Experimental Design and Optimization

1. Experimental Design Stages
   a) Identifying the factors which may affect the results of an experiment;
   b) Designing the experiment so that the effects of uncontrolled factors are minimized;
   c) Using statistical analysis to separate and evaluate

Factors – any aspect of the experimental conditions which affects the result obtained from an experiment. Can you name a factor?

The relationship between factors and responses can be shown mathematically by:

$$y = f(x_1, x_2, x_3 ..., x_k)$$

Where y is the response of interest and x are the factors that affect the response.

Experimental Design and Optimization

d. Experimental Design Approaches
   a) Full Factorial Designs (two levels per factor)
   b) Fractional Factorial Design
   c) Latin Squares
   d) Greco-Latin Squares
   e) Response Surface Designs (more than two levels for one or more factors)
   f) Box-Behnken Designs
   g) Mixture Designs
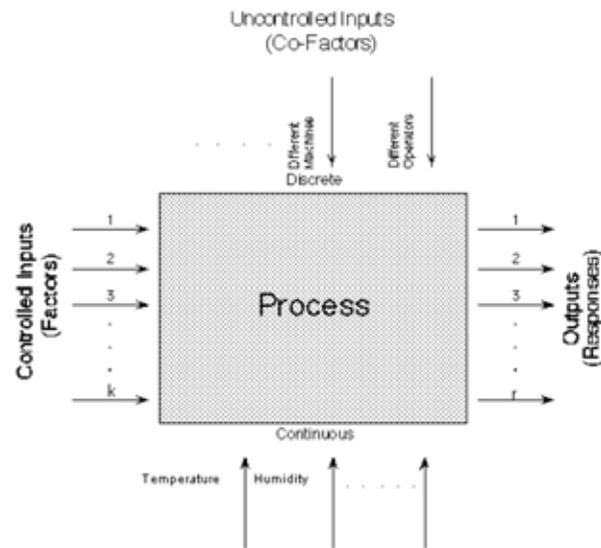
The following types of factors can be distinguished:
   1) Continuous (e.g. temperature)
   2) Discrete (e.g. Experimenter)

- Factors are independent if there is no relationship between them and dependant   if a relationship exists.

Experimental Design and Optimization

Experimental Design and Optimization

Why Bother with Experimental Design?

1. Target
2. Maximizing or Minimizing a Response
3. Reducing Variation
4. Making a Process Robust
5. Seeking Multiple Goals

Experimental Design and Optimization

How do we select an experimental design?

- Selected based the objective and the # of factors
1) Comparative Objective (e.g. one or more factors)
2) Screening Objective
3) Response Surface (method)
4) Optimizing responses
5) Optimal fitting of a regression model

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, 2005.

Experimental Design and Optimization

# of Factors	Comparative Objective	Screening Objective	Response Surface Objective
1	1-factor completely randomized design	-	-
2-4	Randomized block design	Full or fractional factorial	Central composite or Box-Behnken
5 or more	Randomized block design	Fractional factorial	Screen first to reduce the # of factors

Experimental Design and Optimization

3. Important concepts before we tackle Experimental Design:
    a) Replication – multiple measurements to reduce the experimental error.

    $\rightarrow$ averaging $\left(\dfrac{1}{\sqrt{n}}\right)$

    b) Randomization – running the experiments in a random order
        $\rightarrow$ obligatory if systematic errors (bias) cannot be avoided
    c) Blocking – Running experiments in blocks that show a minimum variance within one block.

    Ex: An investigation requires 12 experiments, only 4 expts/day

    Then, experiments should be arranged in three blocks w/ 4 expts each day.
        $\rightarrow$must be random
    d) Factorial Experiments – varying all factors simultaneously at a limited number of factor levels

Experimental Design and Optimization

Randomization – A closer look
$\rightarrow$ typically performed by a computer program
Total sample size (# of runs) = N = k x L x n
where k = # of factors , L = # of levels, n = # of replications

Ex: Completely randomized design
k = 1 factor (x1)
L = 4 levels of that factor (called "1", "2", "3" and "4"
n = 3 replications per level
N = 4 levels *3 replications per level = 12 runs

Experimental Design and Optimization

<u>Randomized Sequence</u>

<u>X1</u>

3
1
4
2
2
1
3
4
1
2
4
3

Note: There are 369,600 ways to run the experiment.

