Analysis of Experiment. New York: John Wiley & Sons.

Morris, M. D., 2000. A Class of Three-Level Experimental Designs for Response Surface Modeling. Technometrics, Volume 42, pp. 111-121.

Mujtaba, A., Ali, M. & Kohli, K., 2014. Statistical optimization and characterization of pH-independent extended-release drug delivery of cefpodoxime proxetil using Box-Behnken design. Chemical Engineering Research and Design, Volume 92, pp. 156-165.

Mukherjee, A., Roy, K. & Jana, D. K., 2019. Qualitative model optimization of almond (Terminalia catappa) oil using soxhlet extraction in type-2 fuzzy environment. Soft Computing, Volume 23, pp. 1-11.

Munro, K., Lanser, J. & Flower, R., 1999. A comparative study of methods to validate formaldehyde decontamination of biological safety cabinets. Applied and Environmental Microbiology, Volume 65, pp. 873-876.

Musee, N., Lorenzen, L. & Aldrich, C., 2008. New methodology for hazardous waste classification using fuzzy set theory Part I. Knowledge asquisition. Journal of Hazardous Materials, Volume 154, pp. 1040-1051.

Myers, R. H. & Montgomery, D. C., 2002. Response Surface Methodology: Process and Product Optimization Using Designed Experiments. New York: John Wiley & Sons.

Na, C.-J.et al., 2018. High-performance materials for effective sorptive removal of formaldehyde in air. Journal of Hazardous Materials, Volume 366, pp. 452-465.

Nair, A. T., Makwana, A. R. & Ahammed, M. M., 2014. The use of response surface methodology for modelling and analysis of water and wastewater treatment processes: a review. Water Science & Technology, Volume 69, pp. 464-478.

Payani, S., Mamatha, C., Chandraprakash, C. & Bhaskar, M., 2019. Protective role of (Bronco- T) against formaldehyde induced antioxidant, oxidative and histopathological changes in lung of male Wister rats. Toxicology Reports, Volume 6, pp. 718-726.

Roy, K., Mukherjee, A. & Jana, D. K., 2019. Prediction of maximum oil-yield from almond seed in a chemical industry: A novel type-2 fuzzy logic approach. South African Journal of Chemical Engineering, Volume 29, pp. 1-9.

Salk, J. E. et al., 1954. Formaldehyde treatment and safety testing of experimental poliomyelitis vaccines. American Journal of Public Health, Volume 44.

Salthammer, T., Giesen, R. & Schripp, T., 2017. A permeation-controlled formaldehyde reference source for application in environmental test chambers. Chemosphere, Volume 184, pp. 900-906.

Samarbaf, S., Birgani, Y. T., Yazdani, M. & Babaei, A. A., 2019. A comparative removal of two dyes from aqueous solution using modified oak waste residues: Process optimization using response surface methodology. Journal of Industrial and Engineering Chemistry, Volume 73, pp. 67-77.

Shinohara, N., Okazaki, Y., Mizukoshi, A. & Wakamatsu, S., 2019. Exposure to benzene, toluene, ethylbenzene, xylene, formaldehyde, and acetaldehyde in and around gas stations in Japan. Chemosphere, Volume 222, pp. 923-931.

Shin, S. K. & Song, J. H., 2011. Modeling and simulations of the removal of formaldehyde using silver nano-particles attached to granular activated carbon. Journal opf Hazardous Materials, Volume 194, pp. 385-392.

Singh, R. et al., 2010. Biosorption optimization of lead (II), cadmium (II) and copper (II) using response surface methodology and applicability in isotherms and thermodynamics modelling. Journal of Hazardous Materials, Volume 174, pp. 623-634.

Soo, E. L. et al., 2004. Response surface methodological study on lipase-catalyzed synthesis of amino acid surfactants. Process Biochemistry, Volume 39, pp. 1511-1518.

UNC, n.d. Finance and Operations: Environment, Health and Safety. [Online] Available at: https://ehs.unc.edu/chemical/formaldehyde/

Vebber, M. C., Crespo, J. d. S. & Giovanela, M., 2019. Self-assembled thin films of PAA/PAH/TiO2 for the photooxidation of ibuprofen. Part I: Optimization of photoactivity using design of experiments and surface response methodology. Chemical Engineering Journal, Volume 360, pp. 1447-1458.

Verma, S. P. & Sarkar, B., 2017. Rhamnolipid based micellar-enhanced ultrafiltration for simultaneous removal of Cd (II) and phenolic compound from wastewater. Chemical Engineering Journal, Volume 319, pp. 131-142.

Wang, H. et al., 2013. Characterization and assessment of volatile organic compounds (VOCs) emissions from typical industries. Chinese Science Bulletin, Volume 58, pp. 724-730.

Wang, M., Xu, S., Li, S. & Zhu, R., 2019. Isolation of formaldehyde-degrading bacteria and the evaluation of the degradation characteristics. Journal of Industrial and Engineering Chemistry, Volume 75, pp. 224-229.

WHO, 2010. WHO Guidelines for Indoor Air Quality: Selected Pollutants, WHO Regional Office of Europe: World Health Organization.

Wongkaew, K. et al., 2016. Mass transfer resistance and response surface methodology for separation of platinum (IV) across hollow fiber supported liquid membrane. Journal of Industrial and Engineering Chemistry, Volume 42, pp. 23-45.

Zhang, X. et al., 2017. Adsorption of VOCs onto engineered carbon materials: A review. Journal of Hazardous Materials, Volume 338, pp. 102-123.

Zhu, L., Deng, J., Yang, Y. & al, e., 2015. Optimization of energy utilization for natural gas dehydration facilities based on RSM. Chemical Engineering (China), Volume 43, pp. 40-43.

Zou, N. et al., 2019. Electrothermal regeneration by Joule heat effect of carbon cloth based MnO2 catalyst for long-term formaldehyde removal. Chemical Engineering Journal, Volume 357, pp. 1-10.