Organic Process Research & Development **2009,** *13,* 798–803

Summary of Lecture Transcripts

Orthogonal Experiments in the Development of Organic Synthetic Processes

Rolf Carlson,* Geir Simonsen, and Alexandre Descomps
Department of Chemistry, Faculty of Science, University of Tromsoe, NO-9037 Tromsoe, Norway

Johan E. Carlson
Department of Computer Science and Electrical Engineering, Luleå University of Technology, SE-971 87 Luleå, Sweden

Abstract:

A new strategy is presented for the design of explorative experiments in synthetic chemistry when the objective is to identify the important experimental variables. The methodology is based on Taylor expansion (response surface) models, and the principles are: A grid of possible settings of the experimental variables is laid out in the experimental domain. These experiments define a candidate design matrix, D_C. From D_C, a candidate model matrix, X_C is defined by appending columns for each variable in the Taylor model X_C is then factored by singular value decomposition (SVD), and $X_C = USV^T$. The rows in X_C that are most parallel to the singular column vectors in V are selected, and the corresponding experiments in D_C are identified. This gives the experimental design. The selected experiments are nearly orthogonal, and they span the dimensions of the model space. The experiments can be run in sequence, and thus, they allow for a systematic search, one experiment at a time. The design principles are illustrated by an example of the dibromination of an acetal. Four variables were studied, and from 12 experiments, all the main effects and all two-factor interaction effects were estimated. From the response surface model, conditions for quantitative yield were predicted, and a mol-scale synthesis carried out under these conditions afforded 98% yield of the isolated pure, >97% product.

Introduction

When a synthetic procedure is to be developed into an optimum process procedure it is often necessary to identify the important experimental variables by a screening design and then to adjust the procedure to an optimum performance by response surface modelling or some kind of gradient search. This can, however, be a tedious task that usually requires a large number of individual experimental runs, and sometimes, there is not time enough to do it.

This paper describes a strategy for designing experiments in organic synthesis when the objective is to find experimental conditions that can give improved yields. The procedure described is intended as a tool when syntheses are transformed from gram scale to hundreds of grams scale or to kilogram scale.

The strategy is based on experiments for which the variable settings in each experiment are near-orthogonal to each other. This allows for a systematic search of the experimental conditions, including also possible interaction effects. The new feature is that the experiments are run sequentially to peel off the dimensions of the search space one by one. It is therefore possible to stop the search when sufficiently good experimental conditions have been found. This is to be contrasted with factorial and fractional factorial designs for which all experimental runs must be completed before the experiment can be evaluated.

Requisites. It is supposed that the experimental procedure that has been used on gram scale affords *promising* results and the experimenter can assign which experimental variables are likely to be influential. It is also assumed that the experimenter can assign a possible operational domain and that it is believed that improved experimental conditions are likely to be found in the vicinity of the hitherto used conditions but that the knowledge of the reactions is insufficient for making any detailed predictions in this sense.

Taylor Expansion Approximation of the Response Function. The outcome y (for example the yield) of a synthetic reaction is dependent on the experimental conditions. These conditions can be specified by the settings, x_i, of the experimental variables (*temperature, concentrations, feed rates, stirring rate, etc.*). We can therefore assume that there is some kind of functional dependence between the result, y, and the experimental settings, $x_1, x_2, ..., x_k$, and that

$$y = f(x_1, x_2, ..., x_k)$$

In most cases it is very difficult to derive an analytical expression for the function f, but if the experimental domain is not too vast, it is reasonable to assume that a truncated Taylor expansion can give a sufficiently good approximation of f, i.e.

* To whom correspondence should be addressed: E-mail: rolf.carlson@chem.uit.no.

10.1021/op800322h CCC: $40.75
Published on Web 05/07/2009

$$y = f(0) + \sum_{i=1}^{k} \frac{\partial f(0)}{\partial x_i} \times x_i + \frac{1}{2!} \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{\partial^2 f(0)}{\partial x_i \partial x_j} \times x_i x_j + \text{higher order terms} + R(0) + e$$

in which $R(\mathbf{0})$ is a remainder term due to the truncation, and e is a random error term. $R(\mathbf{0})$ contains the model error due to truncation, and it becomes smaller the more terms are included in the model.