9

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, 2005.

Experimental Design and Optimization

5. How do we select and scale the Experimental Variables?

 a) Include all important factors

 b) Choose both low and high values

 c) Check for impossible combinations (e.g. low pressure and high gas flows)

 Two-level designs help us make this process easier:

<u>Matrix Notation</u>

	<u>Factor 1 (X1)</u>	<u>Factor 2 (X2)</u>
Trial 1	-1	-1
Trial 2	+1	-1
Trial 3	-1	+1
Trial 4	+1	+1

"+1" and "-1" represent "high level" and "low level", respectively for each factor.

10

Experimental Design and Optimization

5. Full factorial Designs (Screening Design)
   → 2^k – designs, where the base 2 stands for the number of factor levels
   and k expresses the # of factors.
   - with two factors, we can define a visual square.
   - with three factors, we can define a cube.

   → The lower level is usually indicated with a "_" and
   the upper level a "+" sign.

   Factor combinations:
   Ex: $2^2 = ?$ experimental conditions
   $2^3 = ?$

Experimental Design and Optimization

Ex: The effects of reaction temperature and pH in determining the
spectrophotometric response of a standard analyte solution.

Experimental matrix:

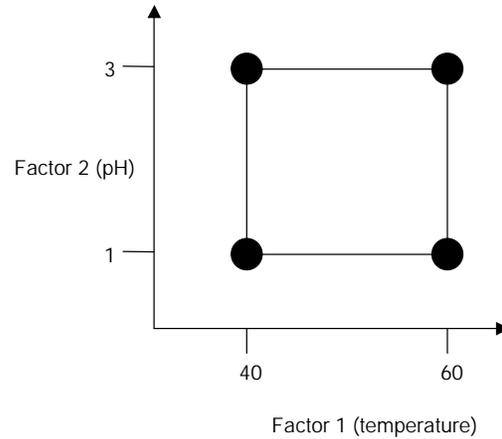| Expt #	Temperature	pH	Response
A	-1	-1	y^1
B	+1	-1	y^2
C	-1	+1	y^3
D	+1	+1	y^4

Factor Levels

(-)	40˚C	pH1
(+)	60˚C	pH3

Experimental Design and Optimization

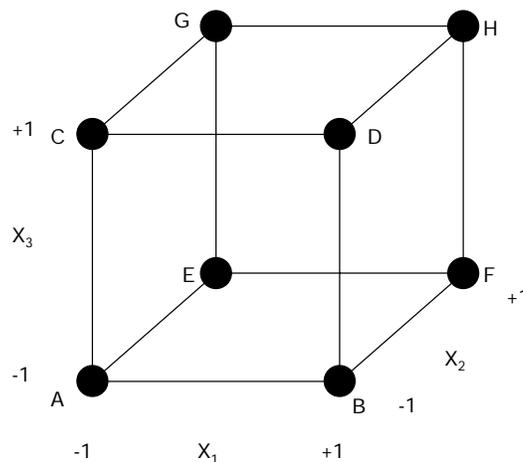Note: The best experimental points in the domain are located in the corners A, B, C and D as follows:

A(40˚C , pH1); B(60˚C, pH1); C(40˚C, pH3); D(60˚C, pH3)

13

Experimental Design and Optimization

If we introduce another variable (e.g. reagent concentration), it is then possible to represent the factors as faces on one or more cubes: 2^3 design.

14

6. Latin Square and Related Designs

-Single factor of primary interest (treatment factor) and several nuisance factors (2 for Latin Square).

-Nuisance factors are used as blocking variables.

-Used in experiments where subjects are allocated treatments over a given time period (time is a major factor).

-Allow experiments with a relatively small number of runs.

-Disadvantages:

Assumes there are no interactions between the blocking variables or between the treatments.

Experimental Design and Optimization

"Block what you can, randomize what you cannot"

Experimental Design and Optimization

Ex: Suppose the treatments are labeled A, B, and C.
In this case the design would be:

Day 1	A	B	C
Day 2	C	A	B
Day 3	B	C	A

Allows the separation of an additional factor from an equal number of blocks and treatments.

