Design of experiment in preliminary optimization study of bioethanol production from *Chlorella pyrenoidosa*

H B Aditiya*, Y Mulyaningsih, H C Theofany, A D Pramiesta and Y Mulyaningsih

Department of Mechanical Engineering, Sampoerna University, Jl. Raya Pasar Minggu, Kav 16, Jakarta 12780, Indonesia

*aditiya.harjon@sampoernauniversity.ac.id

**Abstract.** As an alternative, sustainable energy form to replace fossil gasoline, the interest in producing bioethanol is still advancing. This is critical especially when third generation bioethanol production is focused, in which the fast-growing green algae is exploited as the feedstock. While the bioethanol production route is principally consistent regardless of the feedstock type, implementing optimization in bioethanol production is advantageous to the production as it presents optimum conditions based on the set variables. This study reports the design of experiment in preliminary analysis of optimization study in bioethanol production from *Chlorella pyrenoidosa*. Four factors (feedstock loading; NaOH concentration; hydrolysis time; hydrolysis temperature) are selected to determine the main response of reducing sugar as the main desired component for fermentation stage. Using response surface methodology, the design of experiment consists of 29 experimental runs with linear model as the suggested model.

1. **Introduction**

The endeavour for alternative energy that supports sustainable and environmentally friendly still sparks many researchers and scientists to date. Primarily this is because the energy generation is sourced predominantly by means fossil fuel resources and the derivatives. It has been a popular fact that fossil fuel has caused prolonged issues to the environment. Greenhouse gases, for example, have directly affected the global environment, which resulted from the carbon-based exhaust (CO and CO2) as the side products of fossil fuel combustion. Subsequent impacts rise from here. In the massive scale of exhaust gas emission, these gases are clouded in the atmosphere blocking the earth’s heat that should be released. This leads to global warming, and many disturbances that lead to the unbalance ecosystem follow.

Biofuel is often raised as one promising alternative fuel that is sourced from renewable resources, of which usually take form as agricultural products and vegetal oils. Bioethanol, the biofuel substitute to fossil gasoline, is a specific alternative fuel suitable for spark-ignited internal combustion engines. Current commercial bioethanol production is by means of sugar fermentation from corn and sugar cane. Although this route is available in large scale, there is always a debate on food versus fuel, since these feedstocks are edible sugars [1, 2]. From here, bioethanol production can be routed to the non-edible sources as feedstock, that is called second generation bioethanol production [3]. Types of feedstock in second generation plays important role in the success of this alternative [2, 4]. For instance, the biomass produced from the rich agricultural and forestry biodiversity in Indonesia would highly beneficial if it
were taken more seriously nationwide [5]. Another level in bioethanol production is through sugar-to-ethanol conversion from carbohydrates of algae. This third generation bioethanol production wins over second generation on the land requirement in feedstock cultivation, as algae does not require large area to farm and it can be also done in an open water pond or dedicated algae photo bioreactor [6]. Algae harvesting cycle and growth rate are also notably faster in comparison with agricultural crops. Definitely, using algae as main source in biofuel production does not compete with food stability. Experimental studies involving algae in bioethanol production has been conducted by refs. [7-9].

To pair with the great third generation potential, optimization is a highly valuable method to achieve the most optimum production. Through optimization study the set variables are analyzed throughout the production route, in which the most favourable conditions can be predicted and suggested. In optimization study, response surface methodology (RSM) is favorable compared to the conventional optimization method. RSM utilizes the interactions between the set factors toward the observed response, which results in less trials than that of conventional optimization matrix. This ensues a greater efficiency in both time and cost. Box-Behnken design in RSM allows formulation of design of experiment (DoE) based on three-level incomplete factorial design. In a study with experiment-intensive, this method is highly beneficial to avoid the unaffordability to perform the high-extreme data point, which is costly [10]. Combination of RSM and Box-Behnken is fairly prevalent to be implemented in an optimization study of bioethanol production [11].

