Introduction to Design of Experiments

Presented by

Ang Teck Nam

(For the course KKEK 4283)

University of Malaya
Department of Chemical Engineering
21st October 2011, 10 am

After this talk, you should know…

1. The purpose of experiment
2. The various strategies of experiment
3. How to plan and design experiment
4. Handling of experimental data

Why a well designed experimental plan is essential?

1. Maximize the amount of information from minimal number of experiments
2. Gain knowledge from the experiments
3. Gain control on the reliability of experimental results
4. Facilitate correct evaluation and writing of reports
5. Make communication possible and correct

Main applications of design of experiments

1. Discovering interactions among factors
2. Screening many factors
3. Establishing and maintaining quality control
4. Optimizing a process
5. Designing robust products

Planning of experiment…

Know the different scales of experiment:

Scale	Variability	Control over external factors
Lab	Large	Good
Pilot	Small	Medium
Production	Very small	Little

Guidelines for designing experiment

1. Statement of the problem
2. Selection of the response variable
3. Choice of factors, levels, and range
4. Choice of experimental design
5. Performing the experiment
6. Statistical analysis of the data
7. Conclusion and recommendations

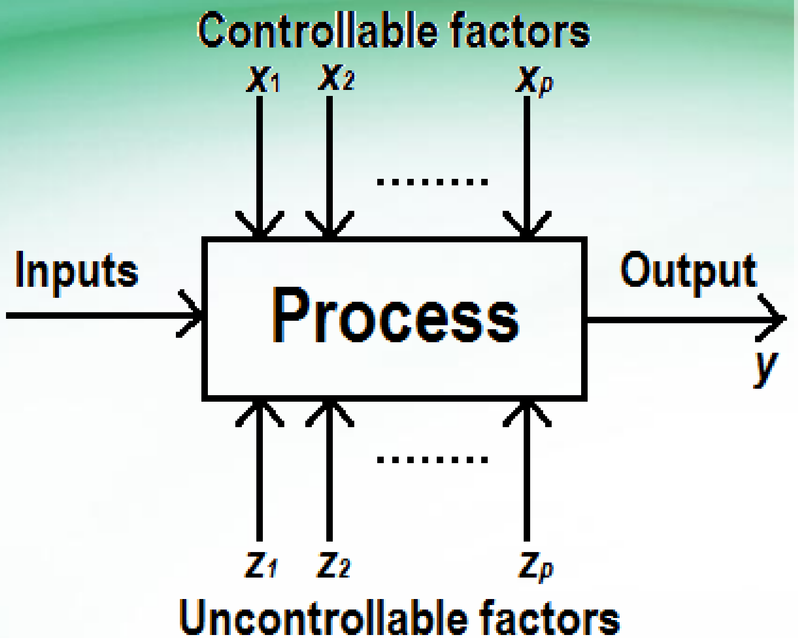
Planning of experiment…

What is the purpose of experiment?

1. Increase yield/productivity
2. Increase knowledge
3. Evaluation and comparison of basic design configurations
4. Evaluation of material alternatives

Guidelines for designing experiment

1. Statement of the problem
2. Selection of the response variable
3. Choice of factors, levels, and range
4. Choice of experimental design
5. Performing the experiment
6. Statistical analysis of the data
7. Conclusion and recommendations



General model of a process or system

Things to consider…

1. Which variables are most influential on the response *y*?
2. Where to set the influential *x*'s so that *y* is almost near the desired nominal value?
3. Where to set the influential *x*'s so that variability in *y* is small?
4. Where to set the influential *x*'s so that the effects of the uncontrollable variables are minimized?

"Think before acting!"