This expression is more conveniently written as a polynomial response surface model:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \beta_{12} x_1 x_2 + ... + \beta_{jk} x_j x_k + \beta_{11} x_1^2 + ... + \beta_{kk} x_k^2 + e$$

To assess the roles played by the experimental variables it will be necessary to obtain estimates of the polynomial coefficients. Interaction effects are often highly significant and should be accounted for in the experimentation. In spite of an increased use of statistically designed experiments in research and production, it is still a common practice, unfortunately, to vary one experimental variable at a time. Such experiments cannot account for any interaction effect, and conclusions from such experiment are often highly erroneous. To avoid this pitfall, it is necessary to run multivariate statistical designs so that possible interaction effects can be identified. In screening experiments, when the objective is to identify the most important variables, it is often sufficient to estimate the linear effects and the two-factor interaction effects. To localise the optimum experimental conditions it is sometimes necessary also to estimate the quadratic coefficients. This is an area where traditional experimental designs (factorial designs, and fractional factorial designs,[2] D-optimal designs,[3] response surface designs[4]) are highly efficient. However, in explorative synthetic chemistry the chemists are quite reluctant to use statistical designs mainly due to the misconception that such designs will contain an excessive number of experimental runs. Still today many new methods that have been established from poor experimental designs are presented. It is in this context the near-orthogonal experiments will play their roles.

Experimental Space and Model Space. The experimenter assigns a tentative Taylor expansion model. We should now distinguish between the *experimental space* and the *model space*. The *experimental space* is defined by the possible settings of the experimental variables. With two variables, x_1 and x_2, this space is two-dimensional and with three variables it is three-dimensional, see Figure 1.

The *model space* is defined by the possible variation of the variables in the Taylor expansion model. Assume that three experimental variables are to be analysed and assume also that it is necessary to consider two-factor interaction effects. The corresponding Taylor model will be

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + e$$

and the model space in this case will be six-dimensional and spanned by $\{x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3\}$. With a full quadratic Taylor polynomial, the model space will be nine-dimensional and spanned by $\{x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3, x_1^2, x_2^2, x_3^2\}$.

Near-Orthogonal Experiments by SVD Design. The following iterative procedure is used to generate the experimental design:

(1) Select a set of candidate experiments that define a grid of points in the experimental domain, i.e. the space spanned by the variable axes. In our first attempts we have used 11 levels of each variable, and the sets of candidate experiments are given by the full 11-level factorial design. For two variables, the grid contains 121 candidate experiments, for three variables, 1331 candidates; for four variables 14641 candidates; for five variables, 161051 candidates; and for six variables 1771561 candidates. We assume that this gives a sufficient spread of the candidate experiments in the experimental domain. This defines the candidate design matrix $\mathbf{D}_c$.

(2) Suggest the response surface model. A candidate model matrix, $\mathbf{X}_c$, is then constructed by appending columns corresponding to each term in the model (cross-products (interaction) and squares). The columns of $\mathbf{X}_c$ define the *model space*. The matrix $\mathbf{X}_c$ is usually very large. $\mathbf{X}_c$ is then factored by singular value decomposition, SVD

$$\mathbf{X}_c = \mathbf{USV}^{\mathrm{T}}$$

The vectors in $\mathbf{U}$ and $\mathbf{V}$ are orthonormal, $\mathbf{S}$ is a diagonal matrix of the singular values, σ_i. The vectors in $\mathbf{V}$ are the eigenvectors of the variance−covariance matrix, $\mathbf{X}^{\mathrm{T}}\mathbf{X}$, and the vectors in $\mathbf{U}$ are the eigenvectors of the correlation matrix, $\mathbf{XX}^{\mathrm{T}}$. The columns of $\mathbf{U}$ define an orthonormal basis for the column space of $\mathbf{X}_c$, and the columns of $\mathbf{V}$ define an orthonormal basis for the row space of $\mathbf{X}_c$. The singular values have the following properties: the eigenvalues of the information matrix, $\mathbf{X}_c^{\mathrm{T}}\mathbf{X}_c$ are equal to σ_i^2 and the eigenvalues to the dispersion matrix $(\mathbf{X}_c^{\mathrm{T}}\mathbf{X}_c)^{-1}$ are equal to σ_i^{-2}. Another important property is that the eigenvector in $\mathbf{V}$ corresponding to the largest singular value points in the direction of the largest variance of the row space of $\mathbf{X}_c$, i.e. the model space.

