Pergamon

0039-9140(94)00149-9

Talanta, Vol. 41, No. 11, pp. 1917–1927, 1994
Copyright © 1994 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0039-9140/94 $7.00 + 0.00

FOUR-LEVEL ORTHOGONAL ARRAY DESIGN AS A CHEMOMETRIC APPROACH TO THE OPTIMIZATION OF POLAROGRAPHIC REACTION SYSTEM FOR PHOSPHORUS DETERMINATION

Wei Guang Lan,[1] Ming Keong Wong,*[1] Ni Chen[1] and Yoke Min Sin[2]
[1]Department of Chemistry, National University of Singapore, 10 Kent Ridge Crescent, Singapore 0511, Republic of Singapore
[2]Department of Zoology, National University of Singapore, 10 Kent Ridge Crescent, Singapore 0511, Republic of Singapore

(*Received* 4 *January* 1994 *Revised* 11 *May* 1994. *Accepted* 14 *May* 1994)

Summary—It is the purpose of the present work to provide information on the four-level orthogonal array design and data analysis for the optimization of analytical procedures. In the theoretical part, the construction and characteristics of the $OA_{16}(4^5)$ matrix is described in detail, followed by the data analysis strategy, in which the significance of the different factors is quantitatively evaluated by an analysis of variance (ANOVA) method including per cent contribution, and the difference among four levels for each factor is determined by Duncan's multiple F test. Furthermore, a third-order polynomial model representing response surface is established to estimate the effects for the factors with significant influences. In the application part, the proposed four-level orthogonal array design and data analysis method were applied to optimize polarographic reaction system for phosphorus determination. By conducting 16 preplanned experiments that span the maximum working range of the system, the best experimental conditions for achieving the largest response can be obtained. The expected value for each experimental trial calculated by the third-order regression equation established is in good agreement with the corresponding experimental value. To confirm the validity of the optimization procedure, additional experiments using the recommended conditions were performed. The results demonstrate that satisfactory results can be acquired. Therefore, the proposed four-level orthogonal array design as a chemometric approach to optimize the polarographic reaction system for phosphorus determination is rather efficient and effective.

β_0	constant item of the polynomial model, where input variable is ϕ_X
β_X	coefficient of the first-order effect for the factor X, where input variable is ϕ_X
β_{XX}	coefficient of the second-order effect for the factor X, where input variable is ϕ_X
β_{XXX}	coefficient of the third-order effect for the factor X, where input variable is ϕ_X
ϵ	standard error
ϕ_X	function of the Z for factor X. When Z is specified as Z_k (level value at level k), ϕ_X is expressed as ϕ_{Xk}
ρ	Duncan's coefficient
df	degree of freedom
E_X	effect of the factor X, when X is at level k, E_X is expressed as E_{Xk}
F	random variable of F distribution; ratio of variance
H_x	difference between the level value at level k and that at level $k+1$, where $k = 1, 2, 3$.
i	number of the experimental trials, $i = 1, 2, \ldots, 16$
j	number of the observed responses, $j = 1, 2, \ldots, J$
k	number of the level settings, unless otherwise specified, $k = 1, 2, 3, 4$
l	number of the replicated times for the different experimental trials at the same level setting, $l = -1, 2, 3, 4$
MS_x	mean square due to the factor X
MS_{error}	mean square due to error
PC	per cent contribution
r_k	average of the responses of the replicated experimental trials for a factor at level k, when specifying the factor X, r_k is expressed as r_{xk}
SP	span between the level means
SS_{total}	sum of squares due to total variance

---

*Author to whom correspondence should be addressed.

SS_x	sum of squares due to the factor X
SS_{error}	sum of squares due to error
SS'	purified sum of squares
SSR	shortest significant ranges
X	factor of interest
X_s	factor with a significant influence
y	individual response
y_{ij}	individual response at the *i*th row and the *j*th observation
$\bar{y}$	average of the responses
$\bar{y}_i$	average of the responses at the *i*th row
$\hat{y}$	polynomial value including β_0 item
Y	polynomial value excluding β_0 item
$\bar{Y}$	average of the polynomial value excluding β_0 item
Z_X	any value between Z_{X1} and Z_{X4} for factor X, when X is at level *k*, Z_X is expressed as Z_{Xk}
$\bar{Z}_X$	average of the sum of the Z_{X1}, Z_{X2}, Z_{X3} and Z_{X4} for factor X

In HCl–Sb(III)–Mo(VI)–ketone medium, PO_4^{3-}, Sb(III) and Mo(VI) form a ternary heteropoly acid, which can be adsorbed on a dropping mercury electrode and reduced to a heteropoly blue. The electrochemical reaction produces a sensitive polarographic absorptive wave (at peak potential of −0.41 V *vs* SCE), which were widely applied to the determination of phosphorus in sea-water, marine sediments and biological tissue.[1-3] Recently, the optimum conditions using this method for phosphorus determination was reported by Chen,[4] in which the so-called one-factor-at-a-time method was employed. However, it is nowadays commonly recognized that the one-factor-at-a-time method is often unreliable and prone to obtaining false optimum.[5-7]As an illustration, let us consider an analytical method in which several factors need to be optimized. In the case of one-factor-at-a-time method, all factors except one are kept constant, that factor is varied, and the level setting at which it yields the optimum response is ascertained. This procedure is then repeated as required for all other factors. However, regardless of which factor being studied, the optimum level value obtained for such experiments is true only if those other conditions are kept identical. Such requirements might be impossible to achieve. Thus, it is desirable that the parameters used be optimized by means of chemometric approach.

We report the use of four-level orthogonal array design as a means of optimizing the medium compositions of the polarographic reaction system for phosphorus determination in the present work. As a chemometric approach, orthogonal array design with different level settings has been introduced and applied in our previous works.[8-11] The precursor of this method was the Latin square design, which was based on the Latin square and Graeco-Latin square.[12] Probably Rao was the first to develop a general theory of orthogonal array design,[13-15] though the term 'orthogonal array' was first coined by Bose and Bush in 1952.[16] However, it did not capture the imagination of English-speaking researchers until this method was successfully applied to engineering areas for quality control by G. Taguchi, a Japanese engineer.[17] For this reason, in western countries, orthogonal array design is often referred to as Taguchi design and/or method.[18-26] However, it should be pointed out that the method of application of orthogonal array design have been independently put forward and developed in China for a long time and a large number of Chinese publications are available.[27-31] In addition, in 1991, Kacker *et al.* critically pointed out that Taguchi's orthogonal arrays are classical designs of experiments.[32,33] Nevertheless, the upsurge of interest in Taguchi's work has led to orthogonal array design being widely applied to engineering areas in the United States and United Kingdom since their introduction in the early 1980s and the mid-1980s, respectively.[16-24]