Guidelines for designing experiment

1. Statement of the problem
2. Selection of the response variable
3. Choice of factors, levels, and range
4. Choice of experimental design
5. Performing the experiment
6. Statistical analysis of the data
7. Conclusion and recommendations

Strategies of experimentation

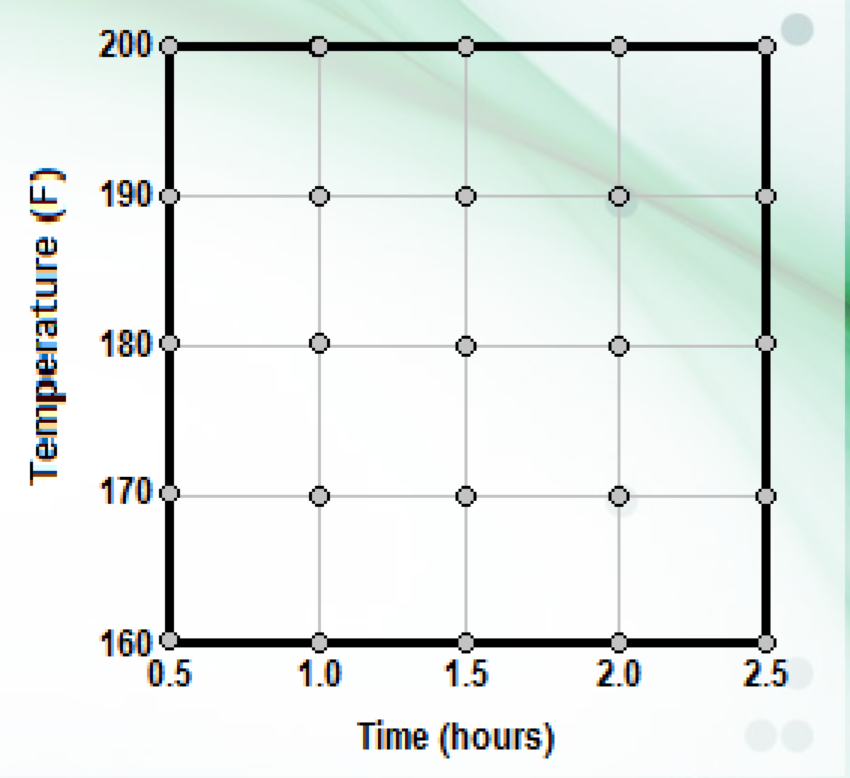
1. ***Best-guess approach***

   Work by select an *arbitrary combination of the factors*, test them and see what happens.

   Cons: Time consuming and no guarantee of best solution.

Example – Best-guess approach

Suppose we are interested in improving the yield of a chemical process. We know that 2 most important process variables that influence the yield are operating temperature (160-200 ^{o}F) and reaction time (0.5-2.5 hours).

