J. Trop. Resour. Sustain. Sci. 3 (2015): 174-179

Journal of Tropical Resources and Sustainable Science

journal homepage: jtrss.org

Growth of Mutant *Synechococcus* SP. PCC 7002: Effects of Multi-Parameters and Prediction of Growth Rate

Nor Fadhillah Mohamed Azmin*, Atikah Mohamed Sharikh, Ummi S. Halmi Shari, Azlin Suhaida Azmi

Department of Biotechnology Engineering, International Islamic University Malaysia, 51000 Gombak, Selangor, Malaysia.

Available online 15 June 2015

Keywords:

Principal Component Analysis (PCA), Partial Least Squares (PLS), cyanobacteria, *Synechococcus*, growth.

⌧*Corresponding author:
Dr. Nor Fadhillah Mohamed Azmin,
Department of Biotechnology Engineering, International Islamic University Malaysia, 51000 Gombak, Selangor, Malaysia.
Email: norfadhillah@iium.edu.my

Abstract

Understanding of the correlative effects of combined variables on the growth rate of the cyanobacteria is fundamental to the exploitation of cyanobacteria as a biological mechanism to produce biofuels. Cyanobacteria (blue-green algae) are phototrophic microorganisms that offers attractive benefits, among which is a direct conversion of CO2 to a range of valuable products such as carbon-based biofuels. One model of cyanobacteria species is the cyanobacterium *Synechococcus* sp. PCC 7002. This paper describes the model developed to investigate the combined impacts of the variables on the growth of the *Synechococcus* sp. PCC 7002. The variables understudy include the temperature of the media, light intensity, the concentration of $NaNO_3$, and the concentration of the NPK. The data is obtained from a lab scale study in which the *Synechococcus* sp. PCC 7002 underwent mutagenesis procedures. It is hypotheses that certain combination of the variables plays a key role in determining the growth rate of *Synechococcus* sp. 7002. The growth rate is determined through the measurement of four response variables, carbohydrate concentration, percentage of CO_2 uptake, cell dry weight (CDW), and optical density (OD). A multivariate PCA model was developed which unearths the underlying relationship between the variables. Promising results were yield from the proposed model. Distinctive correlations between the variables were clearly described by the PCA model.

1. Introduction

Marine cyanobacterium *Synechococcus* sp. PCC (Pasteur Culture Collection) 7002 is a mutated unicellular bacteria species from a wild type strain PR6000. For the past few decades, it has been manipulated to be used widely in many fields such as medical and engineering. In biotechnology application, *Synechococcus* sp. PCC 7002 is considered as one of valuable species due to its ability to directly convert CO2 to a wide range of high-demanded products under photoautotrophic and dark condition [1] and through photosynthesis as the primary step. [2] emphasizes the application of *Synechococcus* sp. to reduce CO_2 content in the atmosphere by using a genetically engineered *Synechococcus* elongatus PCC7942 to produce isobutyraldehyde and isobutanol directly from CO_2 and increased productivity by overexpression of ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco). The performance of *Synechococcus* sp. in various applications is affected by many factors, an in depth study regarding these factors are required to understand the *Synechococcus* sp. behavior to bring each application to an optimization.

For every changes occur in environmental condition where the cyanobacteria inhabit, there exist a direct effect the growth rate of *Synechococcus* sp. A large number of experimental studies and researches have been conducted to observe how changes in parameters such as temperature, light intensity, carbon dioxide (CO2) and iron affect growth rate of *Synechococcus* sp. [3,4,5,6]. A lot more of studies have been done by the other researchers to understand the

ISSN Number: 2289-3946

behavioral changes that result from the changes of these parameters. It is undeniable that studies on factors affecting growth *Synechococcus* sp. has been done worldwide for decades. However, the correlation between parameters has not yet been discovered especially when it involves more than 2 parameters.