THEORETICAL

Matrix

For any orthogonal array design, a matrix, which consists of columns and rows, with various numbers at the intersections of each column and row, must be constructed. Table 1 displays an $OA_{16}(4^5)$ matrix. The constructing method has been given in Ref. 11. This is a four-level orthogonal array matrix which is made up of five columns and 16 rows. Each column represents a factor, which is an independent variable, and each row represents an experimental trial. The numbers at the intersections indicate the level settings that apply to the factors for the experimental trials. From this matrix it can be

Table 1. The $OA_{16}(4^5)$ matrix associated with the analytical results

Trial No.	Column No. 1	2	3	4	5	Experimental $\bar{y} \pm SD\,(n=3)$	Polynomial $Y^{a/*}$	$\hat{y}$†	$(\bar{y}-\hat{y})$
1	1	1	1	1	1	6.36 ± 0.37	−5.84	6.43	−0.07
2	1	2	2	2	2	10.28 ± 0.58	−2.23	10.04	0.24
3	1	3	3	3	3	9.69 ± 0.49	−2.59	9.68	0.01
4	1	4	4	4	4	5.21 ± 0.21	−6.90	5.37	−0.16
5	2	1	2	3	4	12.63 ± 0.72	0.50	12.77	−0.14
6	2	2	1	4	3	11.49 ± 0.49	−0.77	11.50	−0.01
7	2	3	4	1	2	8.08 ± 0.52	−4.45	7.82	0.26
8	2	4	3	2	1	8.92 ± 0.63	−3.24	9.03	−0.11
9	3	1	3	4	2	10.76 ± 0.56	−1.78	10.49	0.27
10	3	2	4	3	1	8.41 ± 0.48	−3.77	8.50	−0.09
11	3	3	1	2	4	9.05 ± 0.35	−3.05	9.22	−0.17
12	3	4	2	1	3	7.81 ± 0.27	−4.48	7.79	0.02
13	4	1	4	2	3	7.06 ± 0.38	−5.24	7.03	0.03
14	4	2	3	1	4	8.48 ± 0.47	−3.63	8.64	−0.16
15	4	3	2	4	1	11.90 ± 0.58	−0.23	12.04	−0.14
16	4	4	1	3	2	7.82 ± 0.39	−4.62	7.65	0.17
r_1	7.88	9.20	8.68	7.68	8.90				
r_2	10.28	9.66	10.66	8.83	9.24	$u \pm SD = 9.00 \pm 2.04$			
r_3	9.01	9.68	9.46	9.64	9.01				
r_4	8.82	7.44	7.19	9.84	8.84	$\bar{y} \pm SD = -3.27 \pm 2.04$			
E_1	−1.92	−0.67	−1.62	−1.63	−0.24	$\bar{\hat{y}} \pm SD = 9.00 \pm 2.04$			
E_2	0.48	−0.18	0.34	−0.47	0.09				
E_3	−0.80	−0.16	−0.86	0.35	−0.15	$\overline{Y-\hat{y}} \pm SD = 0.00 \pm 0.16$			
E_4	−0.96	−2.39	−3.14	0.55	−0.34				

*Polynomial values excluding β_0 item.

†Polynomial values including β_0 item.

noted that each of five columns is varied over four level settings, each level setting repeats four times, and thus a total of 4 × 4 = 16 experimental trials are necessary for each column. Furthermore, in any two columns, the horizontal combinations of any two level numbers appear at the same number of times. That is, each combination of the 16 ordered pairs (1, 1), (1, 2), (1, 3), (1, 4); (2, 1), (2, 2), (2, 3), (2, 4); (3, 1), (3, 2), (3, 3), (3, 4) and (4, 1), (4, 2), (4, 3), (4, 4) appears exactly once. Thus when the four level numbers of a column are at level 1, for any other columns, the corresponding level numbers are at level 1, 2, 3 and 4, respectively. Similar cases can be seen when the four level numbers of this column are at other three level settings. The above features of the $OA_{16}(4^5)$ matrix provide the orthogonality among all the five columns. Orthogonality means that when the effect for the factor X is calculated the influence of the factors except for factor X cancel out, and thus the effect for the factor assigned to each column in the $OA_{16}(4^5)$ matrix can be estimated independently of one another. It can be proven by a statistical method which has been described in detail elsewhere.[11] Therefore, in the case of all the interactions being neglected, an $OA_{16}(4^5)$ design which has only 16 trials can represent the possible $4^5 = 1{,}024$ combinations. The latter number is so large that it is difficult to carry out these experiments in practice.

Assignment of experiments

In fact, the $OA_{16}(4^5)$ matrix can be viewed as a $16/4^5 = 1/64$ fraction of a full 4^5 factorial design. In this matrix, the 16 experimental trials provide a total of 15 degrees of freedom for the entire experiment, owing to each column having four level settings, *i.e.* three degrees of freedom, and thus five columns can be allocated to assign factors. In principle, one column can be assigned a factor. However, in practice, if possible, it is desirable that at least one column be used to assign a dummy factor, in which no variable (actual factor) is assigned so as to measure the error variance. Otherwise, the error variance must come from the repetitions in each experimental trial.

Data analysis strategy

After conducting four-level orthogonal array design and obtaining data of responses, for each factor, the four level means (r_1, r_2, r_3 and r_4) must be first calculated. Then, the ANOVA method is employed to estimate the effects for the factors. Table 2 shows the ANOVA

Table 2. ANOVA formulae including per cent contribution for the $OA_{16}(4^5)$ matrix

Source of variance	Sum of square (SS)	Degree of freedom (*df*)	Mean square (MS)	*F* value	Purified sum of square (SS′)	Per cent contribution (PC)
X	$4J\sum_{k=1}^{4}(r_{xk}-u)^2$	3	$\frac{SS_X}{3}$	$\frac{MS_X}{MS_{error}}$	$SS_X - 3MS_{error}$	$\frac{SS'_x}{SS_{total}} \times 100\%$
Error	$\sum_{i=1}^{16}\sum_{j=1}^{J}(y_{ij}-\bar{y}_i)^2$	$16(J-1)$	$\frac{SS_{error}}{16(J-1)}$		$SS_{total} - \sum SS'_x$	
Total	$\sum_{i=1}^{16}(y_{ij}-u)^2$	$16J-1$				

formulae for testing the magnitude of the different factors in the $OA_{16}(4^5)$ matrix. The calculation of those sums of squares in Table 2 has been discussed in detail elsewhere.[10,11]

When implementing ANOVA, a term called the 'purified sum of squares (SS′)', which is the sum of squares minus the variance due to error, and a term called the 'per cent contribution' which is the relative contribution of SS′ for each factor, or error to the total variance can be computed. Their calculation formulae have been included in Table 2. The per cent contribution due to factor X provides an estimate of the importance of the factor considered, and the per cent contribution due to error provides an estimate of the adequacy of the experiment parameters.[18]

