ISSN 0216-2393

Jurnal Gradien Vol. 10 No. 1 Januari 2014 : 957-962

Analisis Model Regresi Linear Berganda dengan Metode *Response Surface*

*Henoh Bayu Murti, Dian Kurniasari, Widiarti

Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung, Indonesia
**enoch_keepstraight@yahoo.com*

Diterima 10 April 2013; Disetujui 8 Juni 2013

Abstract - Problem in analysis of regression research is the need of more than one independent variables in regression model. This condition often causes complication in estimating an important response, which results multiple linear regression that is not optimal. In this condition, it is needed an optimalization that is essential in multiple linear regression. To be able to use independent variables optimally, response surface method is used. This method determines approaching function that is suitable in estimating the next dependent variable *y*. Optimal point of the result of the optimalization of multiple linear regression using three independent variables of the data of reaction rate for catalytic isomerization of n-pentane to isopentane is on saddle point.

Keywords: **multiple linear regression, response surface**

1. Pendahuluan

Sebagian besar masalah penelitian dalam analisis regresi adalah diperlukannya lebih dari satu variabel bebas dalam model regresi. Hal ini menyebabkan kerumitan dalam memprediksi sebuah respon yang penting. Oleh karena itu, model regresi berganda diperlukan [**1**]. Sering kali model regresi berganda yang telah dibuat tidak optimal. Oleh karena itu, pada model regresi berganda tersebut perlu dilakukan pengoptimalisasian. Optimalisasi sangat penting dalam analisis regresi berganda dikarenakan model yang optimal dapat menentukan fungsi pendekatan yang cocok untuk memprediksi respon variabel tak bebas *y* yang akan datang.

Jumlah variabel-variabel bebas yang banyak pada model regresi linear berganda akan mempengaruhi keoptimalan dalam pemilihan variabel-variabel bebas tersebut. Oleh karena itu, untuk dapat menggunakan variabel-variabel bebas secara optimal maka digunakan metode *response surface* [**2**].

Metode *response surface* merupakan suatu metode gabungan antara teknik matematika dan teknik statistik yang berguna untuk memodelkan dan menganalisis data dimana respon variabel tak bebas *y* yang diteliti dipengaruhi oleh beberapa variabel bebas dan bertujuan untuk mengoptimalkan respon tersebut. Hubungan antara respon variabel tak bebas *y* dan variabel-variabel bebas adalah:

$$y = f(x_1, x_2, \ldots, x_k) + \varepsilon$$

dimana y = variabel respon/variabel tak bebas, x_i = variabel bebas ($i = 1, 2, 3, \ldots, k$) dan ε = error yang diamati pada respon variabel tak bebas y. Jika respon yang diduga diasumsikan sebagai $E(y) = f(x_1, x_2, \ldots, x_k) = \eta$, maka permukaannya dilukiskan oleh $\eta = f(x_1, x_2, \ldots, x_k)$ yang disebut sebagai permukaan respon (*response surface*) [**3**].

Metode Kuadrat Terkecil (MKT) bagi Parameter pada Model Orde I

Setiap pengamatan $(x_{i1}, x_{i2}, \ldots, x_{ik}, y_i)$, memenuhi model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \quad (1)$$

Persamaan 1 bila ditulis dalam bentuk matriks adalah

$$y = X\beta + \varepsilon$$

dimana:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \text{ dan}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Untuk mendapatkan penduga kuadrat terkecil bagi $\hat{\beta}$, maka

$$L = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon'\varepsilon = (y - X\beta)'(y - X\beta)$$

Setelah dilakukan penyederhanaan, penduga kuadrat terkecil bagi $\hat{\beta}$ adalah

$$\hat{\beta} = (X'X)^{-1}X'y \qquad (2)$$

Penyesuaian Model Orde I

Model orde I yang disesuaikan adalah

$$\hat{y} = X\hat{\beta} \qquad (3)$$

Penyesuaian model orde I dalam notasi skalar adalah

$$\hat{y} = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \qquad i = 1, 2, \ldots, n$$