The aim of this study was to develop a multivariate model to determine the correlation of the parameters that affect the growth rate of *Synechococcus* sp. PCA and PLS models were developed for the prediction of the growth rate and CO_2 uptake. The major goals of PCA are to extract as much information as possible from the dataset, to compress the dataset by keeping only important information, to simplify the description of data, and to analyze the structure of the observation and variables by finding similarities and differences in the dataset (Badr & Noureldien, 2013). This method could be important for the rapid prediction of an up-scale process of the microbe.

2. Materials and Methods

2.1. Process Description

The dataset was obtained from experimental tests in laboratory scale. Total run for the experiments was 12 runs whereby each of them was conducted for 7 days so that a significant growth can be obtained for later evaluation. Controlled parameters identified in this experiment were temperature, light intensity, and concentration of sodium nitrate ($NaNO_3$), and NPK fertilizer. Temperature of the media was varied from 30℃ to 40℃ while 2500 lx and 3500 lx being the minimum and maximum light intensity respectively. For $NaNO_3$, the concentration was varied between 0 and 1 whereas 0 to 0.05 was the range for NPK concentration. Five response variables were measured in the assessment of the CO_2 uptake rate are lipid yield (g/ml), the amount of carbohydrate production (g/L), the amount of CO_2 residual after photosynthesis reaction (%v,v), optical density (OD) (growth/hr), and cell dry weight (CDW) (growth/hr. Of all variables, CDW and OD were measured on daily basis while the rest were measured on seven days interval. Under controlled condition, high value of each variable indicates high growth rate and vice versa.

Growth rate of the cell was assumed to be high when low CO_2 was measured the end of the experiment indicating high rate of photosynthesis reaction which led to high product formation. To determine the growth rate of the bacteria, the latter four response variables were considered. The data array W comprising four independent variables and five response variables is depicted in Figure 1. A sub-matrix X represents the controlled variables and a sub-matrix Y represent the response variable.

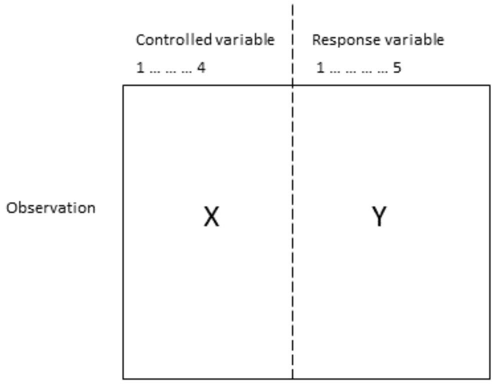

Figure 1: Data matrix *W*, of *Synechococcus* sp. PCC7002

2.2. Data Pre-processing

Data matrix X and Y are both heterogenous datasets because they comprise of variables of different units. Thus, both datasets require standardization prior to performing the multivariate principal component analysis (MPCA). An initial step in standardization involves mean-centering the datasets.

Dataset containing more than one array can be mean-centered in two ways: row mean-centered and column mean-centered. Each plays different roles depending on the data orientation and what is the final purpose of finding the mean value. Eq. (1) is used to determine the mean value for each row whilst Eq. (2) is used to evaluate the mean value for each column.

$$\overline{r_i} = \frac{\sum_{i=1}^{n} r_i}{n}. \qquad \text{Eq. (1)}$$

$$\overline{c_i} = \frac{\sum_{i=1}^{n} c_i}{n}. \qquad \text{Eq. (2)}$$

where $\bar{r}$ and $\bar{c}$ are row mean value and column mean value respectively, and r_i and c_i are

ISSN Number: 2289-3946

elements in the respected row and column . Mean centering is required needed to eliminate the multiple units thus all data can be treated equally.

For data matrix X, mean centering the row means finding the mean of all variables for each run. Thus, doing row mean centering is obviously will not give much information on influences of each parameter on growth as no comparison is made between all runs. In contrast, column mean centering gives average value for each run. Thus, performing column mean centering is more preferable for data matrix X as it compares the readings between runs when parameters are varied. Similar data pre-processing procedures were performed to the data matrix Y.