Coming from there, this paper reports the preliminary stage in optimization of bioethanol production from green algae *Chlorella pyrenoidosa*. The study aims to formulate the design of experiment (DoE) suitable for the optimization study. Response surface methodology with Box-Behnken experimental design is used to formulate the DoE. Four factors are set as experimental variables, with reducing sugar as the main observed response. As a preliminary stage, the outcome from this paper is critical for the subsequent stages in the optimization study.

2. Methods

2.1. *Chlorella pyrenoidosa* green algae collection

*C. pyrenoidosa* was obtained from a domestic supplier in Jakarta, Indonesia. The obtained feedstock was in already in dried, powder form and was processed according to USDA organic standard. The processing involved cold processing and ultrasound. In this study, feedstock loading is one of the DoE inputs, which is set at 20 to 60 g/L of range.

2.2. Alkali hydrolysis

Sodium hydroxide (NaOH) was used as the alkali agent to hydrolyze *C. pyrenoidosa*. This stage is purposed to alter the complex sugar structure of *C. pyrenoidosa* into simple sugar chains that are fermentable. Simple sugars, including glucose, are essential in fermentation stage as the microbial activity of the yeast converts sugar into ethanol.

In preparation of NaOH, solutions of NaOH with different concentrations were planned in 100 ml stock and mixed with distilled water. As the inputs in the DoE, hydrolysis temperature and duration are also considered as variables in this study. The set range for NaOH concentrations is 0.3 M as minimum and 0.6 M as maximum. The set range for hydrolysis temperature is 50 to 90°C, and for hydrolysis time is 15 to 45 minutes.

2.3. Inputs in design of experiment (DoE)

This experiment uses Design Expert 10 software to formulate the design of experiment (DoE), using response surface method (RSM) with Box-Behnken. The selection of Box-Behnken in RSM is due to the practicality in optimization study of biofuel production.

Prior formulating the design of experiment, several inputs are needed. Four factors were set as controlled variables in the experiment: (A) feedstock loading, g/L; (B) NaOH concentration, M; (C) hydrolysis temperature, °C; (D) hydrolysis time, min. Limits of those factors were inputted as minimum
and maximum, which coded as -1 and 1 respectively. In here, the limits are according to the set range of each factor, as described in the previous sections. The response set in this study is the concentration of detected reducing sugars, in g/L. The complete inputs of this study are shown in Table 1 below.

**Table 1.** Design inputs for Box-Behnken response surface methodology DoE.

| Factor code | Factor name                      | Units | Min limit | Max limit | Coded values       | Mean  | Std.dev. |
|-------------|----------------------------------|-------|-----------|-----------|-------------------|-------|----------|
| A           | Feedstock loading                | g/L   | 20        | 60        | -1.000=20; 1.000=60 | 40    | 1.309    |
| B           | NaOH concentration              | M     | 0.3       | 0.6       | -1.000=0.3; 1.000=0.6 | 0.45  | 0.098    |
| C           | Hydrolysis temperature          | °C    | 50        | 90        | -1.000=50; 1.000=90 | 70    | 13.093   |
| D           | Hydrolysis time                 | Minutes | 15     | 45        | -1.000=15; 1.000=45 | 30    | 9.819    |

3. Results and discussions

3.1. Design of experiment summary

Based on the inputs in Table 1 of Box-Behnken response surface methodology, Stat-Ease generates 29 data points to run in the experiment. The generated data points are tabulated in Table 2 below.