The ANOVA method including per cent contribution in Table 2 provides a significant test for the different factor effects by comparing statistics F_x to F_α, which is a critical value at the given degrees of freedom and significant level. However, it does not specify which of the level means is different. Thus, we recommend employing the Duncan's multiple *F* test[34] to compare the difference among level means for those factors with significant effects. In this case, the application of Duncan's multiple *F* test is briefly described as follows: (i) After carrying out ANOVA, for the factor with a significant effect, the four level means are in order from small to large. (ii) According to the degrees of freedom for error [$16(J-1)$, refer to Table 2] and the spans between the level means (SP), the Duncan's coefficient (ρ) is extracted from Duncan's multiple *F* test coefficient tables[34] at the different significant levels. (iii) The shortest significant ranges (SSR) are calculated according to the following formula:

$$\mathrm{SSR} = \rho \times \sqrt{\frac{\mathrm{MS}_{error}}{4J}}, \qquad (1)$$

where $4J$ represents the number of observations in each level. (iv) The difference between two level means is compared with the corresponding SSR, which is calculated by equation (1). If the former exceeds the latter, the difference is significant at the corresponding significant level. Otherwise, it is not significant.

Furthermore, if one expects to estimate the effects for the factors in the model, the following polynomial can be established:

$$y = \beta_0 + \sum_{x_s}\beta_{x_s}\phi_{x_s} + \sum_{x_s}\beta_{x_sx_s}\phi^2_{x_s} + \sum_{x_s}\beta_{x_sx_sx_s}\phi^3_{x_s} + \epsilon. \qquad (2)$$

The polynomial model given in equation (2) excludes the factors without significant influences because the contribution of the nonsignificant factors to responses is trivial and thus can be incorporated into the ϵ item. For the assumptions of the above model and the estimate of the different regression coefficients, refer to Appendix A.

EXPERIMENTAL

Apparatus and reagents

A model JP-1A oscillopolarographic analyzer (China) was used. Dropping mercury electrode and saturated calomel electrode (SCE) were employed as the working and reference electrodes, respectively. The auxiliary electrode was made of platinum.

Standard stock phosphorus solution (1000 $\mu g/cm^3$ P) was prepared from potassium dihydrogen phosphate; further standards were obtained by diluting this solution as required. Hydrochloric acid (5*M*), potassium antimony tartrate solution (0.002*M*), ammonium heptamolybate tetrahydrate solution (0.1*M*), acetone and butanone were used.

Analytical procedures

A cm^3 ammonium heptamolybate tetrahydrate solution, B cm^3 of potassium antimony

Table 3. Assignment of factors and their level values to the polarographic reaction system for phosphorus determination in the $OA_{16}(4^5)$ matrix

	Column No.				
	1	2	3	4	5
	Factor*				
Level†	A	B	C	D	E
1	0.50	2.00	0.50	0.00	3.00
2	1.00	1.50	1.00	1.00	2.00
3	1.50	1.00	1.50	2.00	1.00
4	2.00	0.50	2.00	3.00	0.00

*A: 0.1*M* of ammonium heptamolybate tetrahydrate solution (cm^3); B: 0.002*M* of potassium antimony tartrate solution (cm^3); C: 5*M* of hydrochloric acid (cm^3); D: acetone (cm^3); E: butanone (cm^3).

†The level value for each factor is (i) satisfied with the definition given in equation (A3); (ii) randomly arranged from low to high or *vice versa*.

Table 5. Worksheet of the Duncan's multiple *F* test for the significant factors at different levels*

		SP (SSR at $P = 0.05$)		
		2 (0.40)	3 (0.42)	4 (0.44)
A	r_1 (7.88)	$\underline{r_4\ (8.82)}$	$\underline{r_3\ (9.01)}$	r_2 (10.28)†
B	r_4 (7.44)	$\underline{r_1\ (9.20)}$	$\underline{r_2\ (9.66)}$	$\underline{r_3\ (9.68)}$
C	r_4 (7.19)	r_1 (8.68)	r_3 (9.46)	r_2 (10.66)†
D	r_1 (7.68)	r_2 (8.83)	$\underline{r_3\ (9.64)}$	$\underline{r_4\ (9.84)}$

*Any two level means not underscored by the same line are significantly different ($P < 0.05$). On the contrary, any two level means underscored by the same line are not significantly different ($P > 0.05$).

†Significantly different from the other level means at $P < 0.01$, where $SSR_{(2,32,0.01)}$ is 0.54.

tartrate solution, and C cm^3 of hydrochloric acid were added to a 25 cm^3 volumetric flask and mixed well. Then D cm^3 acetone with E cm^3 of butanone and 5 cm^3 of standard phosphorus solution (100 $\mu g/dm^3$) were added and mixed thoroughly for different experimental trials, the factors (A–E) were varied according to the level setting numbers shown in Table 3). Afterwards, the solution was diluted with de-ionized water to the mark, mixed well and an aliquot of the solution in volumetric flask was transferred to the polarographic cell. The polarogram from -0.30 to -0.80 V was recorded at a scan-rate of 250 mV/sec with an oscillopolarographic analyzer. The peak potential of the derivative curve was at -0.41 V *vs* SCE, at which the peak currents of the different experimental trials $[i_p(\mu A)]$ were recorded and calculated as the response (y) for the four-level orthogonal array design.

RESULT AND DISCUSSION

On the basis of Refs 1–4 five factors were considered in the present work: (i) the volume of 0.1*M* ammonium heptamolybate tetrahydrate solution (A); (ii) the volume of 0.002*M* potassium antimony tartrate solution (B); (iii) the volume of 5*M* hydrochloric acid (C); (iv) the volume of acetone (D) and (v) the volume of butanone (E). The position and four different level setting values of five factors in the $OA_{16}(4^5)$ matrix are given in Table 3. Since no dummies were available, triplicate runs were carried out for each experimental trial to measure the error variance.

After conducting all of the experiments, the mean of the response values ($\bar{y}$) and their standard deviation (SD) obtained are given in Table 1. Using the formulae shown in Table 2, the ANOVA table including per cent contribution was constructed and given in Table 4. From Table 4 it is clear that all the factors except for E are statistically significant at $P < 0.001$, whereas no statistical difference is observed for the factor E ($P > 0.1$). The latter demonstrates that the peak current for the polarographic adsorptive wave of the PO_4^{3-}–Sb(III)–Mo(VI) ternary heteropoly acid is not affected by the volume of butanone. In other words, the addition of the butanone does not enhance response value considered. Therefore, it is actually not necessary to add butanone in the polarographic reaction system for phosphorus determination.