Selisih antara pengamatan y_i yang sebenarnya dan nilai dugaan $\hat{y}_i$ yang sesuai penyesuaian, adalah residual, dikatakan $e_i = y_i - \hat{y}_i$ [**3**]. Residual dinotasikan sebagai

$$e = y - \hat{y} \qquad (4)$$

Metode kuadrat terkecil menghasilkan penduga tak bias dari parameter β dalam model regresi linear berganda [**4**]. Parameter yang penting adalah jumlah kuadrat dari residual

$$SS_E = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2 = e'e \qquad (5)$$

Karena ***X'X***$\hat{\beta}$ = ***X'y***, maka dapat diperoleh formula perhitungan untuk SS_E

$$SS_E = y'y - \hat{\beta}'X'y \qquad (6)$$

Persamaan 6 disebut sebagai error atau jumlah kuadrat residual. Dapat ditunjukkan bahwa penduga tak bias dari σ^2 adalah

$$\sigma^2 = \frac{SS_E}{n-p} \qquad (7)$$

dimana n = banyaknya pengamatan dan p = banyaknya koefisien regresi. Jumlah kuadrat total adalah

$$SS_T = y'y - \frac{(\sum_{i=1}^{n} y_i)^2}{n}$$

$$SS_T = \sum_{i=1}^{n} y_i^2 - \frac{(\sum_{i=1}^{n} y_i)^2}{n} \qquad (8)$$

Maka koefisien dari ketetapan berganda R^2 didefinisikan sebagai

$$R^2 = 1 - \frac{SS_E}{SS_T} \qquad 0 \le R^2 \le 1 \qquad (9)$$

Karena R^2 selalu meningkat saat ditambahkannya variabel bebas pada model, beberapa pembangun model regresi lebih suka menggunakan statistik R^2 yang disesuaikan didefinisikan sebagai

$$R_{adj}^2 = 1 - \frac{SS_E/(n-p)}{SS_T/(n-1)} = 1 - \frac{n-1}{n-p}(1 - R^2) \qquad (10)$$

Metode *Steepest Ascent*

Metode *steepest ascent* adalah sebuah prosedur untuk memprediksi daerah optimal respon pada model orde I. Arah *steepest ascent* adalah arah dimana $\hat{y}$ meningkat dengan paling cepat. Arah ini paralel dengan normal pada respon permukaan yang disesuaikan. Garis melalui pusat daerah yang diinginkan dan normal ke permukaan yang disesuaikan diambil sepanjang lintasan *steepest ascent*. Dengan demikian, langkah-langkah sepanjang lintasan sepadan dengan koefisien-koefisien regresi ($\hat{\beta}_i$) [**3**].

Analisis Permukaan Orde II

Ketika relatif dekat pada optimum, sebuah model yang ada lengkungannya biasanya dibutuhkan untuk memperkirakan respon. Pada kebanyakan kasus, model orde II yang digunakan [**3**].

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum \sum_{i<j} \beta_{ij} x_i x_j + \varepsilon \qquad (11)$$

Lokasi Titik Stasioner

Misalkan ingin menentukan tingkat dari $x_1, x_2, \ldots, x_k$ yang mengoptimalkan respon yang diduga. Titik ini, jika ada, akan menjadi kumpulan dari $x_1, x_2, \ldots, x_k$ di mana turunan parsial $\partial\hat{y}/\partial x_1 = \partial\hat{y}/\partial x_2 = \ldots = \partial\hat{y}/\partial x_k = 0$. Titik ini, katakan $x_{1,s}, x_{2,s}, \ldots, x_{k,s}$, disebut titik stasioner. Titik stasioner dapat mewakili titik maksimum respon, titik minimum respon, atau titik pelan [**3**]. Model orde II dalam notasi matriks dapat dituliskan sebagai

$$\hat{y} = \hat{\beta}_0 + x'b + x'Bx \qquad (12)$$

dimana

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \quad b = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} \text{ dan}$$

$$B = \begin{bmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 & \ldots & \hat{\beta}_{1k}/2 \\ \hat{\beta}_{12}/2 & \hat{\beta}_{22} & \ldots & \hat{\beta}_{2k}/2 \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\beta}_{1k}/2 & \hat{\beta}_{2k}/2 & \ldots & \hat{\beta}_{kk} \end{bmatrix}$$