2.3. Multivariate Principal Component Analysis (MPCA)

Principal component analysis (PCA) is a multivariate technique that simplifies and describes interrelationships between multiple variables and the responses. PCA was performed with the use of the MATLAB (The Mathworks, Natick, MA) software packages. The software was used to systematically extract the variation in both data matrix X and Y. The validity of the models was evaluated by comparing root mean square error of calibration (RMSEC) and root mean square error of prediction (RMSEP). All of the PCA models reported here are based on only two principal components.

3. Results and Discussion

To make description of results obtained from data analysis more convenience, discussion will be divided into two parts; result and discussion for data analysis of controlled parameters followed by results and discussion for response variables.

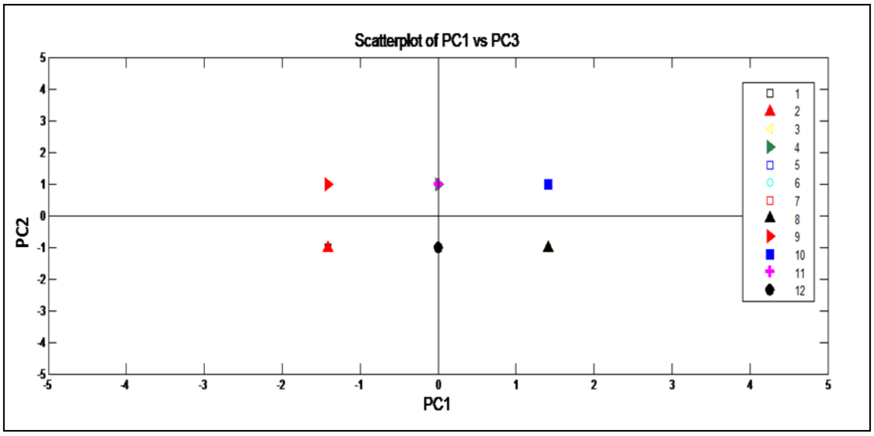

Figure 2: PC1 vs. PC2 of controlled variables

3.1. Multivariate analyses on the controlled variables

Scatterplot of principal component 1 (PC1) vs. principal component 2 (PC2) is depicted in Figure 2. Figure 2 exhibits a cluster of 6 instead of 12 runs. This was because 6 samples were overlapping on the other 6. Run 11 and 12 were positioned on PC1. These runs shared the same values of $NaNO_3$ and NPK percentage. We also observed that run 8 and 10 were positively influenced by PC1. These two runs shared the same readings for temperature and NaNO3 percentage. From this observation, we may conclude that the percentage of $NaNO_3$ is the main factor that affects the runs with respect to PC1. On the other hand, run 9, 11 and 10 were positive with respect to PC2 whilst run 2, 12 and 8 were negative. When we relay this figure to the original data matrix X, we found that the value of one controlled variable i.e. light intensity is of the same value for the positive runs and is of another same value for the negative runs.

3.2. Multivariate analyses on the response variables

Three principal components were retained in this analysis. Scatterplot of PC1 vs. PC2 is depicted in Figure 3. Figures 4 and 5 show the scatterplot of PC1 vs. PC2, and PC2 vs. PC3.