Table 4. An ANOVA table including per cent contribution for experimental responses in the $OA_{16}(4^5)$ matrix

Source	SS	*df*	MS	*F**	SS′	PC(%)
Mo(VI) (A)	35.10	3	11.70	48.75***	34.38	17.63
Sb(III) (B)	40.46	3	13.49	56.21***	39.74	20.38
HCl (C)	76.15	3	25.38	105.75***	75.43	38.68
Acetone (D)	34.64	3	11.55	48.12***	33.92	17.39
Butanone (E)	1.12	3	0.37	1.54	0.40	0.20
Error	7.55	32	0.24		11.15	5.72
Total	195.02	47			195.02	100.00

*Critical value is 6.96(***$P < 0.001$) and 2.27 ($P < 0.1$).

Moreover, from per cent contribution, it can be seen that the most significant effect contributing to the response is C (38.68%), then in order is B (20.38%), A (17.63%), D (17.39%) and E (0.20%). The per cent contribution due to error (unknown and uncontrolled factors) is relatively low (5.72%). This means that no important factors and/or interaction between factors are omitted, so it is reasonable to neglect the interaction between factors in this work.[18]

For those factors with significant influences, the Duncan's multiple F test[34] was implemented and the results obtained are shown in Table 5. From Table 5 it can be seen that (i) the best level means for factors A and C are at level 2, at which the level mean for either factor A or C is significantly different from that at other levels ($P < 0.01$); (ii) the best level mean for factor B is at level 3, but no significant difference is observed among level 1, 2 and 3 ($P > 0.05$) and (iii) the best level mean for factor D is at level 4, but no significant difference is observed between level 3 and 4 ($P > 0.05$).

The above conclusions mean that the recommended conditions for the polarographic reaction system for phosphorus determination are 1.00 cm^3 of ammonium heptamolybate tetrahydrate solution ($0.1M$); 1.00–2.00 cm^3 of potassium antimony tartrate solution ($0.002M$); 1.00 cm^3 of hydrochloric acid ($5M$) and 2.00–3.00 cm^3 of acetone. No butanone is necessary.

Furthermore, according to Appendix A, the effect for each factor at four different levels (E_1, E_2, E_3 and E_4) is calculated and given in Table 1. From the maximum difference between E_k and $E_{k'}$, it is clear that the most important factor is the factor C (3.48), then in order is A (2.40), B (2.23), D (2.18), and E (0.43). These results are similar to those obtained from the ANOVA table including per cent contribution.

By substituting various E_k values given in Table 1 into equation (2), the polynomial value excluding β_0 item (Y) for each experimental trial could be computed and the figures given in Table 1. It should be noted that since the effect for the factor E is relatively small and no statistical difference for this factor is observed ($P > 0.1$), it should be incorporated into the error item and thus not be included in the Y item. In practice, the mean of the difference between $\bar{y}$ and Y [$(\overline{\bar{y} - Y})$, $i = 16$; see Table 1] can be considered as the β_0 item.

Therefore, the following third-order polynomial model containing four significant factors can be generated:

$$y = 12.27 - 1.48\phi_A - 0.64\phi_A^2 + 0.80\phi_A^3 + 0.10\phi_B - 0.68\phi_B^2 - 0.30\phi_B^3 - 1.29\phi_C - 1.06\phi_C^2 + 0.35\phi_C^3 + 0.82\phi_D - 0.24\phi_D^2 - 0.04\phi_D^3 + \epsilon. \quad (3)$$

For each experimental trial, by substituting the ϕ_k values [given in equation (A5) in Appendix

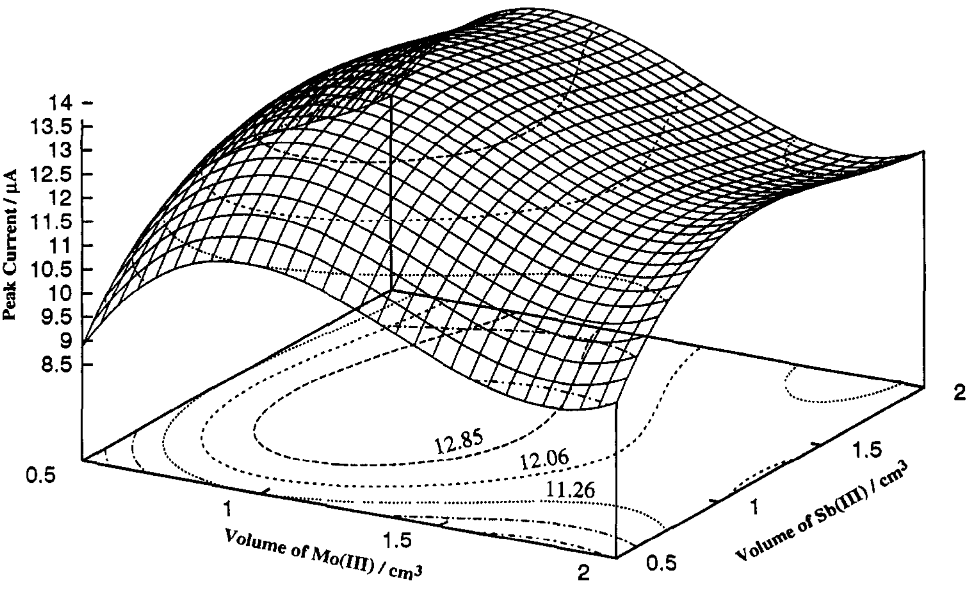

Fig. 1. Response surface for the effect of $0.1M$ of Mo(VI) and $0.002M$ of Sb(III) on the peak current of phosphorus at 1.00 cm^3 of $5M$ of HCl and 2.79 cm^3 of acetone.

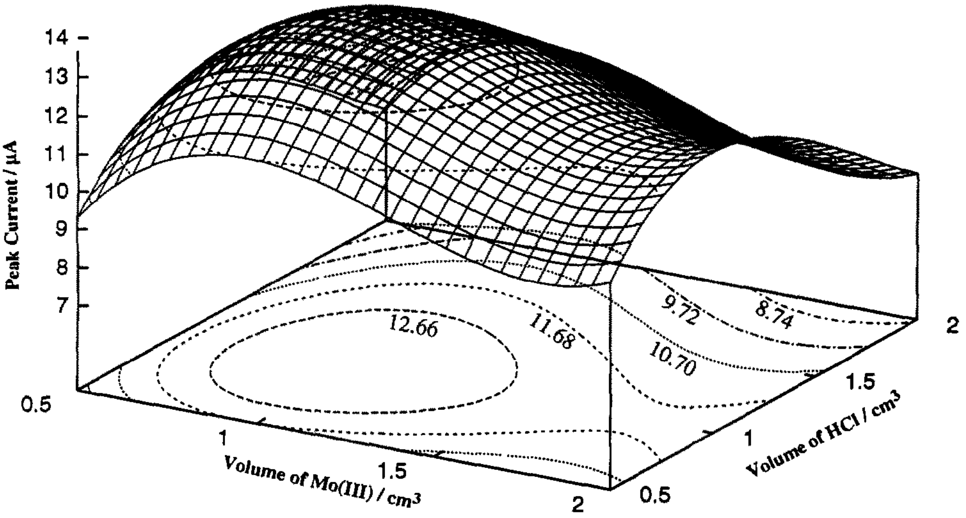

Fig. 2. Response surface for the effect of 0.1*M* of Mo(VI) and 5*M* of HCl on the peak current of phosphorus at 1.21 cm³ of 0.002*M* of Sb(III) and 2.79 cm³ of acetone.

