ORIGINAL PAPER

APPLICATION OF RESPONSE SURFACE METHODS TO DETERMINE CONDITION FOR POPULATION DOUBLING LEVEL PREDICTION

RABIATUL ADAWIYAH ABDUL ROHIM[1], NORHAYATI YUSOP[1],
WAN MUHAMAD AMIR W. AHMAD[1]

Manuscript received: 12.08.2019; Accepted paper: 05.01.2020;
Published online: 30.03.2020.

Abstract. *This paper mainly focuses on the application of response surface methodology (RSM) approach to biological data. RSM is an effective modelling tool to solve linear and nonlinear multivariate regression problem. Response surface methodology (RSM) is advanced tools for modelling and prediction of such complex of biotechnology process. Our aim is to determine condition for population doubling level prediction using response surface methodology analysis. The response was used to develop an empirical model that associated the response to the two process variables. The variables studied were the days of cultured cells (X_1) and passage of number cells (X_2). The response to this experiment was the total population doubling level (Y). The result shows that the predicted values were in good agreement with the experimental values in validation experiments, which confirmed the accuracy of the model. Therefore, this paper can offer the researchers with clear information about RSM methods. The RSM technique is a more effective and easier procedure and interpretation of experiments compared to the other traditional method.*

Keywords: *Response Surface Methodology, Population Doubling, Prediction, Modeling.*

1. INTRODUCTION

RSM is an effective modelling tool to solve linear and nonlinear multivariate regression problem. The RSM technique can stimulate and optimize complex processes because it allows more efficient and easier procedure and explanation of experiments compared to the other traditional method. It also provides a model equation relating the response parameter to the process variables and optimization of the same [1]. We denote the independent variables by $X_1, X_2, X_3 ..., X_k$. It is assumed that these variables are continuous and controllable by the experiment with negligible error.

The response, Y is assumed to be a random variable. RSM is used for the design and analysis of experiments; it seeks to relate an average response to the value of quantitative variables that affect response [2]. The relationship between the dependent variable and independent variables can be represented as $Y = f(X_1, X_2, X_3, ..., X_k) + \varepsilon$ where ε represents the noise or error observed in the response 'y'. If we denote the expected response by $E(Y) = f(X_1, X_2, X_3 ..., X_k) = n$ then surface represented by $n = f(X_1, X_2, X_3, ..., X_k)$ is called

[1] Universiti Sains Malaysia, School of Dental Sciences, 16150 Kubang Kerian, Kelantan, Malaysia.
E-mail: adawiyah5350@yahoo.com; wmamir@usm.my; norhayatiyusop@usm.my.

the response surface. As a result, the RSM was used to determine the optimum input variables that can appropriately reflect subjects' population doubling level.

RSM has important application in the design, development, and formulation of new products, as well as in the development of existing product design. It defines the effect of the independent variables, alone or in combination on the processes. In addition to analyzing the effect of independent variables, this experimental methodology generates a mathematical model, which describes the biochemical processes [3].

The doubling time is defined as the time required for a cell population to double its number of cells, or for a tissue to double its size. If all cells in the population were divided, the doubling time would be equal to the length of the cell cycle t_c, and this is commonly happening in cell cultures during the exponential growth phase when nearly all cells are dividing.

The cell populations in vitro and in vivo are a mixture of dividing cells, non-dividing cells, and dying cells. The doubling time of a cell population will be the result of cell cycle time, growth fraction and the rate of cell loss. This is further complicated by the heterogeneity of cell types in tissues [4].

2. METHODS

This study presents data comparing the isolation, expansion, and characterization of heterogeneous rat MSC from bone marrow and compact bone explants. Attention is given to determine condition for population doubling level prediction. The variables studied were the days of cultured cells (X_1) and passage of number cells (X_2). The response for this experiment was the total population doubling level (Y). The response was used to develop an empirical model that correlated the response to the two process variables. We use the response surface methodology (RSM) in this study to examine the changes in variable behaviour.

