A Multilayer Full Feed Forward Neural Network for Modelling and Optimizing Nutrient Medium Composition for Oxalic Acid Production from Sweet Potato Peels

Amenaghawon, N.A. and Agbedor, M.
Department of Chemical Engineering,
University of Benin, PMB 1154, Benin City, Nigeria.

ABSTRACT

This study was aimed at optimizing the composition of nutrient medium for optimum production of oxalic acid from sweet potato peel using Aspergillus niger. The fermentation experiments were designed using a three-variable Box-Behnken design (BBD). The effect of the nutrient medium components on oxalic acid yield was modelled using a multilayer full feed-forward artificial neural network (MFFF-ANN) trained with the Levenberg-Marquardt algorithm. The nutrient medium components considered and their levels of variation were $\text{KH}_2\text{PO}_4$ (0.5-1 g/L), $\text{MgSO}_4$ (0-0.5 g/L) and $\text{NaNO}_3$ (1-2 g/L). Assessment of the predictive capacity of the ANN showed that it was able to replicate the experimental observations with a high level of precision ($R^2 = 0.9996$, $\text{RMSE} = 0.1845$, $\text{MAE} = 0.084$, $\text{AAD} = 0.030\%$). Analysis of the results showed that oxalic acid concentration was favoured by intermediate levels of the nutrient medium components. To maximize the yield of oxalic acid, the process input variables investigated were optimized by coupling genetic algorithm (GA) with the MFFF-ANN. The best combination of the estimated input variables ($\text{KH}_2\text{PO}_4 = 0.770$ g/l, $\text{MgSO}_4 = 0.194$ g/l and $\text{NaNO}_3 = 1.999$ g/l) gave the highest oxalic acid yield of 29.023 g/l. This study has demonstrated the applicability of a machine learning tool like ANN for modelling and optimising a biochemical process like oxalic acid production.

INTRODUCTION

Oxalic acid is a dicarboxylic acid which can be found in high levels in its natural form in many foods such as potatoes, carrots, beets, broccoli, legumes, grains etc. It is widely used in the pharmaceutical industry, textile and wood industry as a cleaning agent, food industry as a chelating agent, agricultural industry as a preservative and anti-browning agent, polishing of marbles, as well as in tanning, dyes and explosives industries (Edewor-Kuponiyi and Amuda, 2013; Cefola and Pace, 2015; Walaszczyk et al., 2017). It is also a very important chemical in petroleum processing, corrosion inhibition and dental adhesive processing.

Production of oxalic acid can be done through chemical processes (Nakata and He, 2010). These include oxidation of olefins and glycol, radiation processing of carbonate solutions and molasses (Mandal and Banerjee, 2005). These chemicals are known to be unsafe for the environment and may not be cost-effective. Therefore, the need for the biological production of the acid is inevitable (Betiku et al., 2016).

Nakata and He, (2010) reported that oxalic acid can be produced by making use of microbes (fungi, bacteria), plants and animals. Microorganisms such as Aspergillus niger (Strasser et al., 1994),
Penicillium oxalicum (Li et al., 2016) and Burkholderia glumae have been used for oxalic acid production through fermentative processes. Aspergillus niger is preferred to other organisms because of its high productivity and easy accessibility (Chioma and Agwa, 2019).

In order to boost the production of oxalic acid production through the fermentative method, reduction in its cost of production is important and it can be achieved by using cheap agricultural products such as sweet potato peels. Sweet potato (Ipomoea batatas L.) is among the major food crops in the world cultivated in all tropical and subtropical regions particularly in Asia, Africa and the Pacific. Nigeria is the third-largest producer of sweet potatoes in the world followed by China and Uganda (Ahmad et al., 2014). In developing countries such as Nigeria, one of the problems faced is improper disposal of the sweet potato peel and as a result, causes environmental pollution. In order to proffer solution to this pollution problem, value addition to these peels to produce other useful products such as oxalic acid is imperative. The right fermentation technique should be selected and optimization of the fermentation variables that are involved in the process should be carried out. The parameters that affect the yield of oxalic acid include pH, fermentation time, medium composition and nitrogen source. Also, due to the complexity of the metabolic rate in fungus for the improved accumulation of desired product, there is an obvious need to optimize the important process variables (Singh et al., 2011). These parameters are optimized to increase oxalic acid yield and have been done traditionally using the one factor at a time approach which is time-consuming and demanding because of large numbers of experimental runs (Amenaghawon et al., 2015). As a result of the limitation of the one factor at a time approach, methods like the design of experiments (DOE), response surface methodology (RSM) and artificial neural network (ANN) have been adopted in recent times (Emeko et al., 2015). These methods have been previously adopted for optimising numerous biochemical processes such as citric acid production (Imandi et al., 2007; Amenaghawon et al., 2013), bioethanol production (Jambo et al., 2019), biodiesel production (Ayoola et al., 2019), oxalic acid production (Betiku et al., 2016) etc.