A, it should be noted that the level settings (k) for different factors are varied with the level numbers of the intersection as shown in Table 1] into equation (2), the expected value ($\hat{y}$) [polynomial value including β_0 item] can be calculated and the figures are given in Table 1. Correspondingly, the random error item ($\epsilon = \bar{y} - \hat{y}$) for each experimental trial is also calculated and given in Table 1. The results obtained show that the expected value for each experimental trial is in good agreement with the corresponding experimental value. The mean of the error item [$(\overline{\bar{y} - \hat{y}})$, $i = 16$] equals zero (see Table 1), which indicates that the ϵ item is an independent random variable from an $N(0, \sigma^2)$ distribution. Therefore, the third-order regression equation given in equation (3) can adequately and accurately represent the described response surface.

Furthermore, in order to precisely express the quantitative relationship between output response (y) and input variable values (Z_x), equation (A2) (shown in Appendix A) with the known $\bar{Z}_x$ and H_x values should be substituted

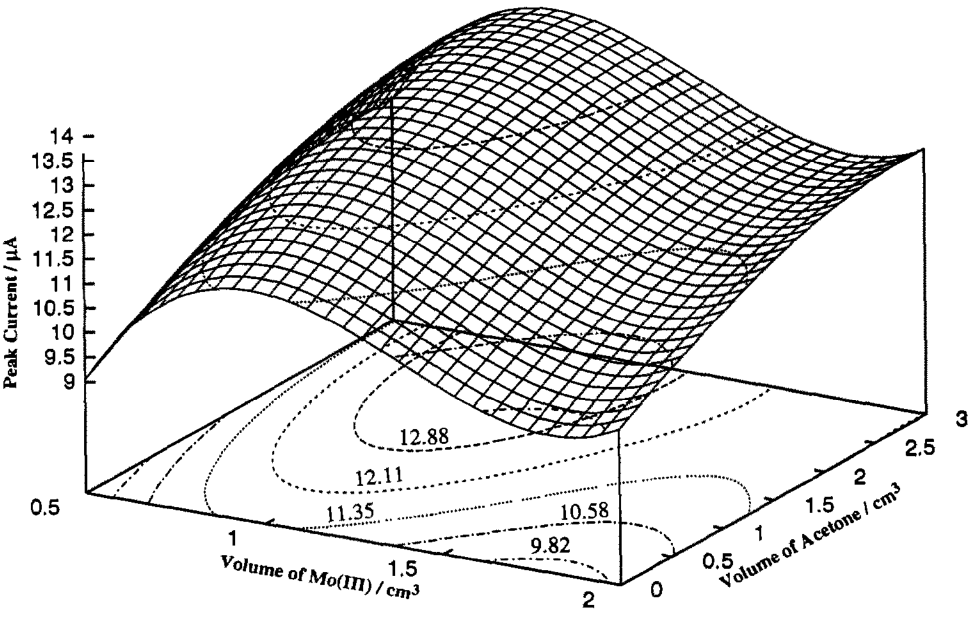

Fig. 3. Response surface for the effect of 0.1*M* of Mo(VI) and acetone on the peak current of phosphorus at 1.00 cm³ of 5*M* of HCl and 1.21 cm³ of 0.002*M* of Sb(III).

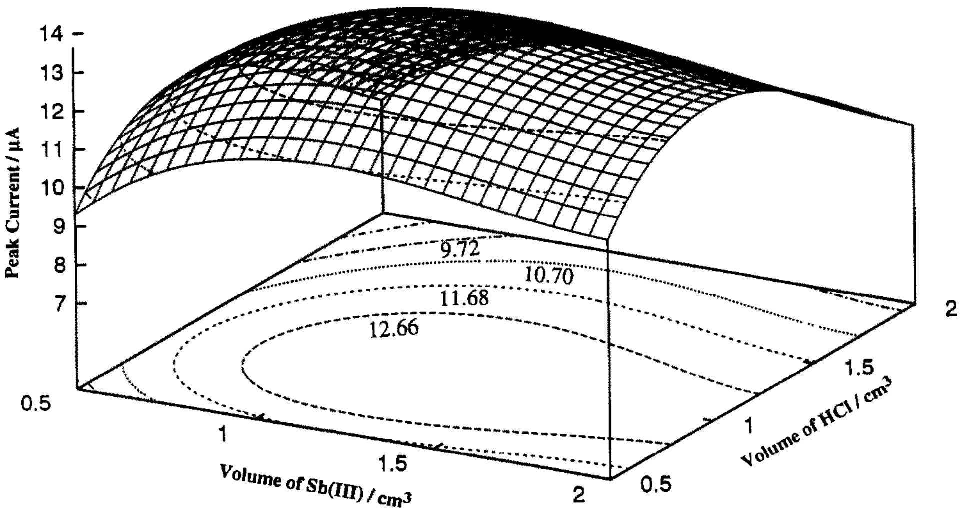

Fig. 4. Response surface for the effect of 0.002*M* of Sb(III) and 5*M* of HCl on the peak current of phosphorus at 0.96 cm³ of 0.1*M* of Mo(VI) and 2.79 cm³ of acetone.

into equation (3), then the following equation can be concluded

$$y = -19.65 + 33.39Z_A + 17.83Z_B + 21.12Z_C + 1.27Z_D - 26.56Z_A^2 - 11.72Z_B^2 - 14.74Z_C^2 - 0.06Z_D^2 + 6.40Z_A^3 + 2.40Z_B^3 + 2.80Z_C^3 - 0.04Z_D^3 + \epsilon. \quad (4)$$

The response surface represented are displayed in Figs 1–6. Moreover, according to the derivative algorithm given in Ref. 11 the following optimum conditions (i) 0.96 cm³ of ammonium heptamolybate tetrahydrate solution (0.1*M*); (ii) 1.21 cm³ of potassium antimony tartrate solution (0.002*M*); (iii) 1.00 cm³ of hydrochloric acid (5*M*) and (iv) 2.79 cm³ of acetone can be obtained. It is obvious that the above results are similar to those obtained from applying Duncan's multiple *F*-test.