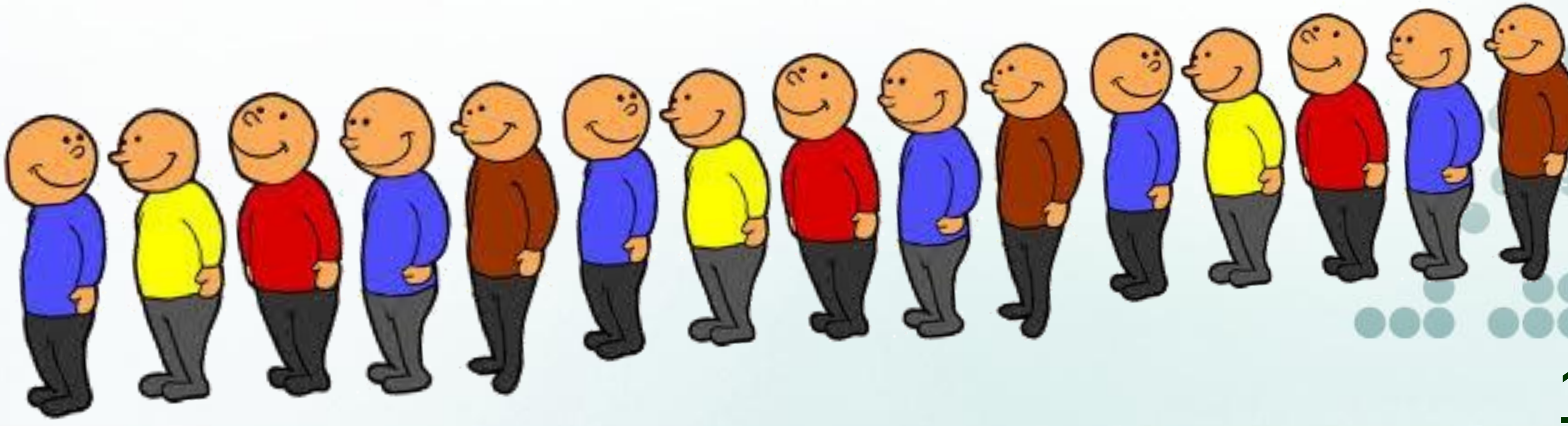


Strategies of experimentation – con’t

2. ***One-factor-at-a-time approach***

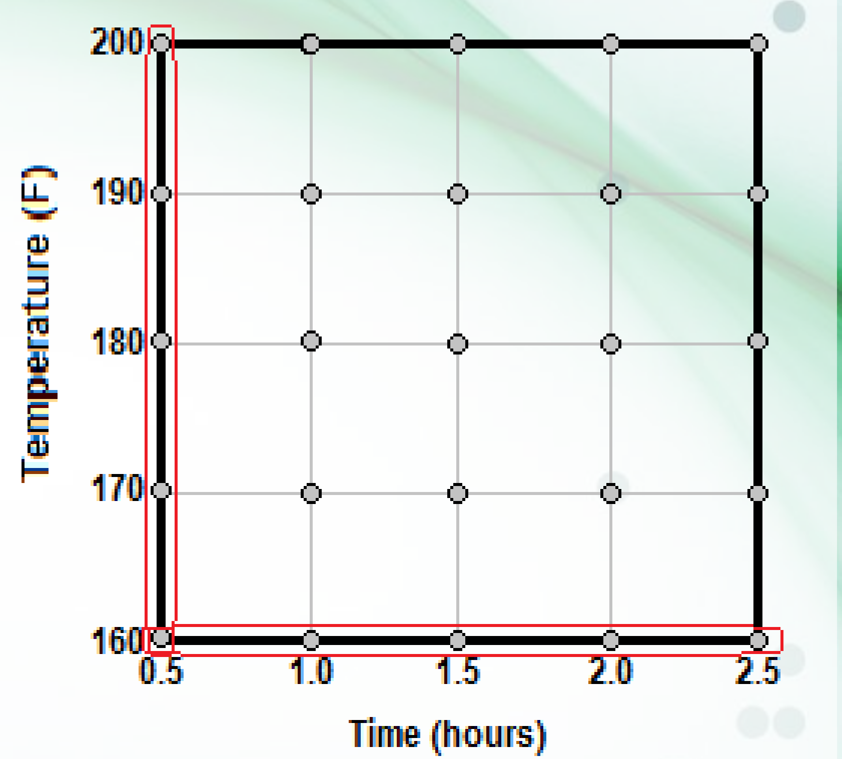
Work by selecting a *starting point* for each factor, and then *successively varying* each factor over its range with the other factors *held constant* at the starting point.

Cons: Require many experiment runs and fail to consider any possible interaction between factors.



Example – One-factor-at-a-time approach

Suppose we are interested in improving the yield of a chemical process. We know that 2 most important process variables that influence the yield are operating temperature (160-200 ^oF) and reaction time (0.5-2.5 hours).