Experimental Design and Optimization

Ex: Four formulations of a drug are to be studied with regard to bioequivalence by treating 4 subjects for 4 periods (4 x 4 Latin Square). Formulations are coded by A, B, C and D.

Subject #	Period 1	2	3	4
1	A	B	C	D
2	D	C	A	B
3	B	D	C	A
4	C	A	B	D

7. Greco-Latin Squares

-Two Latin Squares that are superimposed on each other (2 treatment factors instead of 1, with 4 factors overall).

	A_1	A_2	A_3	A_4
B_1	C_1D_3	C_2D_4	C_3D_1	C_4D_2
B_2	C_4D_2	C_1D_1	C_2D_3	C_3D_4
B_3	C_3D_1	C_4D_2	C_1D_3	C_2D_4
B_4	C_2D_4	C_1D_3	C_3D_2	C_4D_1

Number of runs for a 2^k Full Factorial Design

Number of Factors	Number of Runs
2	2
3	8
4	18
5	32
6	64
7	128

Experimental Design and Optimization

Note: When the # of factors is 5 or greater…it requires a large number of runs (more time, less efficient!)

Fractional Replication

# of experiments is reduced by a number *p* according to 2^{k-p},

With the most common being *(p=1)*

Ex: A design with 2^3 (8 treatment combinations) can be run a

2^{3-1} = 4 treatment combinations

Experimental Design and Optimization

Fractional Factorial is based on an algebraic method of calculating the contributions of factors to the total variance with less than a full factorial # of expt's.

Ex: Measuring the scaled absorbance for a fixed amount of analyte as a function of pH, dielectric constant and mg L^{-1} of catalyst.

$$Y_{1i}=b_0 + b_1x_{1i} + b_2x_{2i} + b_3x_{3i} + e_{1i} \qquad \text{Eq. 1}$$

$$D = \begin{vmatrix} 8 & 72 & 4 \\ 8 & 78 & 8 \\ 10 & 72 & 8 \\ 10 & 78 & 4 \end{vmatrix} \qquad Y = \begin{vmatrix} 0.5 \\ 7.5 \\ 3.5 \\ 8.5 \end{vmatrix}$$

Experimental Design and Optimization

The model represented by Eq. 1 may be fit using:

$$X = \begin{vmatrix} 1 & 8 & 72 & 4 \\ 1 & 8 & 78 & 8 \\ 1 & 10 & 72 & 8 \\ 1 & 10 & 78 & 4 \end{vmatrix}$$

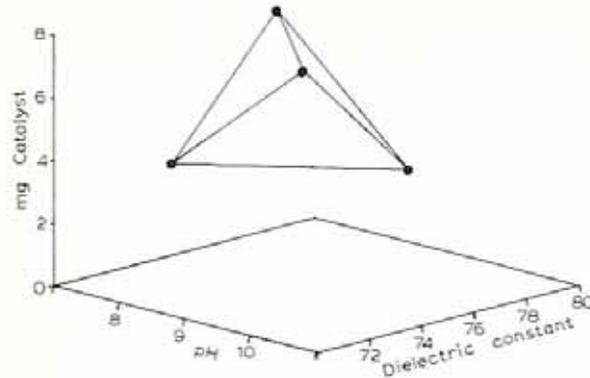Fitted model: $y_1 = -80.5 + 1.00_{x1} + 1.00_{x2} + 0.250_{x3}$

Fig. 10. A $(1/2)2^3$ fractional factorial design.

Otto, M. Chemometrics: Statistical and Computer Applications, Wiley-VCH, 1999.

Experimental Design and Optimization

8. Central Composite Designs

-Contains imbedded factorial or fractional factorial design with center points augmented with a group of axial points.

-Contains twice as many start points as there are factors in the design.

For the # of runs:

$r = 2^{k-p} + 2k + n_0$, where

k=# factors, p=# for reduction of the full design and n_0 = # of experiments in the center of the design
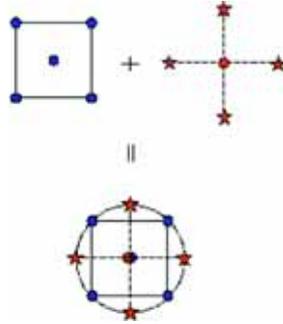
Experimental Design and Optimization

Diagram of Central Composite Design generation for 2 factors.