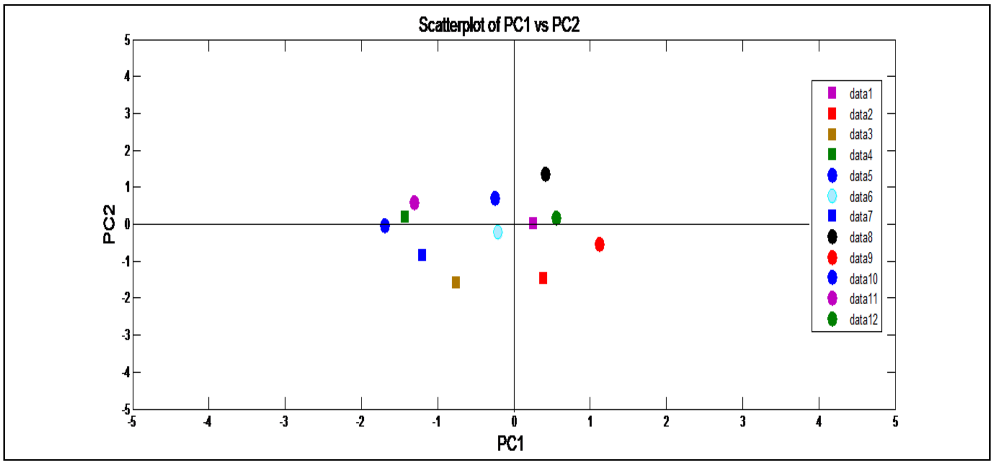

Figure 3: PC1 vs. PC2 of response variables

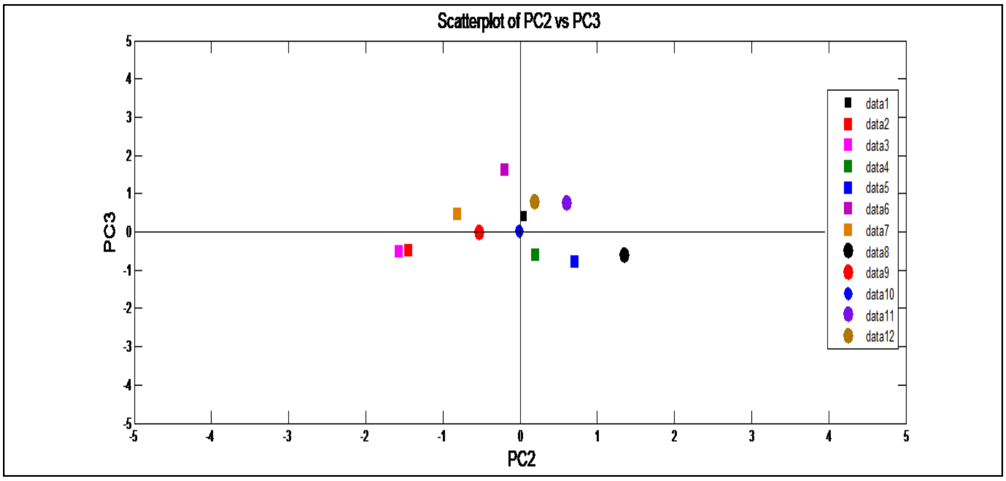

Figure 4: PC1 vs. PC3 response variables

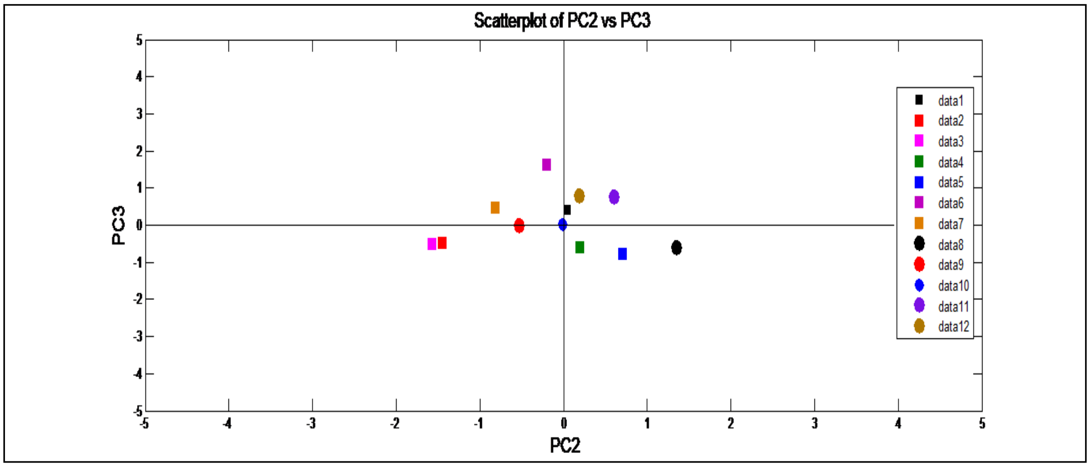

Figure 5: PC2 vs. PC3 response variables

ISSN Number: 2289-3946

In all scatterplots, samples distribution diverged into four quadrant. In clockwise direction, quadrant 1 (Q1) in top right, followed by quadrant 2 (Q2), quadrant 3 (Q3) and quadrant 4 (Q4) in the top left of plotting area. Table 1 shown below was constructed and referred to evaluate in terms of each sample. Positive quadrant in all components was regarded as Q1 while negative quadrant of all components was represented as Q3.