In general, a first-order model takes the form $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$ where $(\beta_0, \beta_1, \beta_2)$ are regression coefficients, (x_1, x_2) are independent or predictor variables, ε is random error and y is a dependent variable or a response variable. This analysis can be performed by MINITAB. At the first stage of analysis by using MINITAB, we have to define the custom response surface design for our datasets and follow with analysis response surface design. After running the analysis, MINITAB will produce important information related to the study of the model. The second step is performing the response surface counter plot and response surface plot (3D) is drawn to visualize the input-output relationships [5].

The statistical testing of the model, which includes a linear coefficient, is performed by ANOVA analysis with F-test to obtain the empirical correlation between input and output parameters. To examine the goodness of fit of the model, each term of the model is tested statistically which confirmed the significance of F-values with $p < 0.05$. The values of R^2, adjusted R^2, and predicted R^2 are obtained [6].

3. RESULTS AND DISCUSSION

3.1 DEVELOPING OF REGRESSION MODEL EQUATION

These data are used to determine the coefficient of the polynomial equation. These estimated coefficients for both the coded and actual values are shown in table 1 along with the coefficient of determination R^2, adjusted R^2 and predicted R^2. The models in term of the coded and actual value of variables are shown in equation (1) and equation (2).

$$Y = 40.26 + 13.55X_1 + 45.26X_2 + 34.91X_1^1 - 15.68X_2^2 \quad (1)$$

$$Y = 0.685829 - 0.639578X_1 + 3.48318X_2 + 0.00450865X_1^1 - 0.0324030X_2^2 \quad (2)$$

Table 1. Coefficient of the model.

Factor	**Coefficient**	
	Coded	**Uncoded (actual)**
Constant	40.26	0.685829
X_1	13.55	-0.639578
X_2	45.26	3.48318
X_1^1	34.91	0.00450865
X_2^2	-15.68	-0.0324030
R^2 : 99.90%		
Adjusted R^2 : 99.89%		
Predicted R^2 : 99.88%		

The positive signs in the models signify synergetic effects of factor while the negative sign indicates the antagonistic effect. The analysis of variance (ANOVA) of the regression model (Table 1) revealed R^2 value of 0.999, indicating that the model can explain 99.90% of the data variation and only 1% of the total variation was not explained by the model. For the model adequate, R^2 value should not be less than 0.75 [7]. However Koocheki et al., (2009) posited that a large value R^2 does not always imply that the regression model is a good one and such inference can only be made based on a similarly high value of adj R^2 [8].

The value of the adjusted determination coefficient (Adj R^2 = 0.9989) therefore confirmed that the model was highly significant, which indicated a good agreement between the experimental and predicted values of population doubling level. Thus the model is adequate for prediction in the range of experimental variables. According to Rai et al., (2016), adj R^2 and prediction R^2 should be within 20% to be in good agreement [9].

This requirement is satisfied in this study with a prediction the R^2 value is 0.9989. The model, therefore, offers 99.89% variability in prediction population doubling beyond the experimental range of process condition.

Table 2. Analysis of variance (ANOVA) for Total Population Doubling Level.

Source	df	Seq SS	Adj SS	Adj MS	F-value	P-value
Regression model	4	61374.9	61374.9	15343.7	9988.17	0.000
Linear	2	61096.4	22318.8	11159.4	7264.34	0.000
X_1	1	54935.7	12.1	12.1	7.87	0.008
X_2	1	6160.8	116.1	116.1	75.57	0.000
Square	2	278.4	278.4	139.2	90.63	0.000
X_1^1	1	207.5	203.9	203.9	132.72	0.000
X_2^2	1	70.9	70.9	70.9	46.15	0.000
Residual Error	39	59.9	59.9	1.5		
Total	43	61434.8				

Table 2 shows the ANOVA of each term of the quadratic model. A term is significant if the F-value is large and $p < 0.05$. From the table, the linear term X_1 and X_2 are significant and quadratic term X_1^1 and X_2^2 also significant. All terms have a significant effect on population doubling.

3.2 ANALYSIS OF RESPONSE SURFACE

The 3D response surface plot and the counterplot derived from the mathematical models were applied to determine the optimal condition. The relationship between independent and dependent variables was graphically represented by 3D response surface and 2D contour plots generated by the model (Figure 1). The following plot shows how variable days of cell culture and the variable passage of number cells affect the quality (contours) of total population doubling cells. The higher passage of number cells identifies a higher total population doubling.