- Start points are at some distance α from the center, with α depending on # factors in the factorial part of the design.

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, 2005.
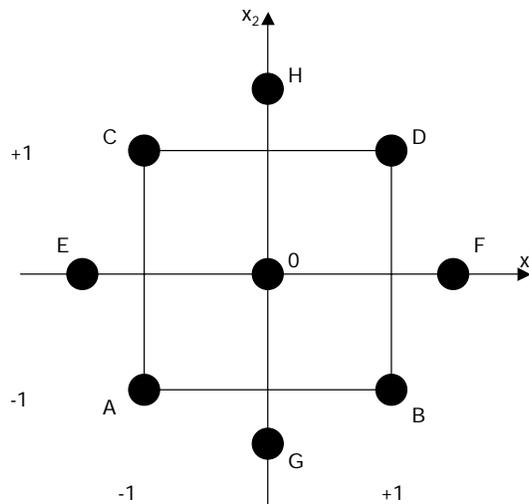
Experimental Design and Optimization

Diagram of Central Composite Design consisting of a full factorial 2-level and star design.
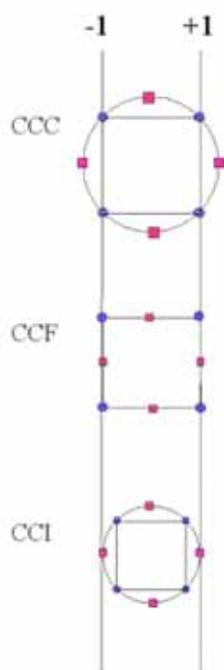
Experimental Design and Optimization

Central Composite Design Type	Comments
Circumscribed (CCC)	Original form, the star points establish new extremes for both low and high factor settings.
Inscribed (CCI)	A scaled down version of the CCC with factor levels % α
Face Centered (CCF)	Requires 3 levels of each factor, $\alpha = +/- 1$

Experimental Design and Optimization

CCC explores the largest process space and the CCI explores the smallest process space.

Both the CCC and CCI are rotatable designs, but the CCF is not.

In the CCC design, the design points describe a circle *circumscribed* about the factorial square.

For three factors, the CCC design points describe a sphere around the factorial cube.

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/, 2005.

Experimental Design and Optimization

9. Box-Behnken Designs – alternative to common CC designs

-Has three levels per factor, but avoids the corners of the space

-Fills the combinations of center and extreme levels

-Combines a fractional factorial with incomplete block designs to avoid extreme vertices & to present a "rotatable" design.

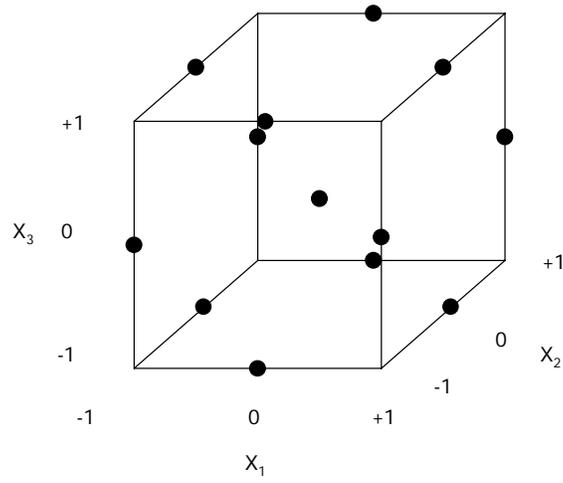-Not useful when the experimenter is interested in predicting extreme responses.

Experimental Design and Optimization

Run	Factors			Response
	x_1	x_2	x_3	
1	+1	+1	0	y_1
2	+1	-1	0	y_2
3	-1	+1	0	y_3
4	-1	-1	0	y_4
5	+1	0	+1	y_5
7	-1	0	+1	y_7
8	-1	0	-1	y_8
9	0	+1	+1	y_9
10	0	+1	-1	y_{10}
11	0	-1	+1	y_{11}
12	0	-1	-1	y_{12}
13,14,15	0	0	0	$Y_{13......}$

Otto, M. Chemometrics: Statistical and Computer Applications, Wiley-VCH, 1999.

Experimental Design and Optimization

Box-Behnken design with 3-levels per factor.