Lastly, to confirm the validity of the optimization procedure, additional experiments using the experimental conditions recommended by Duncan's multiple *F*-test (A_2, B_3, C_2, D_4 and E_4) were performed. The result demonstrated that when compared with that results obtained from 16 experimental trials in the $OA_{16}(4^5)$

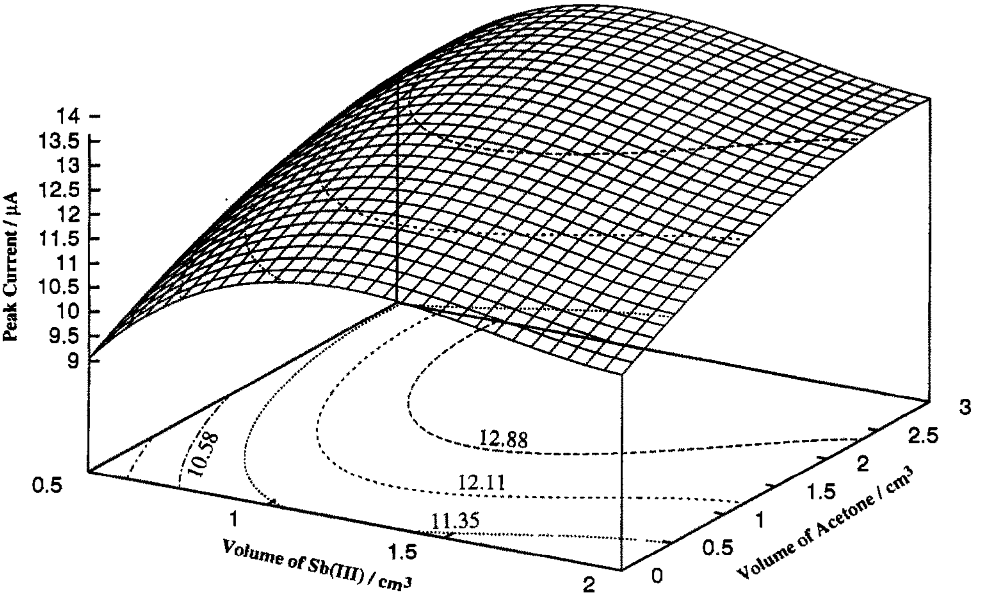

Fig. 5. Response surface for the effect of 0.002*M* of Sb(III) and acetone on the peak current of phosphorus at 0.96 cm³ of 0.1*M* of Mo(VI) and 1.00 cm³ of 5*M* of HCl.

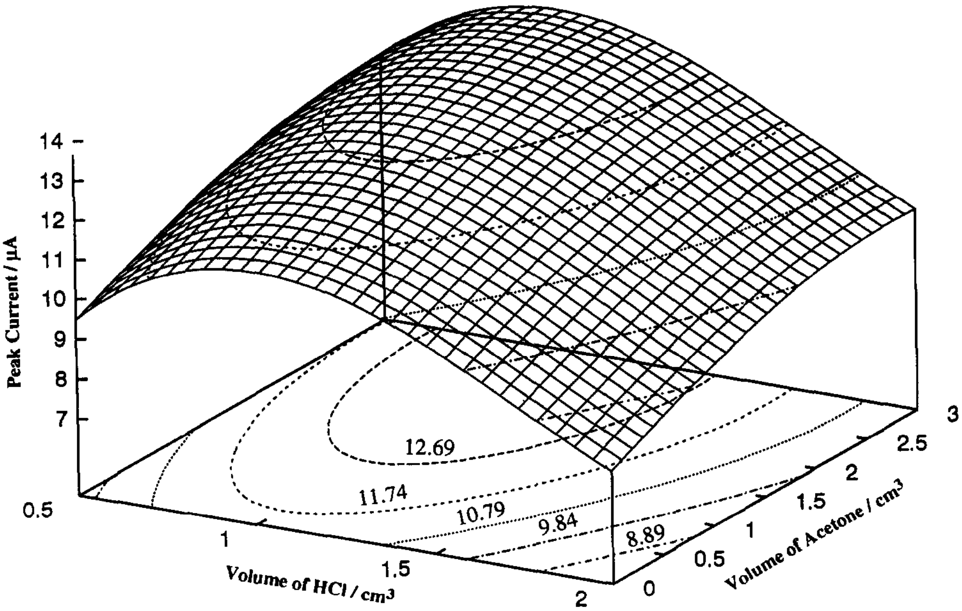

Fig. 6. Response surface for the effect of 5*M* of HCl and acetone on the peak current of phosphorus at 0.96 cm^3 of 0.1*M* of Mo(VI) and 1.21 cm^3 of 0.002*M* of Sb(III).

matrix, the largest response value (13.72 ± 0.75, $n = 3$) could be achieved. On the other hand, the theoretical value predicted by using equation (4) according to the best experimental conditions obtained from Duncan's multiple *F*-test and the optimum conditions obtained from derivative algorithm are 13.48 and 13.62, respectively. Both data are in good agreement with the mean of the former experimental value (13.72).

CONCLUSIONS

Four-level orthogonal array design with the $OA_{16}(4^5)$ matrix provides a systematic procedure which could be used to obtain the overall optimum within the range of experimental conditions investigated. When interaction effects can be neglected, owing to the orthogonality of the experimental design, the preplanned 16 experimental trials can yield the same information as that obtained from the corresponding full factorial design which requires $4^5 = 1{,}024$ experimental trials. By using the ANOVA technique including per cent contribution and the Duncan's multiple *F* test, a large amount of quantitative information on the significance of a factor and the control of the optimum range for a significant factor can be obtained. Furthermore, by means of the formulae given in Appendix A, the third-order regression equation representing the response surface can be easily established and thus the effects for the significant factors can be quantitatively estimated in the mathematical model.

Either polynomial model established with derivative algorithm or Duncan's multiple *F* test can be used to decide the best/optimum experimental condition for each variable considered. However, the computational burden of the former method is much larger than that of the latter one. Indeed, by using the former method, the optimum value can be precisely obtained. But in many cases it may not always be necessary to choose the best experimental conditions according to the optimum value acquired because the departure from the optimum value obtained may not always cause the significant variation. For instance, in the present study, when the volume of potassium antimony tartrate solution changed from 1.0 to 2.0 ml, according to equation (4), it can be calculated that for output response only 0.53 can be reduced. It is obvious that the variation of 0.53 can be neglected. For that reason, if no quantitative relationship between output response and input variable values is acquired, we recommend using Duncan's multiple *F* test instead of the polynomial model with derivative algorithm for estimating the best experimental conditions.