Turunan dari $\hat{y}$ berhubungan dengan elemen-elemen dari vektor x yang disamakan dengan 0 adalah

$$\frac{\partial\hat{y}}{\partial x} = b + 2Bx = 0 \qquad (13)$$

Titik stasioner untuk solusi persamaan tersebut adalah

$$x_s = -\tfrac{1}{2}B^{-1}b \qquad (14)$$

Selanjutnya, Persamaan (13) disubstitusikan dalam Persamaan (14), sehingga diperoleh respon yang diprediksi pada titik stasioner sebagai

$$\hat{y}_s = \hat{\beta}_0 + \tfrac{1}{2}x_s{}'b \qquad (15)$$

Karakterisasi *Response Surface*

Karakteristik *response surface* digunakan untuk menentukan jenis titik stasioner, apakah maksimum, minimum atau titik pelana. Titik stasioner dapat diidentifikasi dengan mentransformasikan fungsi respon dari titik asal $x(0, 0, ..., 0)$ ke titik stasioner x_0 dan sekaligus merotasikan sumbu koordinatnya, sehingga dihasilkan fungsi respon sebagai berikut.

$$\hat{y} = \hat{y}_s + \lambda_1 w_1^2 + \lambda_2 w_2^2 + \cdots + \lambda_k w_k^2 \qquad (16)$$

dengan w_i: variabel independen baru hasil transformasi, $\hat{y}_s$: nilai dugaan y pada titik stasioner x_0 dan λ_i: konstanta yang merupakan *eigen value* dari matriks B, $i = 1, 2, \ldots, k$ [**3**].

2. Metode Penelitian

Data yang digunakan merupakan data kecepatan reaksi untuk pengisomerasian katalis dari n-pentana ke isopentana dan banyaknya pengamatan yang digunakan sebanyak n = 24 dengan *Rate r* sebagai variable tak bebas y, dan variabel bebas x_1, x_1 dan x_3. Data tersebut diperoleh dari buku Nonlinear Regression karya G.A.F. Seber dan C.J. Wild (1989) yang diterbitkan oleh John Wiley & Sons, Inc., New York.. Untuk mempermudah dalam memproses data dan untuk memperoleh grafik *response surface* yang menarik untuk disajikan, penelitian ini menggunakan *software* R 2.15.3. Adapun langkah-langkah dalam penelitian ini adalah sebagai berikut.

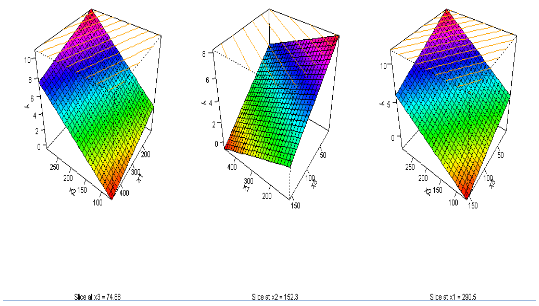
1. Menentukan sebuah pendekatan yang cocok untuk $y = f(x_1, x_2,...., x_k)$ menggunakan metode kuadrat terkecil dengan polinomial orde satu.
2. Menentukan daerah optimal dengan menggunakan metode *steepest ascent* pada model orde satu.
3. Ketika lengkungan ditemukan, selanjutnya mencari sebuah pendekatan yang baru untuk $y = f(x_1, x_2,...., x_k)$ dengan polinomial orde II.
4. Menentukan daerah optimal dengan menentukan lokasi titik stasioner pada model orde II. Melakukan analisis permukaan respon.