Thus, this study aimed to optimise the nutrient medium composition for the production of oxalic acid from sweet potato peels using Aspergillus niger. A three-variable Box-Behnken design was used to design the experiments involving three independent variables (KH2PO4, NaNO3, and MgSO4 concentration) for optimum oxalic acid production.

MATERIALS AND METHODS

Material Collection and Preparation

Sweet potato peels used for this experiment were obtained from a local market in Ekoosdin, Benin City, Edo State Nigeria. The peels were washed with clean water to remove dirt after which it was dried in an electric laboratory oven at 60 °C for 2 hours. The dried peels were grinded with a laboratory mill and sieved to obtain 1 mm particle size.

Microorganism and Inoculum

Aspergillus niger obtained from the Department of Microbiology of the University of Benin, Benin City, Edo State, Nigeria, was used throughout the study as a fermenting organism. Aspergillus niger spore was obtained from cultures grown on Sabouraud dextrose agar (SDA) for 5 to 7 days at 30 °C. This fungus was maintained on SDA plates at 4 °C and subcultured regularly.

Medium Composition

The unoptimised fermentation medium as described by Betiku et al. (2014) was employed in this study. It was composed of sweet potato peel as carbon source, 1.6 g/L of urea, 0.025 g/L of KCl, 0.50 g/L of KH2PO4, and 0.025 g/L
MgSO$_4$.7H$_2$O. The remaining components (KH$_2$PO$_4$, NaNO$_3$, and MgSO$_4$) were optimised according to the experimental design.

Production of Oxalic Acid by Solid-State Fermentation

The solid substrate (10 g) was added in the flask and wetted with the nutrient medium to 80% moisture. The contents were thoroughly mixed, cotton plugged and autoclaved at 121 °C and 15 psi for 15 minutes. After cooling, the autoclaved substrate including the media was inoculated with 2 ml of inoculum of density $2 \times 10^7$ spores/ml and then incubated at 30 °C for 9 days. At the end of fermentation, the medium was diluted with 100 ml of distilled water, filtered, and the filtrate was used for subsequent analysis.

Oxalic acid assay

The concentration of oxalic acid produced was determined using a catalytic kinetic spectrophotometric method (Zhi-Liang et al., 1996). This technique is based on the acid catalytic effect of the redox reaction between rhodamine B and dichromate at the maximum absorption wavelength of 555 nm in sulphuric acid. For the assay, 10 mL of the sample was withdrawn from the fermentation medium and filtered with Whatman No. 4 filter paper. Subsequently, to 1 mL from the filtrate was added 0.5 mL of 0.06 mol/L potassium dichromate (KzCrO$_4$), 0.20 mL of 0.25 mol/L sulphuric acid (H$_2$SO$_4$) and 0.1 mL of $3.28 \times 10^{-4}$ mol/L rhodamine B in a 10 mL test tube and then diluted to the mark with water and mixed thoroughly. The mixture was placed in a water bath at 90 °C. After 8 min, the reaction was quenched by cooling with tap water and the absorbance of the mixture read at 555 nm against the blank solution. The quantity of oxalic acid produced was determined using a standard calibration curve prepared using oxalic acid (Adesina et al., 2014).

Design of Experiment

A three-level three-factor Box- Behnken design employed to design the fermentation experiments. The independent variables considered for this study and their ranges include KH$_2$PO$_4$ (0.5-1.0 g/l), NaNO$_3$ (1-2 g/l), and MgSO$_4$ (0-0.5 g/l) which are denoted as $X_1$, $X_2$ and $X_3$ in Table 1. The experimental design was developed using Design Expert® 7.0.0 (Stat-ease, Inc. Minneapolis, USA), statistical software.

<table>
<thead>
<tr>
<th>Table 1: Coded and actual values of factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>NaNO$_3$ (g/l)</td>
</tr>
<tr>
<td>MgSO$_4$ (g/l)</td>
</tr>
<tr>
<td>KH$_2$PO$_4$ (g/l)</td>
</tr>
</tbody>
</table>

ANN Model Development

The ANN model was developed in a commercial neural network software Neural Power, version 2.5 (C.P.C-X Software USA). A multilayer full feedforward neural network consisting of an input layer, a hidden layer and an output layer was adopted and trained with the Levenberg-Marquadt algorithm. The training process was done using 70% of the experimental data. Of the remaining 30%, 15% was used for validating the model while the last 15% was used for testing the model. Validation and testing were done to assess the predictive capability of the ANN model. The optimum number of neurons in the hidden layer was determined using an iterative process.
This was done to avoid underfitting and overfitting problems. The hyperbolic-tangent transfer function was chosen for the hidden layer. The selection of the optimum number of neurons was done based on the coefficient of determination ($R^2$) and the root mean square error (RMSE). The best alternative was chosen to be the one with the highest $R^2$ value and the smallest RMSE value (Betiku et al., 2014).