Although four-level orthogonal array design was employed to optimize the polarographic reaction system for phosphorus determination in the present work, as a chemometric optimization approach, four-level orthogonal array design has much broader potential applicability. It can be applied to the optimization of nearly all the analytical procedures such as the conditions for various spectrometric determination and flow injection methods, the reaction system of spectrophotometric and polarographic techniques, operation parameters of gas chromatographic and high performance liquid chromatographic analysis, and so forth.

Acknowledgements—We would like to thank Dr J. Y. Guo (Department of Zoology, National University of Singapore) for his technical assistance to display a response surface. One of us (W. G. Lan) wishes to thank the National University of Singapore for the award of a studentship to read for a PhD. degree. This work was supported by a grant from the National University of Singapore (RP920634).

REFERENCES

1. W. G. Lan, N. Chen and S. K. Yang, *Acta Oceanolog. Sinica* (in Chinese), 1990, **12,** 596.
2. W. G. Lan, N. Chen and S. K. Yang, *Marine Environm. Sci.* (in Chinese), 1990, **9,** 101.
3. W. G. Lan, *J. Xiamen Fish. College* (in Chinese), 1990, **12,** 21.
4. L. Chen, *Talanta*, 1992, **39,** 765.
5. H. Van Ryswyk and G. R. Van Hecke, *J. Chem. Educ.*, 1991, **68,** 878.
6. S. Wold, M. Sjostrom, R. Carlson, T. Lundstedt, S. Hellberg, B. Skagerberg, C. Wikstrom and J. Ohman, *Anal. Chim. Acta*, 1986, **191,** 17.
7. O. Grossmann, *Anal. Chim. Acta*, 1986, **186,** 185.
8. H. B. Wan, W. G. Lan, M. K. Wong and C. Y. Mok, *Anal. Chim. Acta*, 1994, **289,** 371.
9. H. B. Wan, W. G. Lan, M. K. Wong, C. Y. Mok and Y. H. Poh, *J. Chromatogr.*, 1994, in press.
10. W. G. Lan, M. K. Wong, N. Chen and Y. M. Sin, *Analyst*, 1994, in press.
11. W. G. Lan, M. K. Wong, N. Chen and Y. M. Sin, *Analyst*, 1994, in press.
12. G. C. Tarry, *R. Ass. Franc. Av. Sci.*, 1901, **1,** 122; 1901, **2,** 170.
13. C. R. Rao, *Bull. Calcutta Math. Soc.*, 1946, **38,** 67.
14. C. R. Rao, *J. Roy. Statist. Soc. Suppl.*, 1947, **9,** 128.
15. C. R. Rao, *Proc. Edinburgh Math. Soc.*, 1947, **8,** 119.
16. R. C. Bose and K. A. Bush, *Ann. Math. Statist.*, 1952, **23,** 508.
17. G. Taguchi, *System of Experimental Design*, Vols 1–2, Kraus, New York, 1987.
18. P. J. Ross, *Taguchi Techniques for Quality Engineering-Loss Function, Orthogonal Experiments, Parameter and Tolerance Design.* McGraw-Hill, New York, 1988.
19. A. Bendell, J. Disney and W. A. Pridmore (eds), *Taguchi Methods*—Applications in World Industry. IFS/Springer, Bedford (U.K.)/Berlin, 1989.
20. American Supplier Institute, *Taguchi Methods—Selected Papers on Methodology and Applications.* ACI Press, Dearborn, Michigan, 1988.
21. W. Charteris, *J. Soc. Dairy Technol.*, 1992, **45,** 33.
22. R. H. Lochner and J. E. Matar, *Design for Quality—An Introduction to the Best of Taguchi and Western Methods of Statistical Experimental Design.* Quality Resources, New York, 1990.
23. T. B. Barker, *Engineering Quality by Design—Interpreting the Taguchi Approach.* Marcel Dekker, New York, 1990.
24. T. Bendell (ed.), *Taguchi Methods—Proceedings of the 1988 European Conference.* Elsevier, London, 1989.
25. P. J. Oles and A. Yankovich, *LC-GC*, 1989, **7,** 579.
26. P. J. Oles, *J. Assoc. Off. Anal. Chem.*, 1993, **76,** 615.
27. Mathematics Department, Beijing University, *Orthogonal Design* (in Chinese). Beijing University Press, Beijing, 1976.
28. Y. Y. Wu, *Experimental Design* (in Chinese). Chinese Productivity Center, Taiwan, 1974.
29. Chemistry Department, Beijing University (Analytical Chemistry Group Report), *Fenxi Huaxue*, 1976, **4,** 9.
30. B. Li, X. Zhang, X. Chen and X. Zhou, *Spectroscopy and Spectral Analysis* (in Chinese), 1989, **9,** 52.
31. Editorial Board for Statistics and Management, *The Proceedings of the Conference on Optimization and Orthogonal Array Designs* (in Chinese). Chinese Field Statistical Institute, Beijing, 1993.
32. R. N. Kacker, E. S. Lagergren and J. J. Filliben, *J. Res. Natl Inst. Stand. Technol.*, 1991, **96,** 577.
33. R. N. Kacker, E. S. Lagergren and J. J. Filliben, *J. Quality Technol.*, 1991, **23,** 107.
34. D. B. Duncan, *Biometrics*, 1955, **11,** 1.

APPENDIX A. Establishment of the Polynomial Model and Estimate of the Regression Coefficients[11]

First let us assume the following statistical model: (i) for a four-level factorial design, the polynomial model representing the response surface can be expressed as a third-order regression equation:

$$y = \beta_0 + \sum_{X=A}^{E} \beta_X \phi_X + \sum_{X(\neq X')=A}^{E} \beta_{XX'} \phi_{XX'} + \sum_{X=A}^{E} \beta_{XX} \phi_X^2 + \sum_{X=A}^{E} \beta_{XXX} \phi_X^3 + \epsilon, \quad \text{(A1)}$$

where

$$\phi_x = \frac{Z_x - \bar{Z}_x}{H_x} \quad \text{(A2)}$$

in equation (A2), assume that

$$H_x = Z_{x2} - Z_{x1} = Z_{x3} - Z_{x2} = Z_{x4} - Z_{x3} \quad \text{(A3)}$$

$$\text{then } \bar{Z}_x = \frac{Z_{x1} + Z_{x2} + Z_{x3} + Z_{x4}}{4} = Z_{x1} + 1.5H_x. \quad \text{(A4)}$$

Therefore

$$\phi_{x1} = -1.5;\ \phi_{x2} = -0.5;\ \phi_{x3} = 0.5;\ \phi_{x4} = 1.5; \quad \text{(A5)}$$

and

$$\sum_{k=1}^{4} \phi_{xk} = 0;\ \sum_{k=1}^{4} \phi_{xk}^2 = 5;\ \sum_{k=1}^{4} \phi_{xk}^3 = 0. \quad \text{(A6)}$$