When the number of candidate experiments (rows in $\mathbf{X}_c$), is larger than the number of columns (the dimension of the model space) the maximum rank, r, of $\mathbf{X}_c$ equals the dimensions of the model space. In that case, when all singular values, $\sigma_1,...,\sigma_r$ are distinctively different from zero, the singular vectors, $\mathbf{v}_i$, ($i = 1,..., r$) will span the model space. It was shown by Eckhart

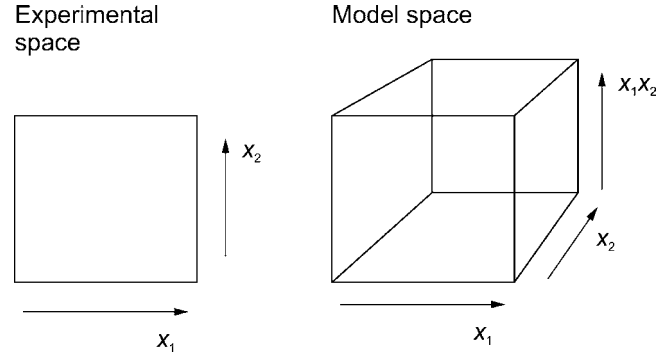

Figure 1. **Experimental space with two variables and model space with three variables.**

and Young[5] as early as in 1936 that SVD gives an optimal low-rank approximation of any matrix.

(3) The next step is to identify which row vector, $\mathbf{x}_i$, in $\mathbf{X}_c$ is most parallel to the first singular vector, $\mathbf{v}$, (i.e., corresponding to the largest singular value), in $\mathbf{V}$ as evaluated from the maximum absolute value of the scalar product $|\mathbf{x}_j\ \mathbf{v}_1|_{\mathrm{Max}}$. Then, identify which row in the candidate design matrix, $\mathbf{D}_c$, corresponds to this first selected row, $\mathbf{x}_i$, in $\mathbf{X}_c$. This yields the first experiment in the experimental design matrix. This experiment will represent a direction through the candidate design accounting for the largest variance, thus being of importance when finding a minimum set of experiments that efficiently span the variations of the model space.

(4) When the first experiment has been chosen, the next step is to remove the component in this direction from all remaining rows in $\mathbf{X}_c$. The resulting matrix, $\mathbf{X}_{c-1}$ will have the rank exactly one less than $\mathbf{X}_c$ and the corresponding rows, $_k$ are computed as

$$\hat{x}_k = \mathbf{x}_k - (\mathbf{x}_i,\mathbf{x}_k^{\mathrm{T}})/(\mathbf{x}_i,\mathbf{x}_i^{\mathrm{T}})\cdot\mathbf{x}_i$$

$\mathbf{X}_{c-1}$ is then factored by SVD and the row that is most parallel to the first singular vector is determined. The corresponding row in $\mathbf{D}_c$ is identified. This gives the second experiment in the design.

This procedure is repeated until the desired experiments have been selected. When r experiments have been selected experiments, they will span the model space.

The singular vectors, $\mathbf{v}_i$, are orthogonal, and the selected rows in $\mathbf{X}_c$ will be as orthogonal as possible. The selected experiments will thus peel off the dimensions of the model space, one experiment by one. Since the experiments are near-orthogonal, each new experiment will provide as much new information as possible. This permits a systematic search of the model space. The design is interruptible, and the experimenter can stop when a satisfactory result has been obtained. When enough experiments have been run, it is possible to fit the suggested model.

The principle for the selection of experiments is illustrated in Figure 2.

The algorithm for generating the design is illustrated in Figures 3 and 4.

We have up to now determined designs with 3, 4, and 5 variables for fitting linear, second-order interaction models, and quadratic models. The candidate experiments were defined by 11-level full factorial designs. These designs are summarised in the Appendix.