**Verification of the ANN model**

The proficiency of the ANN model ANN in predicting the actual experimental observations was assessed using some statistical indicators such as coefficient of determination ($R^2$), root mean square error (RMSE), mean absolute error (MAE) and average absolute deviation (AAD). Equations 1 to 4 were used to compute these statistical indices. The coefficient of determination $R^2$ described by Equation 1 indicates the degree of fit for the model (Nath and Chattopadhyay, 2007). The closer the $R^2$ value is to 1, the better the model fits the actual data (Sin et al., 2006). The RMSE, MAE and AAD between predicted and experimental values must be as small as possible (Amenaghawon et al., 2017).

\[
R^2 = 1 - \frac{\sum_{i=1}^{n}(x_{a,i} - x_{p,i})^2}{\sum_{i=1}^{n}(x_{p,i} - x_{a,ave})^2}
\]

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{p,i} - x_{a,i})^2}
\]

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |(x_{a,i} - x_{p,i})|
\]

\[
AAD = \frac{1}{n} \left(\sum_{i=1}^{n} \left(\frac{|x_{a,i} - x_{p,i}|}{x_{a,i}}\right)\right) \times 100
\]

Where $n$ is the number of experimental data, $x_{p,i}$ is the predicted value, $x_{a,i}$ is the experimental values, $x_{a,ave}$ is the average experimental values, and $x_{p,ave}$ is the average predicted values.

**RESULTS AND DISCUSSION**

**Analysis of the ANN Model**

Statistical assessment of the Levenberg-Marquadt algorithm for training the network showed that it was adequate for the purpose. This is seen in the very high $R^2$ value ($R^2=0.9996$) and low RMSE value (RMSE=0.1845).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.9996</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.1845</td>
</tr>
</tbody>
</table>

Figure 1 shows the results of an investigation carried out to determine the optimum number of neurons in the hidden layer. Since the optimum number of neurons was chosen based on the $R^2$ value, it can be seen that a network 3 neurons were deemed optimum. This is because the trend in Figure 1 showed increasing $R^2$ values with an increase in the number of neurons up to 3. However, increasing the number of neurons beyond 3 did not result in improvement in the $R^2$ value. Thus, a neural network with three input factors, four neurons in the hidden layer and 1 factor in the output layer denoted as 3-3-1 was chosen for the study as shown in Figure 2.
The predictions of the ANN model were compared with those of the actual experiments and the results are shown in Table 3. The values predicted by the ANN model were very similar to those of the experiments indicating the validity and reliability of the ANN model. The validity of the ANN model was further assessed by evaluating the goodness of fit characteristics of the model and the results are shown in Table 3. The $R^2$ value was almost equal to one ($R^2 = 0.9996$) indicating that the model was able to explain 99.96% the variability in the results. Furthermore, the RMSE, MAE and AAD values were 0.1845, 0.084 and 0.030 respectively. The small magnitude of these values indicates very little variation between the model predictions and the experimental observations. This is an indication of a very good fit between the model and the experimental data (Yi et al., 2009).
Table 3: Experimental and ANN predicted oxalic acid concentration