Based on Table 1, sample 1 and 12 were allocated in positive quadrant of PC1, PC2, and PC3. This indicated that the growth pattern of organism in sample 1 resembled or close to the one in sample 12. Thus, they are said to be highly correlated to each other. Relating this observation to the original data matrix W, sample 1 and 12 shared the same similarity such that both were tested under 2500 lux of light intensity and 0 %w/v of NPK fertilizer. In terms of temperature and NaNO3, sample 1 was tested at a higher ambient temperature and concentration of fertilizer. This finding concluded that lower amount of light intensity and NPK fertilizer influenced growth of *Synechococcus* sp. in a positive way.

Table 1: Sample position in four quadrants

Sample	Quadrant I	Quadrant II	Quadrant III	Quadrant IV
1	√√√			
2		√√	√	
3			√√√	
4		√	√	√
5		√	√	√
6			√	√√
7				
8	√	√√		
9		√√	√	
10			√	√√
11	√			√√
12	√√√			

Another significant information extracted from the scatterplots was how parameters tested could decline the growth rate of *Synechococcus* sp. According to Table 1, only one sample that showed negative growth rate which was sample 3. Sample 3 was tested at 40°C with 2500 lux of light intensity, 0 %w/v of $NaNO_3$, and 0.05%w/v of NPK fertilizer. Compared to first finding, the concentration of fertilizer used was higher to both sample 1 and 12. Therefore, it can be deduced that under 2500 lux of light intensity, high concentration of NPK fertilizer retarded growth rate of *Synechococcus* sp.

4. Conclusions

Study to identify parameters that affect growth of *Synechococcus* sp. using PCA conclude that in the presence of low concentration of NPK fertilizer, low light intensity improved organism's growth. Conversely, with the presence of high concentration NPK fertilizer, growth rate reduces.

Acknowledgement

The authors gratefully acknowledge the financial support from International Islamic University Malaysia (IIUM).

References

[1] Vu, T. T., Hill, E. A., Kucek, L. A., Konopka, A. E., Beliaev, A. S., & Reed, J. L. (2013). Computational Evaluation of *Synechococcus* sp. PCC 7002 Metabolism for Chemical Production. Biotechnology Journal, 8(5), 619-630. doi: 10.1002/biot.201200315

[2] Atsumi, S., Higashide, W., & Liao, J. C. (2009). Direct Photosynthetic Recycling of Carbon Dioxide to Isobutyraldehyde. Nat Biotech, 27(12), 1177-1180. doi: http://www.nature.com/nbt/journal/v27/n12/suppinfo/nbt.1586_S1.html

[3] C. Coman, V. B., C.Sicora. (2013). The Study of ps II Photochemical Activity in

ISSN Number: 2289-3946

Synechococcus sp. Pcc 7002 Under Different Light Intensities. Annals of RSCB, XVII(2), 7.

[4] Ludwig, M., & Bryant, D. A. (2012). *Synechococcus* sp. Strain PCC 7002 Transcriptome: Acclimation To Temperature, Salinity, Oxidative Stress and Mixotrophic Growth Conditions. Frontiers in Microbiology, 3. doi: 10.3389/fmicb.2012.00354

[5] Aiko Kimura, T. H., Eugena H. Morita, and Hidenori Hayashi. (2002). A High Temperature-sensitive Mutant of *Synechococcus* sp. PCC 7002 With MOdifications in The Endogeneous Plasmid, pAQ1. Plant Cell Physical, 43(2), 7.

[6] Badr, N., & Noureldien, N. A. (2013). Examining Outlier Detection Performance for Principal Component Analysis Method and Its Robustification Methods. 6(2), p. 573..

ISSN Number: 2289-3946