A Note on Computations. The selection procedure described above is new and has not yet been implemented in any commercial software. We have used the MATLAB software[6] for determining the design matrices. The singular vectors, $\mathbf{v}_i$ in

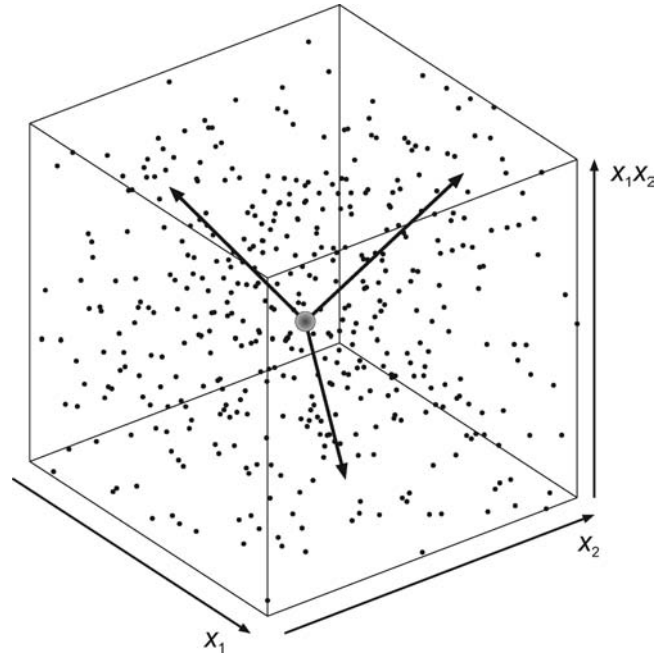

Figure 2. **Orthogonal vectors defining experiments in a three-dimensional model space.**

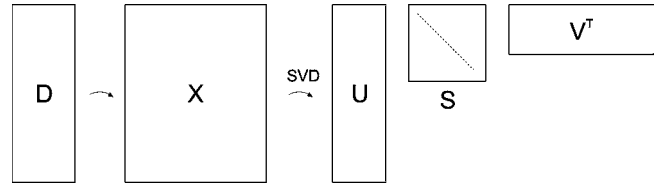

Figure 3. **Singular value decomposition of the candidate model matrix $\mathbf{X}_c$.**

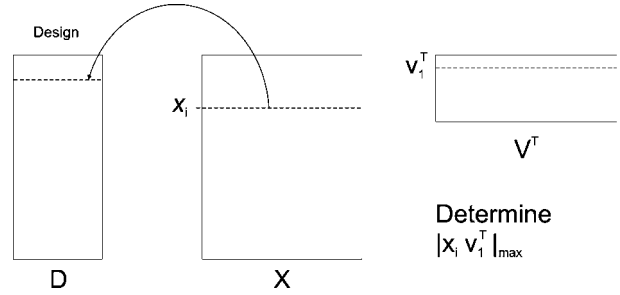

Figure 4. **Selection of experiments that are parallel to the singular vectors.**

$\mathbf{V}$ are identical to the loading vectors $\mathbf{p}_i$ obtained in principle component decomposition of a matrix $\mathbf{X}$ and $\mathbf{X} = \mathbf{TP}^{\mathrm{T}}$. The matrix $\mathbf{P}$ is defined by the loading vectors, $\mathbf{P} = [\mathbf{p}_1\ \mathbf{p}_2 \ldots\ \mathbf{p}_r]$. For this reason, any commercial software that can perform principal component analysis[7] can be used to determine the singular vectors.

Distribution of the Selected Experimental Points in the Model Space. We show an example with three experimental variables. The distribution of the experimental points in the

(1) *Optimising Organic Reactions*, presented at the Scientific Update Conference, Basel, Switzerland, 29−30 October , 2007.
(2) Box, G. E. P.; Hunter, J. S.; Hunter, W. G. *Statistics for the Experimenters: Design, Innovation, and Discovery*; Wiley-Intersciences: Hoboke, NJ, 2005.
(3) (a) Nalimov, V. V.; Golikova, T. I.; Mikeshina, N. G. *Technometrics* **1970**, *12*, 799–812. (b) FedorovV. V. *Theory of Optimal Experiments*; Academic Press: New York, 1972.
(4) Box, G. E. P.; Draper, N. R. *Response Surfaces, Mixtures, and Ridge Analysis*; Wiley-Intersciences: Hoboken, NJ, 2007.
(5) Eckhart, C.; Young, G. *Psychometrika* **1936**, *1*, 211–218.
(6) *MATLAB*; The MathWorks, Inc.: Natick, MA 01760, U.S.A, 2007.
(7) Some examples of commercial software are: *SIMCA*, available from Umetrics Inc. 17 Kiel Avenue, Kinnelon, NJ 07405, U.S.A.; *Unscrambler*, available from CAMO Smart, 1480 Route 9 North Suite 209, Woodbrodge, NJ 07405, U.S.A.; *SIRIUS*, available from Pattern Recognition Systems AS, Bergen High_Tech Center, Thorm. Gt 55, NO-5008 Bergen, Norway.