Number of experiments = 25

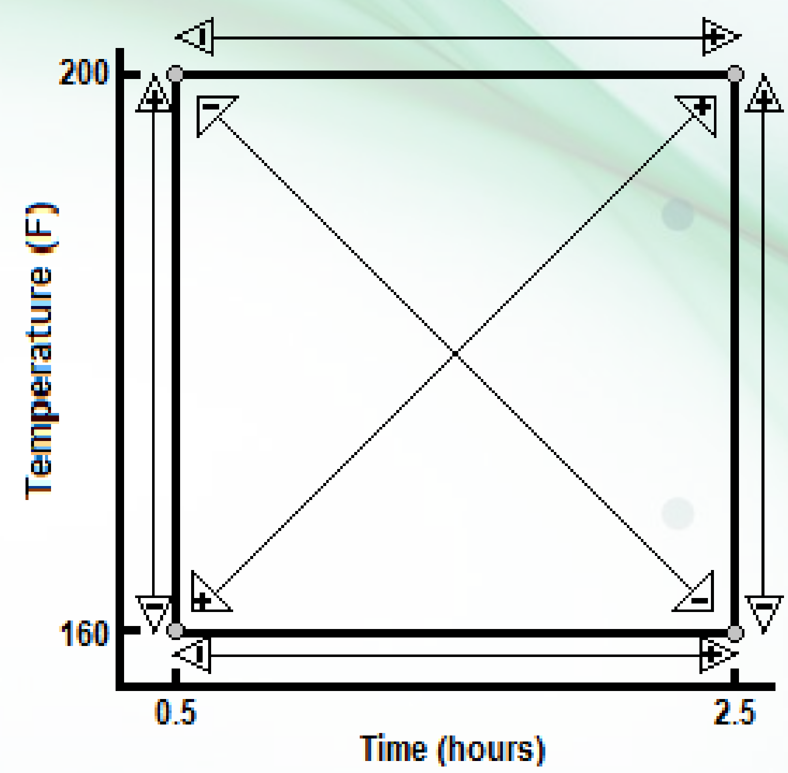
Strategies of experimentation – con't

3. ***Factorial experiment***

- Reduction or minimization of total number of trials.
- Factors under investigation are *varied together.*
- Study of interaction effect of factors.

Example – Factorial experiment

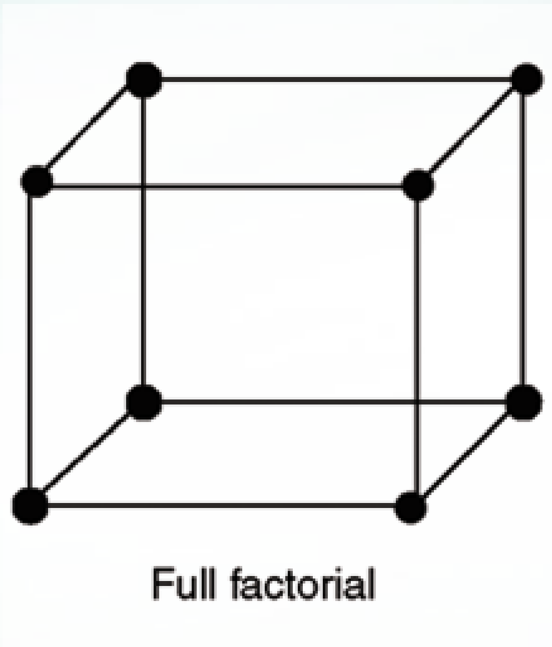
Suppose we are interested in improving the yield of a chemical process. We know that 2 most important process variables that influence the yield are operating temperature (160-200 ^oF) and reaction time (0.5-2.5 hours).

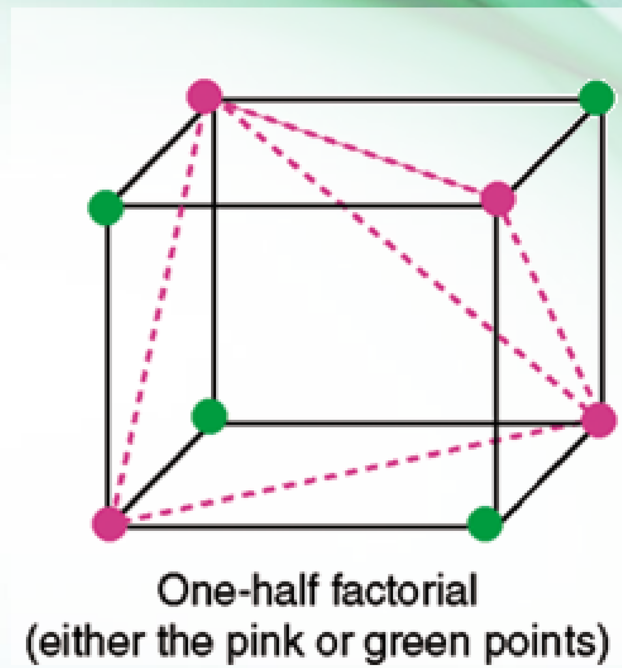


Number of experiments:

$2^2 = 4$

Factorial experiment
- Fractional Factorial Experiment



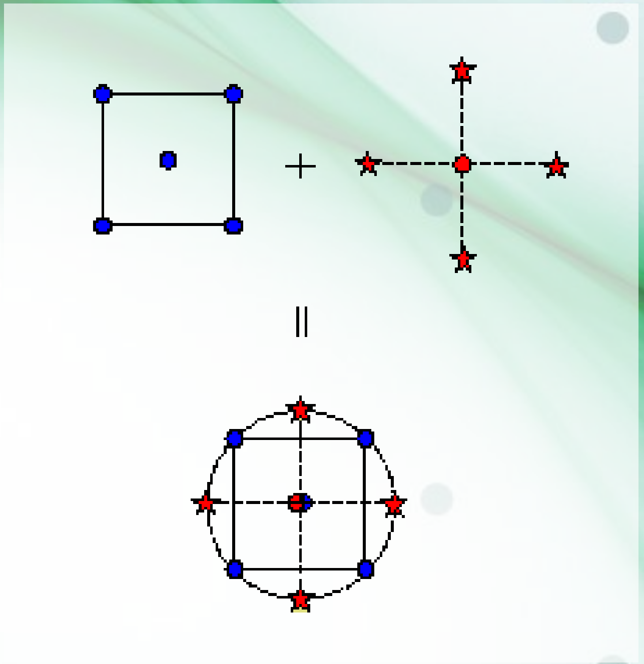


Full factorial and one-half factorial in three dimensions.

Variants of factorial experiment

1. ***Central composite design (CCD)***

   Contains an embedded factorial or fractional factorial design with centre points that is augmented with a group of 'star points' that allow estimation of curvature.

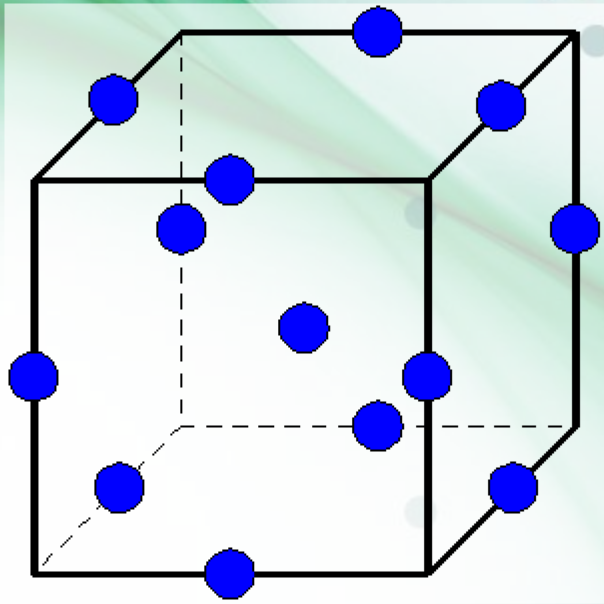


Generation of a CCD for two factors

Variants of factorial experiment

2. ***Box-Behnken design (BBD)***

Contains factorial design with incomplete block designs. The treatment combinations are at the midpoints of edges of the process space and at the center.



A BBD for three factors

Guidelines for designing experiment

1. Statement of the problem
2. Selection of the response variable
3. Choice of factors, levels, and range
4. Choice of experimental design
5. Performing the experiment
6. Statistical analysis of the data
7. Conclusion and recommendations

Response Surface Methodology (RSM)

RSM is a collection of *mathematical and statistical techniques* useful for modeling and analysis of problems where a response of interest is influenced by several variables and the objective is to *optimize this response*.

- Defines the effects of the independent variables and interaction of variables on the process.
- Generates a mathematical model describing the process.
- The response surface is visualized by response surface and contour plots.

What about dealing with large number of variables?

Screening experiment to determine the influential factors

Response surface experiment to determine the shape of the factor effects (linear or curved)

Another example…

Application of statistical experimental designs for the optimization of medium constituents for the production of L-Asparaginase by *Serratia marcescens*

Aim: To optimize medium constituents for L-Asparaginase production by *S. marcescens in solid-state fermentation using sesame oil cake as the sole substrate*

Adapted from Vuddaraju SP, Nikku MY, Chaduvula AIR, Dasari VRRK, Donthireddy SRR (2010) Application of Statistical Experimental Designs for the Optimization of Medium Constituents for the Production of L-Asparaginase by Serratia Marcescens. J microbial Biochem Technol 2: 89-94.