**Table 2.** Design of experiment through response surface methodology, Box-Behnken design.

| Run no. | Factor A (g/L) | Factor B (M) | Factor C (°C) | Factor D (min) | Run no. | Factor A (g/L) | Factor B (M) | Factor C (°C) | Factor D (min) |
|---------|----------------|--------------|---------------|----------------|---------|----------------|--------------|---------------|----------------|
| 1       | 40             | 0.45         | 90            | 15             | 16      | 20             | 0.3          | 70            | 30             |
| 2       | 20             | 0.45         | 90            | 30             | 17      | 20             | 0.45         | 70            | 15             |
| 3       | 60             | 0.45         | 70            | 45             | 18      | 40             | 0.6          | 70            | 15             |
| 4       | 20             | 0.6          | 70            | 30             | 19      | 40             | 0.45         | 70            | 30             |
| 5       | 40             | 0.45         | 50            | 15             | 20      | 40             | 0.3          | 90            | 30             |
| 6       | 40             | 0.3          | 70            | 45             | 21      | 40             | 0.6          | 70            | 45             |
| 7       | 40             | 0.3          | 50            | 30             | 22      | 40             | 0.45         | 70            | 30             |
| 8       | 60             | 0.6          | 70            | 30             | 23      | 40             | 0.6          | 50            | 30             |
| 9       | 20             | 0.45         | 70            | 45             | 24      | 60             | 0.45         | 70            | 15             |
| 10      | 40             | 0.45         | 70            | 30             | 25      | 40             | 0.45         | 70            | 30             |
| 11      | 60             | 0.45         | 90            | 30             | 26      | 40             | 0.6          | 90            | 30             |
| 12      | 20             | 0.45         | 50            | 30             | 27      | 40             | 0.3          | 70            | 15             |
| 13      | 60             | 0.3          | 70            | 30             | 28      | 40             | 0.45         | 50            | 45             |
| 14      | 60             | 0.45         | 50            | 30             | 29      | 40             | 0.45         | 90            | 45             |
| 15      | 40             | 0.45         | 70            | 30             |         |                |              |               |                |

Factor A: feedstock loading; Factor B: NaOH concentration; Factor C: hydrolysis temperature; Factor D: hydrolysis time

As per definition, despite the many experiment runs that must be conducted (total 29 runs from Table 2), in comparison with one-factor-at-time method Box-Behnken RSM reduces a significant amount of
number of runs. If it were to consider the same three-level with four factors in conventional optimization DoE, there would be 12 x 12 DoE matrix, totalling of 144 experiment data points. The three-level factor from Box-Behnken (coded -1, 0, and 1) provides accuracy and reliability in the resulted response.

3.2. Evaluation of the design of experiment

The prediction model based from the four factors as input is further estimated. To begin the evaluation, prediction model under second-order polynomial (quadratic) was selected. The predicted model follows the following Eq. 1.

\[ y = \beta_0 + \sum_{i=1}^{k} \beta_{i} x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j}^{k} \beta_{ij} x_i x_j + e \]  

(1)

where \( \beta_0 \) is the intercept value, \( x_i \) is the linear term of the model, representing each of the assigned factor. \( \beta_{ii} \) is the coefficient for quadratic term, while \( \beta_{ij} \) is the coefficient for the interaction term. Quadratic term is defined by \( x_i^2 \), while the interaction term is defined by \( x_i x_j \). \( e \) indicates any random error. From four factors (A, B, C, D, see Table 1), it results to 14 terms which comes from the linear, interaction and quadratic. They are: A, B, C, D, AB, AC, AD, BC, BD, CD, A^2, B^2, C^2, D^2.

According to the prediction model for the four factors, no aliases found for the selected quadratic model to predict the design. This characteristic is favourable as this indicates the likability of the estimated coefficients for the predicted model. It is also predicted that using quadratic model it would only result degrees of freedom of 10 and 4 for lack-of-fit and pure error, respectively. This is appropriate as it suggests a valid lack-of-fit test, as it predicted. At each term, variance of inflation factor (VIF) shows value of unity. Value of VIF illustrates the lack of orthogonality of the design according to the increasing variance of the coefficient model. VIF of 1 is ideal, as greater value of VIF implies large variance of the coefficients due to collinearity.