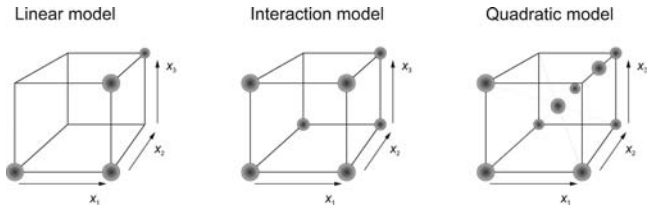

Figure 5. **Distribution of experimental points in SVD designs.**

Scheme 1

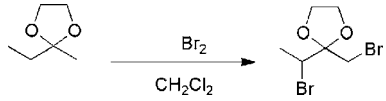

Table 1. **Experimental variables and the levels of their settings**

variables	levels of the settings −1	0	+1
x_1: reaction temperature/°C	0	1.5	30
x_2: concentration of acetal/M	0.2	0.3	0.4
x_3: stirring rate/rpm	250	325	400
x_4: rate of bromine addition/meq min^{-1}	20	50	70

Table 2. **Experimental design and yields obtained**

	design				yield
exp. no.	x_1	x_2	x_3	x_4	y
1	1.0	1.0	1.0	1.0	87.4
2	1.0	−1.0	−1.0	−1.0	95.8
3	−1.0	1.0	−1.0	1.0	79.5
4	−1.0	−1.0	1.0	−1.0	63.7
5	−1.0	−1.0	0.2	1.0	53.9
6	−1.0	1.0	1.0	−1.0	68.7
7	1.0	1.0	1.0	−1.0	58.8
8	1.0	−1.0	1.0	−1.0	93.5
9	1.0	−1.0	−1.0	1.0	94.0
10	−1.0	−1.0	−1.0	−1.0	77.1
11	1.0	−1.0	1.0	1.0	80.9
12	0	0	0	0	88.6

experimental domain of SVD designs for a linear model, an interaction model, and a quadratic model are shown in Figure 5.

From Figure 5 it is seen how such designs in this case (three variables) can be used in a sequential manner; a linear model can be fitted from four experiments. If this is unsatisfactory, an interaction model can be established by adding a few complementary experiments. A quadratic model can be established from the interaction model design by adding a few complementary experiments in the interior of the search space.

An Example: Bromination of an Acetal. We show an example of a SVD design in the bromination of the ethylene acetal from 2-butanone, see Scheme 1.

Laboratory-scale (10 mmol) experiments had afforded yields in the range 80−84%. Four variables were investigated, and their variations were chosen to embrace the hitherto known best conditions. The variables and their settings are given in Table 1. As interactions are likely, a second-order interaction Taylor model was assigned. The design and the yields obtained are given in Table 2. The experiments carried out by the design were run on larger scale (0.1−0.2 mol). The evolution of the yield was monitored by gas chromatography (internal standard technique). After 4 h the increase in yield had become insignificant, and the yields given in Table 2 were obtained after 4 h.

The second orthogonal experiment, no. 2, gave a highly increased yield compared to what was previously known as the "best" conditions. Under severe time constraint, the study could have stopped here. By using all the experiments in the design, the coefficients of the Taylor polynomial were determined using PLS regression[8], and the estimated model is

$$y = 77.71 + 8.92x_1 - 0.71x_2 - 3.11x_3 - 0.18x_4 - 6.83x_1x_2 - 1.24x_1x_3 + 2.66x_1x_4 + +0.69x_2x_3 + 6.27x_2x_4 + 1.64x_3x_4 + e$$

where e is a random error term.