3. Hasil dan Pembahasan

Dari perhitungan melalui *software* R 2.15.3 diperoleh nilai $\hat{\beta}_0$ sebesar 4.117, $\hat{\beta}_1$ sebesar -0.00893, $\hat{\beta}_2$ sebesar 0.03572 dan $\hat{\beta}_3$ sebesar -0.03864. Sehingga diperoleh model orde satu yang disesuaikan sebagai berikut:

$$\hat{y} = 4.117 - 0.00893x_1 + 0.03572x_2 - 0.03864x_3$$

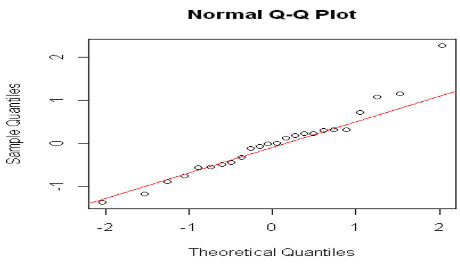
Gambar 1 menunjukkan bahwa respon permukaannya berbentuk bidang datar dan kontur plotnya paralel pada garis lurus. Pada saat variabel bebas x_1 dan x_3 mengalami peningkatan dan variabel bebas x_2 mengalami penurunan keadaan ini mengakibatkan penurunan pada nilai respon y.



Gambar 1. Kontur Plot dari Model Orde I

Pada Gambar 2, yaitu gambar uji kenormalan residual ditemukan adanya pencilan, yang menandakan bahwa residual dari model orde pertama tidak berdistribusi normal. Oleh karena itu, model orde kedua digunakan.

Jumlah kuadrat dari residual SS_E sebesar 14.45761, nilai dari penduga tak bias σ^2 sebesar 0.7228807, jumlah kuadrat total SS_T sebesar 242.6444, nilai ketetapan berganda R^2 sebesar 0.940 dan nilai R^2_{adj} sebesar 0.931.



Gambar 2. Normal Probability Plot dari Residual

Steepest Ascent

$\hat{\beta}_3$ merupakan nilai mutlak koefisien regresi terbesar dengan nilai $|-0.03864|$ sehingga $\Delta x_3 = 1$ dipilih sebagai ukuran langkah dalam metode *steepest ascent*. Ukuran langkah dalam variabel Δx_1 dan Δx_2 adalah

$$\Delta x_1 = \frac{\hat{\beta}_1}{\hat{\beta}_3/\Delta x_3} = \frac{-0.00893}{-0.03864/1} = 0.231 \quad \text{dan}$$

$$\Delta x_2 = \frac{\hat{\beta}_2}{\hat{\beta}_3/\Delta x_3} = \frac{0.03572}{-0.03864/1} = -0.9246$$

Untuk mendapatkan nilai natural dari variable-variabel bebasnya, x_1, x_2 dan x_3 digunakan langkah sebagai berikut.

$$x_1 = \left(\frac{Nilai\ Maks.x_1 - Nilai\ Min.x_1}{2}\right)\Delta x_1 + \left(\frac{Nilai\ Maks.x_1 + Nilai\ Min.x_1}{2}\right)$$

$$x_1 = \left(\frac{470.9 - 106.6}{2}\right)\Delta x_1 + \left(\frac{470.9 + 106.6}{2}\right)$$

$$x_1 = 182.15\Delta x_1 + 288.75$$

$$x_2 = \left(\frac{Nilai\ Maks.x_2 - Nilai\ Min.x_2}{2}\right)\Delta x_2 + \left(\frac{Nilai\ Maks.x_2 + Nilai\ Min.x_2}{2}\right)$$

$$x_2 = \left(\frac{294.4 - 68.3}{2}\right)\Delta x_2 + \left(\frac{294.4 + 68.3}{2}\right)$$

$$x_2 = 113.05\Delta x_2 + 181.35$$

$$x_3 = \left(\frac{Nilai\ Maks.x_3 - Nilai\ Min.x_3}{2}\right)\Delta x_3 + \left(\frac{Nilai\ Maks.x_3 + Nilai\ Min.x_3}{2}\right)$$