(ii) Suppose all the five columns in the $OA_{16}(4^5)$ matrix are assigned an independent variable, namely A–E, respectively. Meanwhile, suppose that there are no interactions between variables, then a polynomial model representing a response surface can be expressed as

$$y=\beta_0+\sum_{X=A}^{E}\beta_X\phi_X+\sum_{X=A}^{E}\beta_{XX}\phi_X^2+\sum_{X=A}^{E}\beta_{XXX}\phi_X^3+\epsilon. \qquad (A7)$$

Moreover, suppose that

$$E_X=\beta_X\phi_X+\beta_{XX}\phi_X^2+\beta_{XXX}\phi_X^3. \qquad (A8)$$

Then, according to equations (A7) and (A8), the response y_l for each experimental trial in the $OA_{16}(4^5)$ matrix can be described as follows:

$$y_l=\beta_0+E_{Ak}+E_{Bk}+E_{Ck}+E_{Dk}+E_{Ek}+\epsilon_l \qquad (A9)$$

where k represents the level setting numbers which are varied with the intersections in the $OA_{16}(4^5)$ matrix. ϵ_l is a statistical error which is an independent random variable from an $N(0,\sigma^2)$ distribution. Therefore,

$$\frac{1}{16}\sum_{l=1}^{16}\epsilon_l\cong 0;\ \frac{1}{4}\sum_{l=1}^{4}\epsilon_l\cong 0. \qquad (A10)$$

Now, let us randomly select a factor at a level, *e.g.* factor C at level 3, according to the $OA_{16}(4^5)$ matrix (shown in Table 1) and equation (A9), it can be concluded that

$$R_{C3}=y_3+y_8+y_9+y_{14}=4\beta_0+\sum_{k=1}^{4}E_{Ak}+\sum_{k=1}^{4}E_{Bk}$$
$$+4E_{C3}+\sum_{k=1}^{4}E_{Dk}+\sum_{k=1}^{4}E_{Ek}+(\epsilon_3+\epsilon_8+\epsilon_9+\epsilon_{14}). \qquad (A11)$$

When combined with equation (A8), equation (A11) can be rewritten as

$$R_{C3}=4\beta_0+\beta_A\sum_{k=1}^{4}\phi_{Ak}+\beta_{AA}\sum_{k=1}^{4}\phi_{Ak}^2$$
$$+\beta_{AAA}\sum_{k=1}^{4}\phi_{Ak}^3+\beta_B\sum_{k=1}^{4}\phi_{Bk}+\beta_{BB}\sum_{k=1}^{4}\phi_{Bk}^2$$
$$+\beta_{BBB}\sum_{k=1}^{4}\phi_{Bk}^3+4(\beta_C\phi_{C3}+\beta_{CC}\phi_{C3}^2$$
$$+\beta_{CCC}\phi_{C3}^3)+\beta_D\sum_{k=1}^{4}\phi_{Dk}+\beta_{DD}\sum_{k=1}^{4}\phi_{Dk}^2$$
$$+\beta_{DDD}\sum_{k=1}^{4}\phi_{Dk}^3+\beta_E\sum_{k=1}^{4}\phi_{Ek}+\beta_{EE}\sum_{k=1}^{4}\phi_{Ek}^2$$
$$+\beta_{EEE}\sum_{k=1}^{4}\phi_{Ek}^3+(\epsilon_3+\epsilon_8+\epsilon_9+\epsilon_{14}). \qquad (A12)$$

Then, substituting equation (A6) and (A10) into (A12) will lead to

$$R_{C3}=4\beta_0+5\sum_{X(\neq C)=A}^{E}\beta_{XX}+4(\beta_C\phi_{C3}+\beta_{CC}\phi_{C3}^2$$
$$+\beta_{CCC}\phi_{C3}^3). \qquad (A13)$$

Suppose that

$$4T_{CX}=4\beta_0+5\sum_{X(\neq C)=A}^{E}\beta_{XX}. \qquad (A14)$$

equation (A13) can be rewritten as

$$r_{C3}=T_C+\beta_C\phi_{C3}+\beta_{CC}\phi_{C3}^2+\beta_{CCC}\phi_{C3}^3. \qquad (A15)$$

Correspondingly

$$r_{C1}=T_C+\beta_C\phi_{C1}+\beta_{CC}\phi_{C1}^2+\beta_{CCC}\phi_{C1}^3 \qquad (A16)$$

$$r_{C2}=T_C+\beta_C\phi_{C2}+\beta_{CC}\phi_{C2}^2+\beta_{CCC}\phi_{C2}^3 \qquad (A17)$$

$$r_{C4}=T_C+\beta_C\phi_{C4}+\beta_{CC}\phi_{C4}^2+\beta_{CCC}\phi_{C4}^3. \qquad (A18)$$

Equation (A17) minus equation (A16), then combined with equation (A6), will lead to

$$r_{C2}-r_{C1}=\beta_C-2\beta_{CC}+3.25\beta_{CCC}. \qquad (A19)$$

Similarly

$$r_{C3}-r_{C2}=\beta_C+0.25\beta_{CCC} \qquad (A20)$$

$$r_{C4}-r_{C3}=\beta_C+2\beta_{CC}+3.25\beta_{CCC}. \qquad (A21)$$

Consequently

$$\beta_C=\frac{1}{24}\{27(r_{C3}-r_{C2})-(r_{C4}-r_{C1})\} \qquad (A22)$$

$$\beta_{CC}=\frac{1}{4}\{(r_{C4}-r_{C3})-(r_{C2}-r_{C1})\} \qquad (A23)$$

$$\beta_{CCC}=\frac{1}{12}\{2(r_{C4}-r_{C1})-6(r_{C3}-r_{C2})\}. \qquad (A24)$$

By using a similar method, the following formula for each factor can be inferred that

$$\beta_X=\frac{1}{24}\{27(r_{X3}-r_{X2})-(r_{X4}-r_{X1})\} \qquad (A25)$$

$$\beta_{XX}=\frac{1}{4}\{(r_{X4}-r_{X3})-(r_{X2}-r_{X1})\} \qquad (A26)$$

$$\beta_{XXX}=\frac{1}{12}\{2(r_{X4}-r_{X1})-6(r_{X3}-r_{X2})\}. \qquad (A27)$$

Because r_{X1}, r_{X2}, r_{X3} and r_{X4} can be easily calculated according to the results obtained from implementing four-level orthogonal array design, it is quite simple to establish a polynomial model as shown in equation (A7) for representing the response surface of four-level orthogonal array design.