The model is interpreted as follows. To increase the yield: The temperature, x_1 should be adjusted to its high level (30 °C); the concentration, x_2 should be low; the stirring rate, x_3, should be low; and the rate of addition of bromine, x_4, should be low. With these setting, the interaction effect would have a maximum beneficial influence. The predicted yield is actually 102%. We can understand the model as follows: The reaction is slightly exothermal, and to prevent unwanted temperature increase, bromine should be added slowly to the acetal at a not too high concentration. To dissipate heat from the reaction mixture, stirring is necessary, but it is probably sufficient at any level in the experimental domain. With a rapid bromine addition to a concentrated solution of the substrate, minor amounts <5% of higher brominated products were observed. A response surface projection showing the variation in yield vs x_1 and x_2 when x_3 and x_4 were set to their low level is seen in Figure 6

We have tested the suggested improved conditions in a scale-up run using 1 mol of substrate, see Experimental Section. The isolated yield was 98%, and the purity was >97% (GC, ^{1}H NMR).

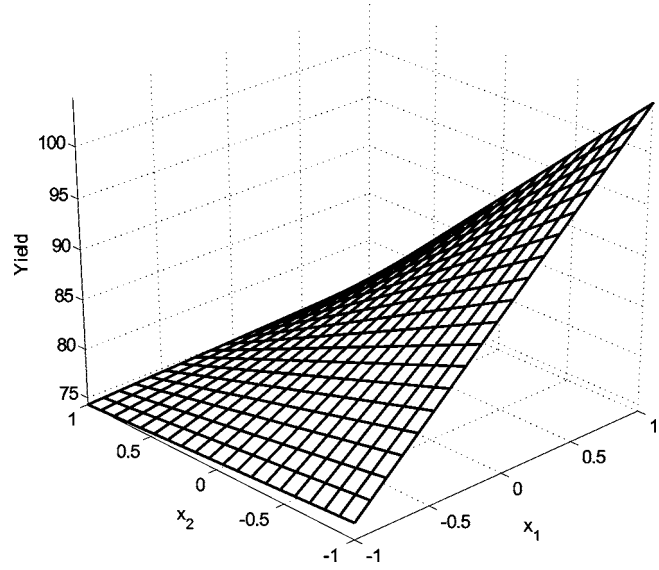

Figure 6. **Response surface projection: yield, *y*, vs the reaction temperature, x_1, and the initial concentration of the acetal, x_2. The stirring rate, x_3, and the rate of bromine addition, x_4, are set to their low values.**

Discussion

The experimental designs based on near-orthogonal experiments are intended as tools in explorative synthetic experimentation when the objective is to rapidly determine useful experimental conditions. Since the experimental settings in different experimental runs are nearly orthogonal to each other, the suggested strategy makes it possible to run the experiments sequentially, one by one, in order to systematically investigate the experimental space. It may well be possible that sufficiently good experimental conditions can be found after only a few experimental runs. In this respect, the designs based on orthogonal experiments are interruptible. We assume that this feature will make the suggested strategy attractive when time constraints impose limitations as to the number of possible experiments. We have previously shown that a design based on orthogonal experiments can be used for designing combinatorial libraries.[9] In this context it was demonstrated that such designs are A-Optimal: they minimise the trace of the dispersion matrix $(\mathbf{X}^T\mathbf{X})^{-1}$. If the experimental settings are adjusted exactly as specified by the singular vectors in $\mathbf{V}^T$, the designs become D-Optimal. This is possible when all variables are continuous over their range of variation and when a Taylor expansion model with only linear terms is attempted. If some variables are discrete and investigated on only two levels, ± 1, or if a higher-order model is attempted, it is unlikely that the experimental vectors can be adjusted to be parallel to the singular vectors. In such cases, the algorithm presented above can be used.

It was pointed out by one reviewer that the designs presented in this paper have inferior statistical properties compared with fractional factorial designs and D-Optimal designs. We agree with this criticism. When compared with fractional factorial designs or D-Optimal designs, the designs based on near-orthogonal model vectors have larger condition numbers, $\lambda_{Max}/\lambda_{Min}$ (the ratio of the largest and smallest eigenvalues of the dispersion matrix $(\mathbf{X}^T\mathbf{X})^{-1}$. It should, however, be borne in mind when and where an experimental design is laid out. If the objective is to fit a model with high precision in the estimated model parameters, factorial, designs, fractional factorial designs, composite response surface designs or D-Optimal designs should be used. The objective is then the model fit. If, on the other hand, the objective is to rapidly find improved experimental conditions and to have some information as to the most influencing variables, the designs based on near-orthogonal experiments are likely to be sufficiently good.