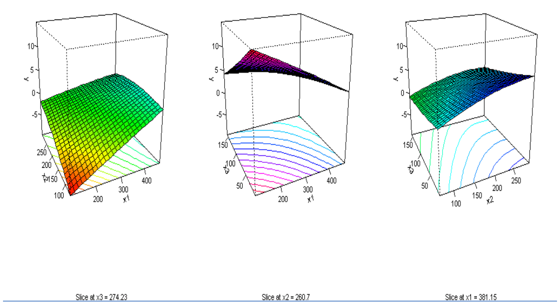
$$x_3 = \left(\frac{157.1 - 10.5}{2}\right)\Delta x_3 + \left(\frac{157.1 + 10.5}{2}\right)$$

$$x_3 = 73.3\Delta x_3 + 83.8$$

Nilai respon y pada *Steepest Ascent* disajikan pada Tabel 1.

Lokasi Titik Stasioner

Gambar 3 merupakan gambar kontur plot dari model orde II bagi dua variabel bebas. Pada gambar tersebut dapat dilihat bahwa bentuk kontur permukaan respon merupakan bentuk titik pelana.



Gambar 3. Kontur Plot dari Model Orde II

Lokasi titik stasioner dapat ditentukan dengan menggunakan Persamaan 14 dimana

$$b = \begin{bmatrix} -0.0008911 \\ 0.0914962 \\ -0.0840521 \end{bmatrix}$$ dan

$$B = \begin{bmatrix} -0.00006905 & -0.00004774 & 0.00007512 \\ -0.00004774 & -0.00008558 & -0.00002938 \\ 0.00007512 & -0.00002938 & 0.00006081 \end{bmatrix}$$

sehingga titik stasioner yang diperoleh adalah

$$x_s = -\frac{1}{2}B^{-1}b$$

$$x_s = \begin{bmatrix} 493.4524 \\ 198.3845 \\ 177.3786 \end{bmatrix}$$

Respon yang diprediksi pada titik-titik stasioner x_{1s} = 493.4524, x_{2s} = 198.3845 dan x_{3s} = 177.3786 adalah

$$\hat{y}_s = \hat{\beta}_0 + \frac{1}{2}x_s'b$$

$$\hat{y}_s = 0.5113 + \frac{1}{2}[493.4524 \quad 198.3845 \quad 177.3786] \begin{bmatrix} 0.0008911 \\ 0.0914962 \\ -0.0840521 \end{bmatrix}$$

$$\hat{y}_s = 1.912701$$

Nilai-nilai eigen λ_1, λ_2 dan λ_3 adalah akar dari persamaan determinan

$$|B - \lambda I| = 0$$

$$\begin{bmatrix} -0.00006905 - \lambda & -0.00004774 & 0.00007512 \\ -0.00004774 & -0.00008558 - \lambda & -0.00002938 \\ 0.00007512 & -0.00002938 & 0.00006081 - \lambda \end{bmatrix} = 0$$

Sehingga diperoleh persamaan karakteristik sebagai berikut.

$$-\lambda^3 - 0.00009302\lambda^2 + (5.690079 \times 10^{-9})\lambda + (3.593445 \times 10^{-13}) = 0$$

Akar-akar dari persamaan pangkat tiga tersebut adalah λ_1 = 0.0001219, λ_2 = -0.0000452 dan λ_3 = -0.0001083. Sehingga diperoleh bentuk resmi dari model yang dicocokan sebagai

$$\hat{y} = 1.912701 + 0.0001219w_1^2 - 0.0000452w_2^2 - 0.0001083w_3^2$$

Karena nilai λ_1 positif dan nilai eigen λ_2 dan λ_3 negatif maka dapat disimpulkan bahwa titik stasioner adalah titik pelana.