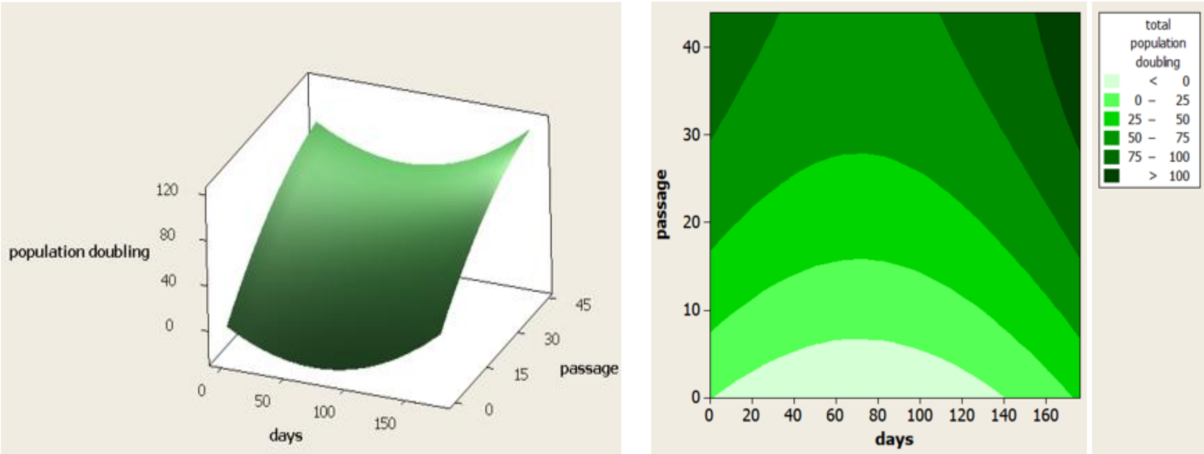

Figure 1. Contour and Surface Plot of Passage Vs Days and Total Population Doubling.

An elliptical or saddle nature of the contour plots indicated the significance of the interactions between the corresponding variables. From the response surface plots, it was easy and convenient to understand the interactions between the two factors and locate their optimum levels [1]. Contour plots are useful for establishing the response values and operating conditions are required.

4. CONCLUSION

The predicted values were in good agreement with the experimental values in validation experiments, which established the accuracy of the model. This paper can provide the researchers with clear information about RSM methodology and the outcomes obtained from the RSM methodology can help to study more about the changing behavior and changes according to their association. The RSM technique is more efficient and easier arrangement and interpretation of experiments compared to the other traditional method.

REFERENCES

[1] Amir, W.M., Shafiq, M., Mokhtar, K., Aleng, N.A., Rahim, H.A., Ali, Z., *Journal of Modern Applied Statistical Methods*, **15**(1), 44, 2016.

[2] Montgomery, D.C., *Design and Analysis of Experiments: Response Surface method and Designs*, 6th ed., John Wiley & Sons, Inc., Hoboken, 2005.

[3] Wang, Z.W., Liu, X.L., *Bioresource* Technology, **99**(17), 8245, 2008.

[4] Norouzzadeh, M., Kalikias, Y., Mohamadpur, Z., Sharifi, L., Mahmoudi, M., *International Research Journal of Applied and Basic Sciences*, **10**(3), 299, 2016.

[5] Yang, L., Jiang, J.G., Li, W.F., Chen, J., Wang, D.Y., Zhu, L., *Journal of Separation Science*, **32**(9), 1437, 2009.

[6] Khayet, M., Cojocaru, C., Essalhi, M., *Journal of Membrane Sci*ence, **368**(1-2), 202, 2011.

[7] Le Man, H., Behera, S.K., Park, H.S., *International Journal of Environmental Science & Technology*, **7**(1), 157, 2010.
[8] Koocheki, A., Taherian, A.R., Razavi, S., Bostan, A., *Food Hydrocolloids*, **23**(8), 2369, 2009.
[9] Rai, A., Mohanty, B., Bhargava, R., *Food Chemstry*, **192**, 647, 2016.