Experimental Section

Chemicals. 2-Ethyl-2-methyl-1,3-dioxoloane (99%) was obtained from Aldrich, dichloromethane (*Puriss.*), and bromine (*Puriss.*) were obtained from Merck, and 1,2-dichlorobenzene (*Puriss.*) was obtained from Fluka and used as delivered.

GC Analyses. A Varian 3400 gas chromatograph equipped with a flame ionisation detector coupled to a Varian 4400 integrator was used. The column was SPB-5, 30 m, 0.35 mm i.d., operated with the following temperature program: 70 °C, 5 min; 10 °C min^{-1}; 180 °C. The yields in the screening experiments were determined from the peak areas using 1,2-dichlorobenzene as an internal standard.

^{1}H NMR spectra were recorded at 400 MHz and ^{13}C NMR at 100 MHz using a Varian Mercury spectrometer.

General Procedure for the Screening Experiments in Table 2. The settings of the experimental variables, x_1-x_4 are given in Table 1.

The reactions were run in a four-necked 1 L mantled cylindrical reactor. The reaction temperature was controlled by circulating ethanol through the cooling mantle using a Julabo F70 thermostat. The flask was mounted with an anchor-shaped Teflon stirrer for which the stirring rate was adjusted using a Peaktech 2780 Tachometer, a reflux condenser connected to a HBr trap, a 250 mL pressure-equalised dropping funnel with a nitrogen inlet, and a temperature probe (Pt 100 sensor) dipping into the reaction mixture.

2-Ethyl-2-methyl-1,3-dioxolane (11.62 or 23.24 g, 0.10 or 0.20 mol, respectively) and an accurately weighed amount (ca. 6 g) of 1,2-dichlorobenzene (internal standard) were placed in the reactor and dissolved in dichloromethane to give the initial concentration, x_2, of the acetal.

The stirring rate was adjusted to x_3, and the temperature was adjusted to x_1. After 10−15 min, the temperature had reached the set value. Bromine (2 equiv) dissolved in 50 mL of dichloromethane was placed in the dropping funnel, and a slow stream of nitrogen was passed through the flask via the side arm of the dropping funnel. The rate of bromine addition was adjusted to x_4. Samples, 0.5 mL, were withdrawn at regular time intervals, washed with 5% aqueous sodium bisulfite, filtered through a plug of cotton, diluted with dichloromethane (2 mL) and analysed by GC. After 4 h (measured from the start of bromine addition) the changes in yields had become insignificant. These results are shown in Table 2.

Synthesis of 2(1-Bromoethyl)-2-(bromomethyl)-1,3-dioxoloane. The reactor was a four-necked, 2-L mantled cylindrical flask equipped as for the screening experiments, but using a 500 mL dropping funnel. The flask was charged with 2-ethyl-2-methyl-1,3-dioxolane (116.2 g, 1.00 mol) and 1 L of dichloromethane. The stirring rate was adjusted to 300 rpm, and the temperature was adjusted to 30 °C. When the temperature was stabilised, bromine (100 mL, 2 mol) dissolved in 250 mL of dichloromethane was added over 20 min, and the mixture was stirred at 30 °C for 4 h.

Workup: Water (300 mL) was added, and the mixture was stirred. Powdered sodium bisulfite was added carefully until the yellowish colour of unreacted bromine had disappeared. The organic layer was separated and washed a second time with 300 mL of water, and finally with 300 mL of saturated aqueous sodium bicarbonate to remove any remaining trace of Hbr. The organic layer was dried over anhydrous magnesium sulfate overnight. Filtration and evaporation of the solvent gave 267.9 g (98%) of 2-(1-bromoethyl)-2-bromomethyl-1,3-dioxolane. The product was >97% pure. The product can be further purified by distillation, bp 82 °C/13 mbar. Elemental analysis: Calcd. (C 26.30%, H 3.68%, Br 58.33%). Found (on the crude product) (C 25.71%, H 3.77%, Br 57.89%). ^{1}H NMR δ 1.69 (d, $J = 7.0$ Hz, 3H), 3.57 (d, $J = 11.1$ Hz., 1 H), 3.79 (d, $J = 11.1$ Hz, 1H), 4.13−4.18 (m, 4 H), 4.45 (q, $J = 7.0$ Hz, 1H); ^{13}C NMR δ 20.8, 35.0, 50.6, 67.7, 109.2.