Tabel 1. Respon *y* pada *Steepest Ascent*

Δx_1	Δx_2	Δx_3	x_1	x_2	x_3	Respon y
0.231	-0.924	1	330.837	76.821	157.1	4.043
0.462	-1.849	2	372.924	-27.707	230.4	3.969
0.693	-2.773	3	415.011	-132.236	303.7	3.895
0.924	-3.698	4	457.098	-236.764	377.0	3.821
1.155	-4.623	5	499.185	-341.293	450.3	3.748
1.386	-5.547	6	541.272	-445.822	523.6	3.674
1.617	-6.472	7	583.360	-550.350	596.9	3.600
1.848	-7.396	8	625.447	-654.879	670.2	3.527
2.079	-8.321	9	667.534	-759.408	743.5	3.453
2.310	-9.246	10	709.621	-863.937	816.8	3.379
2.541	-10.170	11	751.708	-968.465	890.1	3.305
2.772	-11.095	12	793.795	-1072.994	963.4	3.232
3.003	-12.020	13	835.883	-1177.523	1036.7	3.158
3.234	-12.944	14	877.970	-1282.051	1110.0	3.084
3.465	-13.869	15	920.057	-1386.580	1183.3	3.010
3.696	-14.793	16	962.144	-1491.109	1256.6	2.937
3.927	-15.718	17	1004.231	-1595.637	1329.9	2.863
4.159	-16.643	18	1046.318	-1700.166	1403.2	2.789
4.390	-17.567	19	1088.405	-1804.695	1476.5	2.715
4.621	-18.492	20	1130.493	-1909.224	1549.8	2.642
4.852	-19.417	21	1172.580	-2013.752	1623.1	2.568
5.083	-20.341	22	1214.667	-2118.281	1696.4	2.494
5.314	-21.266	23	1256.754	-2222.810	1769.7	2.421
5.545	-22.190	24	1298.841	-2327.338	1843.0	2.347
5.776	-23.115	25	1340.928	-2431.867	1916.3	2.273
6.007	-24.040	26	1383.016	-2536.396	1989.6	2.199
6.238	-24.964	27	1425.103	-2640.925	2062.9	2.126
6.469	-25.889	28	1467.190	-2745.453	2136.2	2.052
6.700	-26.814	29	1509.277	-2849.982	2209.5	1.978
6.931	-27.738	30	1551.364	-2954.511	2282.8	1.904

4. Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dilakukan, dapat diperoleh beberapa kesimpulkan sebagai berikut.

1. Pada kasus regresi linear berganda dengan tiga variabel bebas menggunakan data kecepatan reaksi untuk pengisomerisasian katalis dari n-pentana ke isopentana, nilai respon *y* pada *steepest ascent* untuk semakin menurun.
2. Titik stasioner pada kasus regresi linear berganda dengan tiga variabel bebas menggunakan data kecepatan reaksi untuk pengisomerisasian katalis dari n-pentana ke isopentana berada pada titik pelana.
3. Kontribusi baru metode *response surface* terhadap perkembangan model regresi linear berganda adalah
   - Metode *response surface* dapat digunakan dalam pengoptimalisasian yang sangat penting dalam analisis regresi berganda sehingga model regresi linear berganda yang optimal dapat menentukan fungsi pendekatan yang cocok untuk memprediksi respon variabel tak bebas *y* yang akan datang.
   - Pendugaan respon *y* pada model regresi linear berganda digambarkan dengan menggunakan plot permukaan respon. Sehingga pendugaan respon *y* pada model regresi linear berganda dapat diketahui dalam bentuk permukaan respon yang berupa titik stasioner. Titik stasioner dapat mewakili titik maksimum respon, titik minimum respon, atau titik pelana.

Daftar Pustaka

[1] McClave, J.T. dan Benson, P.G. 1985. *Statistics for Business and Economics 3rd Edition*. Dellen Publishing Company, California.

[2] Walpole, R.E., Myers, R.H. dan Myers, S.L. 1998. *Probability and Statistics for Engineers and Scientist Sixth Edition*. Prentice Hall, New Jersey.

[3] Montgomery, D.C. 2001. *Design and Analysis of Experiments 5th Edition*. John Wiley & Sons, Inc., New York.

[4] Caley, K.M., Kamneva, N.Y. dan Reminga, J. 2004. *Response Surface Methodologi*. CASOS Technical Report.

[5] Seber, G.A.F. dan C.J. Wild. 1989. *Nonlinear Regression*. John Wiley & Sons, Inc., New York.