<table>
<thead>
<tr>
<th>Run no</th>
<th>Coded value of factors</th>
<th>Actual value of factors</th>
<th>Response (g/l)</th>
<th>Experimental value</th>
<th>ANN prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X1 X2 X3</td>
<td>X1 X2 X3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0 1 -1</td>
<td>1.5 0.50 0.50</td>
<td>10.83</td>
<td>10.83</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-1 -1 0</td>
<td>1.0 0.00 0.75</td>
<td>3.55</td>
<td>3.55</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-1 0 1</td>
<td>1.0 0.25 1.00</td>
<td>7.31</td>
<td>7.31</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-1 1 0</td>
<td>1.0 0.50 0.75</td>
<td>19.56</td>
<td>19.56</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-1 0 -1</td>
<td>1.0 0.25 0.50</td>
<td>6.08</td>
<td>6.08</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0 0 0</td>
<td>1.5 0.25 0.75</td>
<td>28.14</td>
<td>28.39</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0 0 0</td>
<td>1.5 0.25 0.75</td>
<td>28.12</td>
<td>28.39</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0 1 1</td>
<td>1.5 0.50 1.00</td>
<td>12.80</td>
<td>12.80</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1 1 0</td>
<td>2.0 0.50 0.75</td>
<td>15.58</td>
<td>15.58</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1 0 -1</td>
<td>2.0 0.25 0.50</td>
<td>12.05</td>
<td>12.05</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0 0 0</td>
<td>1.5 0.25 0.75</td>
<td>28.15</td>
<td>28.39</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0 -1 -1</td>
<td>1.5 0.00 0.50</td>
<td>5.02</td>
<td>5.02</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0 -1 1</td>
<td>1.5 0.00 1.00</td>
<td>7.69</td>
<td>7.69</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0 0 0</td>
<td>1.5 0.25 0.75</td>
<td>28.50</td>
<td>28.39</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1 -1 0</td>
<td>2.0 0.00 0.75</td>
<td>18.66</td>
<td>18.66</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1 0 1</td>
<td>2.0 0.25 1.00</td>
<td>15.37</td>
<td>15.37</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0 0 0</td>
<td>1.5 0.25 0.75</td>
<td>28.99</td>
<td>28.39</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Goodness of fit statistics for the ANN model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.9996</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.1845</td>
</tr>
<tr>
<td>MAE</td>
<td>0.0840</td>
</tr>
<tr>
<td>AAD (%)</td>
<td>0.0300</td>
</tr>
</tbody>
</table>

Figure 3 shows the parity plot that compares the oxalic acid concentration predicted by the ANN model to the actual experimental values. Its purpose is also to assess the level of fit between the experimental and model-predicted results. The results showed that there was an acceptable level of fit between the experimental and model-predicted results. This is evident from the fact that the data points all clustered around the 45° diagonal line showing that there was minimal deviation between experimental and predicted values.
Effect of Nutrient Medium Components on Oxalic Acid Production

Response surface plot shows the visual observation among (KH$_2$PO$_4$, MgSO$_4$ and NaNO$_3$) for the oxalic acid production. Figure 4 shows the effect of MgSO$_4$ and KH$_2$PO$_4$ on oxalic acid concentration. The response surface plot showed an elliptical shape. This is an indication that intermediate levels of MgSO$_4$ and KH$_2$PO$_4$ were favourable for oxalic acid production. Furthermore, the elliptical shape shows that there was a significant interaction between the level of MgSO$_4$ and KH$_2$PO$_4$. These observations have been reported in previous studies (Emeko et al., 2015; Betiku et al., 2016). Figure 5 shows the effect of NaNO$_3$ and KH$_2$PO$_4$ on oxalic acid concentration. There was also a very significant interaction between the two factors as seen in the elliptical shape of the 3D plot. Similarly, intermediate levels of NaNO$_3$ and KH$_2$PO$_4$ resulted in maximum oxalic acid concentration. It was observed that the presence of these nutrients in the fermentation medium facilitated the production of oxalic acid. For instance, magnesium has been reported to enhance the growth and metabolic activity of *Aspergillus niger* (Shankaranand and Lonsane, 1994).
Optimization of Oxalic Acid Production

The optimum conditions of the input variables were determined by incorporating the genetic algorithm with the ANN model. The optimum nutrient medium composition is shown in Table 4. The concentration of KH₂PO₄, MgSO₄ and NaNO₃ were obtained as 0.77 g/l, 0.194 g/l and 1.999 g/l. This resulted in a maximum oxalic acid concentration of 29.023 g/l.

Table 5: Optimized medium composition

<table>
<thead>
<tr>
<th>KH₂PO₄</th>
<th>MgSO₄</th>
<th>NaNO₃</th>
<th>Oxalic acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.770 g/l</td>
<td>0.194 g/l</td>
<td>1.999 g/l</td>
<td>29.023 g/l</td>
</tr>
</tbody>
</table>

CONCLUSION

This study adopted a multilayer full feed-forward artificial neural network to model and optimise the nutrient medium composition for producing oxalic acid from sweet potato peels. The ANN model was very modelling the oxalic acid production process as seen in its very good statistical parameters (R²=0.9996, RMSE=0.1845, MAE= 0.084, AAD=0.030). Oxalic acid production is favoured at intermediate levels of KH₂PO₄, MgSO₄ and NaNO₃. The optimum values of KH₂PO₄, MgSO₄ and NaNO₃ were 0.770 g/l, 0.194 g/l and 1.999 g/l respectively. Under this condition concentration of the oxalic acid was obtained as 29.023 g/l. This study has thus demonstrated the suitability of ANN in optimising oxalic acid production.

REFERENCES


Optimisation of Solid-State Fermentation of Banana Peels for Citric Acid Production. 
_Nigerian Journal of Technology, 34_(4), pp. 716-723._