(8) *Pirouette for Windows*, available from Infometrix Inc., P.O.Box 1528, Woodinville, WA 98072, U.S.A. The *MODDE 8.0* program was used. It is available from Umetrics Inc., 17, Kiel Ave, Kinnelon. NJ 07404, U.S.A.

(9) Carlson, R.; Carlson, J. E.; Grennberg, A. *J. Chemom.* **2001**, *15*, 455–474.

Appendix: A Design Matrices

A.1. Linear Models

Table A.1.1. **Three variables**

exp. no.	x_1	x_2	x_3
1	1	1	1
2	−1	−1	−1
3	1	−1	1
4	1	1	−1

Table A.1.2. **Four variables**

exp. no.	x_1	x_2	x_3	x_4
1	1	1	1	1
2	−1	−1	−1	−1
3	1	−1	1	−1
4	1	1	−1	−1
5	1	−1	−1	1

Table A.1.3. **Five variables**

exp. no.	x_1	x_2	x_3	x_4	x_5
1	1	1	1	1	1
2	−1	−1	−1	−1	−1
3	1	−1	−1	−1	−1
4	−1	1	−1	1	−1
5	1	1	−1	−1	1
6	1	−1	−1	1	1

A.2. Models with Linear Terms and Cross Terms

Table A.2.1. **Three variables**

exp. no.	x_1	x_2	x_3
1	1	1	1
2	−1	1	−1
3	1	−1	−1
4	−1	−1	1
5	−1	−1	−1
6	1	−1	1
7	1	1	−1

Table A.2.2. **Four variables**

exp. no.	x_1	x_2	x_3	x_4
1	1	1	1	1
2	1	−1	−1	−1
3	−1	1	−1	1
4	−1	−1	1	−1
5	−1	−1	0	1
6	−1	1	1	−1
7	1	1	−1	−1
8	1	−1	1	−1
9	1	−1	−1	1
10	−1	−1	−1	−1
11	1	−1	1	1

Table A.2.3. **Five variables**

exp. no.	x_1	x_2	x_3	x_4	x_5
1	1	1	1	1	1
2	−1	−1	−1	1	1
3	−1	0	1	−1	−1
4	1	1	−1	−1	−1
5	1	−1	1	−1	1
6	1	−1	−1	1	−1
7	−1	1	−1	−1	1
8	−1	1	1	1	−1
9	−1	−1	−1	−1	−1
10	1	1	−1	1	1
11	−1	−1	1	1	−1
12	−1	1	−1	1	−1
13	−1	−1	1	−1	1
14	1	1	1	1	−1
15	1	1	1	−1	1
16	1	−1	1	−1	−1

A.3. Quadratic Models, Including Linear Terms and Cross Terms

Table A.3.1. **Three variables**

exp. no.	x_1	x_2	x_3
1	1	1	1
2	1	−1	−1
3	−1	1	−1
4	−1	−1	1
5	0	0	0
6	1	1	−1
7	−1	−1	−1
8	−1	1	1
9	1	0	1
10	1	0	0

Table A.3.2. **Four variables**

exp. no.	x_1	x_2	x_3	x_4
1	−1	1	1	−1
2	−1	−1	−1	1
3	1	−1	1	1
4	1	1	−1	−1
5	0	1	0	1
6	1	−1	−1	−1
7	−1	−1	1	−1
8	−1	0	−1	−1
9	1	1	1	−1
10	−1	0	1	1
11	1	0	−1	1
12	1	1	1	1
13	−1	1	−1	1
14	−1	1	1	0
15	1	0	0	−1

Table A.3.3. **Five variables**

exp. no.	x_1	x_2	x_3	x_4	x_5
1	1	1	1	1	1
2	−1	1	1	−1	−1
3	1	−1	−1	1	−1
4	−1	−1	−1	−1	1
5	1	−1	1	−1	−1
6	−1	1	−1	1	−1
7	−1	−1	1	1	1
8	1	1	−1	−1	1
9	0	0	0	0	0
10	−1	1	1	−1	1
11	1	−1	−1	1	1
12	1	−1	1	−1	1
13	1	1	−1	−1	−1
14	1	1	1	1	−1
15	1	1	−1	1	1
16	−1	1	1	1	1
17	−1	−1	−1	1	0
18	−1	1	−1	−1	0
19	0	1	1	1	1
20	1	1	1	0	−1
21	1	0	−1	−1	−1

Received for review December 30, 2008.

OP800322H