Design evaluation continued further by analysing the standard design of error of the predicted quadratic model for the design. Figure 1 depicts the example of the predicted standard error of design for this study at an instant experimental condition. The shading of the contour plot reflects the error intensity; the darker the color the greater the predicted error. The error contour is depicted according to the set factors. NaOH concentration and feedstock loading concentration set at the axes of the contour, while the remaining two factors, hydrolysis temperature and time, impact the intensity of standard error of design as they are adjusted accordingly. This is illustrated at the difference between Figure 1.A. and 1.B. The former displays the low, favorable predicted error contour resulted from the mild, mid-range of all the four factors. Figure 1.B., produced by the end-extremes of hydrolysis temperature and time (taken from run #29), however shows a greater predicted error contour as the darker region suggests. This is only logical due to the nature of Box-Behnken design in RSM. Figure 2 illustrates how the design of experiment is constructed from the distributed dots, signalling experiment runs. As this design excludes the extremes (i.e. points at each corner of the box), it is only rational that the predicted standard error of design becomes higher as the factors approaching the extreme values. According to this contour plot, experiments based on this design will lead to higher reliability with factors at intermediate level, as indicated by the standard error of design value less than 0.5. After all, the evaluations above are predictive of the set quadratic model for the design. The model can only be fully valid once all the runs are conducted and values for responses (i.e. reducing sugar concentration) are filled. As this is preliminary result, similarly, optimization will fully take place after responses have been filled.
4. Conclusion
Design of experiment in the pursue to optimize bioethanol production from *Chlorella pyrenoidosa* was formulated in this study. This study considers Box-Behnken design in response surface methodology, and it yields total of 29 experimental runs. This approach has reduced tremendous number of experimental runs (from 144 runs conventionally), which elevates efficiency and overall cost-effectiveness. The design was further analyzed based on the predicted model, using second-order polynomial. Based on the prediction analysis, the design shows a fair suitability to be proceeded in modelling the response via quadratic polynomial. As this study reports the preliminary stage of optimization, the responses should be submitted to the design, meaning all runs must be conducted next to ensure the validity of the prediction. Additionally, optimization should be commenced in the following of obtaining all the responses values.
Acknowledgement

Authors would like to acknowledge the support from Centre of Research and Community Service, Sampoerna University, under Internal Research Grant scheme no. 022/IRG/SU/AY.2019-2020.

References

[1] Rastogi M and Shrivastava S 2017 Recent advances in second generation bioethanol production: an insight to pretreatment, saccharification and fermentation processes Renewable and Sustainable Energy Reviews 80 330-340

[2] Aditiya H B, Mahlia T M I, Chong W T, Nur H and Sebayang A H 2016 Second generation bioethanol production: A critical review Renewable and sustainable energy reviews 66 631-653

[3] Domínguez E, Romaní A, Domíngues L and Garrote G 2017 Evaluation of strategies for second generation bioethanol production from fast growing biomass Paulownia within a biorefinery scheme Applied Energy 187 777-789

[4] Aditiya H B, Mahlia T M I, Chong W T, Nur H and Sebayang A H 2016 Second generation bioethanol production: A critical review Renewable and sustainable energy reviews 66 631-653

[5] Putrasari Y, Praptijanto A, Santos W B and Lim O 2016 Resources, policy, and research activities of biofuel in Indonesia: A review Energy Reports 2 237-245

[6] Slade R and Bauen A 2013 Micro-algae cultivation for biofuels: cost, energy balance, environmental impacts and future prospects Biomass and bioenergy 53 29-38

[7] Aditiya H B, Theofany H C, and M Yheni 2019 High-pressurizing green algae in third generation bioethanol production Journal of Physics: Conference Series, 4th Annual Applied Science and Engineering Conference (AASEC) 1402

[8] Theofany H C, Yheni M, Aditiya H B and Sepwin N S 2019 Microwave irradiation pre-treatment in third generation bioethanol production from tropical green algae In Journal of Physics: Conference Series 1402(4) 044046

[9] Yheni M, Theofany H C, Aditiya H B and Sepwin N S 2019 Preliminary study of acidic hydrolysis in third generation bioethanol production using green algae In Journal of Physics: Conference Series 1402(4) p 044047

[10] Ferreira S C, Bruns R E, Ferreira H S, Matos G D, David J M, Brandao G C and Dos Santos W N L 2007 Box-Behnken design: an alternative for the optimization of analytical methods Analytica chimica acta 597(2) 179-186

[11] Sebayang A H, Hassan M H, Ong H C, Dharma S, Silitonga A S, Kusumo F and Bahar A H 2017 Optimization of reducing sugar production from Manihot glaziovii starch using response surface methodology Energies 10(1) 35