Another example – con't

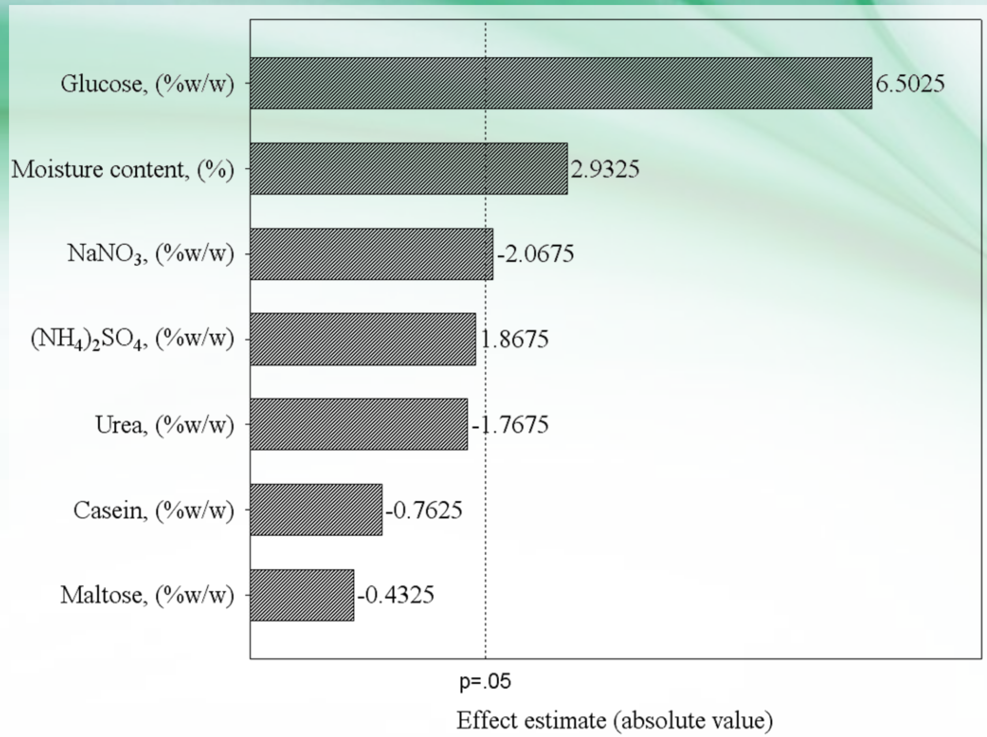
Methodology…

1. Identification of important medium constituents using Plackett-Burman design.
2. Response surface experiment to produce a prediction model to determine curvature, detect interactions among the factors, and optimize the process.

1. Identification of important medium constituents using Plackett-Burman design.

Variable	Unit	Low level	High level
Glucose	%, w/w	1.0	5.0
Maltose	%, w/w	1.0	5.0
$(NH_4)_2SO_4$	%, w/w	0.5	2.0
Urea	%, w/w	0.5	2.0
$NaNO_3$	%, w/w	0.5	2.0
Casein	%, w/w	0.5	2.0
Moisture content	%	30	80

Response: L-Asparaginase production (U/gds)



Pareto graph showing effect of various variables on L-Asparaginase production based on the observations of Plackett-Burman design.

Three variables were found to be significant ($p<0.05$), they are:

- ***a) Moisture content of substrate***
- ***b) Glucose***
- ***c) $NaNO_3$***

2. Response surface experiment to produce a prediction model to determine curvature, detect interactions among the factors, and optimize the process.

$$y = \beta_0 + \sum_{i=1}^{p} \beta_i X_i + \sum_{i=1}^{p} \sum_{\substack{j=1 \\ i \neq 1}}^{p} \beta_{ij} X_i X_j$$

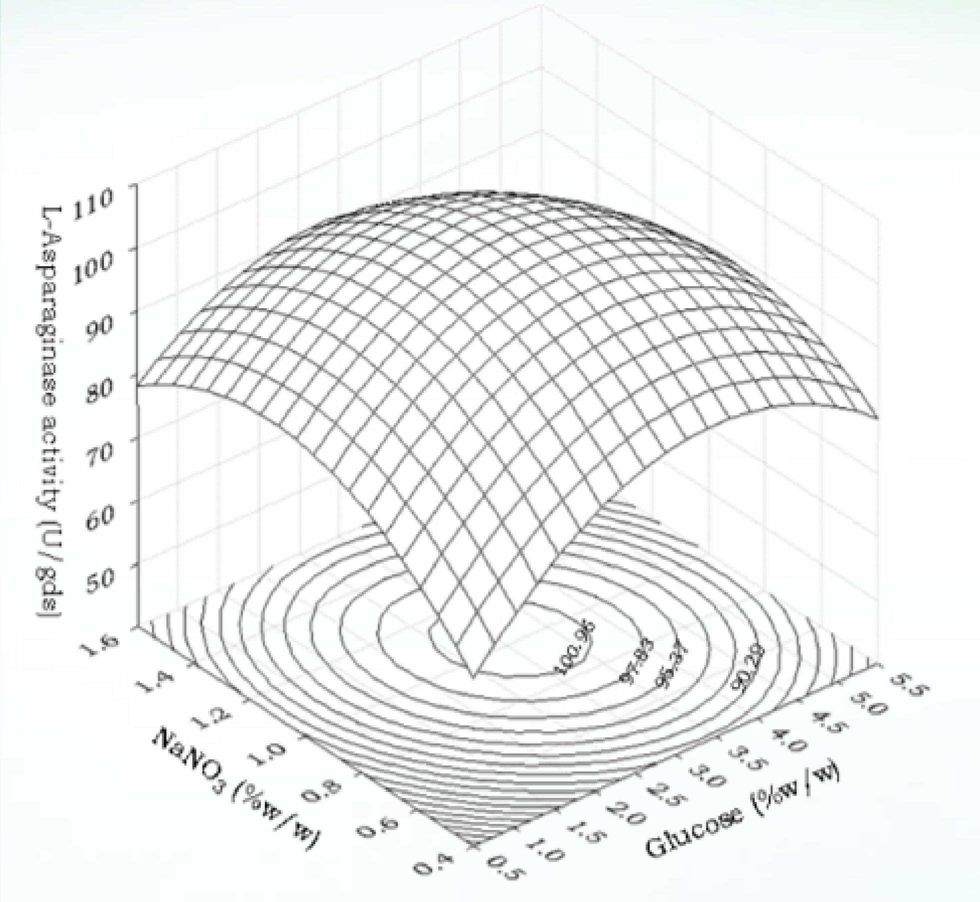
Where

β_o = the overall mean response

β_i = the main effect for factor (*I* = 1, 2, …, *p*)

β_j = the two-way interaction between the *i*th and *j*th factors

2. Response surface experiment to produce a prediction model to determine curvature, detect interactions among the factors, and optimize the process.



Response surface and contour plot of glucose vs. $NaNO_3$ on L-Asparaginase activity (moisture content was kept constant at 70%).

Rough guide for the selection of design…

Number of factors	Screening experiment	Response surface experiment
2 - 4	Full or fractional factorial	CCD or BBD
5 or more	Fractional factorial or Plackett-Burman	Screen first to reduce number of factors

***Suggestion**:*
Choose a design that requires somewhat fewer runs than the budget permits, so that center point runs can be added to check for curvature in a 2-level screening design.

Available DOE software

1. Design-Expert® (Stat-Ease Inc.)
2. DOE++® (ReliaSoft Corp.)
3. Minitab® (Minitab Inc.)
4. MATLAB ® (The MathWorks Inc.)

Useful references…

1. *A Brief Introduction to Design of Experiments* by Jacqueline K. Telford

   [Retrieved from www.jhuapl.edu/techdigest/TD/td2703/telford.pdf]

2. *Design and Analysis of Experiments (John Wiley & Sons, Inc.)* by Douglas C. Montgomery

Q & A session

Thank you for your time!