UNIT - 5

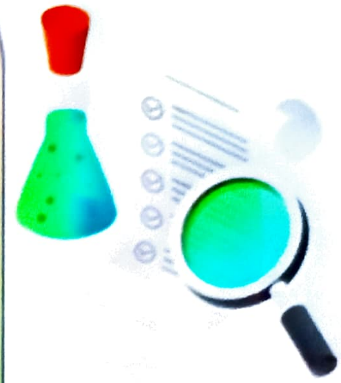
DESIGN AND ANALYSIS OF EXPERIMENTS

Points to be covered in this topic

1. **Factorial Design:**
   - Definition
   - 2^2, 2^3 design
   - Advantage of factorial design
2. **Response Surface methodology :**
   - Central composite design
   - Historical design
   - Optimization Techniques

Design and Analysis of experiments

- The Design of Experiments is carried by various methods using different tools and techniques in which Response Surface Method is one.
- These strategies were originally developed for the model fitting of physical experiments, but can also be applied to numerical experiments.
- The objective of Design of Experiments is the selection of the points where the response should be evaluated.
- Most of the criteria for optimal design of experiments are associated with the mathematical model of the process.



- Generally, these mathematical models are **polynomials with an unknown structure,** so the corresponding experiments are **designed only for every particular problem.**
- In a traditional Design of Experiments, **screening experiments are performed in the early stages of the process,** when it is likely that many of the **design variables initially considered have little or no effect on the response.**
- The purpose is to identify the design variables that have **large effects for further investigation.**
- A particular combination of runs defines an experimental design.
- The possible settings of each independent variable in the N dimensional space are called **levels.**
- Different methodologies are used such as **Full Factorial Design, Central Composite Design, D-optimal designs.**

SCHEME OF EXPERIMENTATION

- Investigation is the process that is used to find something better from the existing system.
- Experiments a are performed to analyse the situation and obtaining output responses which may further modified according to need.
- Before performing any experiment for a particular type of situation there must be target which may be termed as the aim of experiment that must be framed in mind.

DEFINITION OF AIM OF EXPERIMENT

When the aim is well defined in concurrence with the situation, the problem should be analysed with the help of the following questions:

- **What is known?**
- **What is unknown?**
- **What do we need to investigate?**
- **To be able to plan the experiments in a rational way the problem has to be concrete.**
- **Which experimental variables can be investigated?**
- **Which responses can be measured?**

EXPERIMENTS

- Experiments are actions that are carried out in order to examine the behaviour of the system and the influences of factors on the system under study.
- In statistical experimental design all experiments are determined in advance. This means that in each experiment all the factors have defined values.
- The main and interaction effects are calculated on the basis of the obtained responses and the factors' signs and the interactions' levels. These calculations are called factor analysis and can only be performed if all the experiments of an experimental design are carried out.

Before starting any experiment, four preliminary activities must be carried out:

- **Identification of all factors** that may affect the system.
- **Selection of the most significant factors** on the basis of known facts, data from literature, experience.
- Selection of **significant response parameters.**

The following steps are carried out in the technical activities
(1) Selection of suitable experimental design
(2) Experimental work
(3) Factor analysis

Some well-known types of experimental designs include:

- **A Full factorial designs for screening of factors and all possible interactions.**
- **Fractional factorial designs for screening of factors and some interactions.**
- **A Plackett- Burman designs for screening of factors without interactions.**
- **Different Response Surface designs (central composite designs, multifactor, multilevel designs, cantered cube designs) for modelling.**
- **Mixture designs for the examination and modelling of different mixtures.**

TERMINOLOGY

- **Experimental domain:** The experimental 'area' that is investigated (defined by the variation of the experimental variables).
- **Factors:** Experimental variables that can be changed independently of each other.
- **Continuous variables:** Independent variables that can be changed continuously.

- **Discrete variables:** Independent variables that are changed step-wise.
- **Responses:** The measured value of the results from experiments.
- **Residual:** The difference between the calculated and the experimental result.

APPLICATIONS OF EXPERIMENTAL DESIGNS

- Experiments are the thought **process of humans as an ordinary person** or as a **scientist or an engineer in the search of better solution** of problems that are faced by them.

- Experimental Design is an **important tool and techniques** which is used to **improve the product realization.**

- The elements of experimental activities are in **new manufacturing process design and development,** and **process management.**
- Experimental Design methods are of **fundamental importance in scientific** and **engineering design activities** where **existing products were improved after performing experiment** and a new one is developed.

PRINCIPLES OF EXPERIMENTS

- **Statistical deigns of experiments** refers to the **process of planning the experiment**, the data is **analysed by using statistical methods** such as different type of tests involved in statistics e.g. **chi square test, p-test resulting in valid and objective conclusions from the data.**

There are three basic principles of design of an experiment that is

- **Randomization:** Randomization means that the **both the allocation of the experimental material and the order in which the individual runs or trials of the experiment are to be performed are randomly determined.**

- **Replication:** Replication means an **independent repeat of each factor combination while performing an experiment.** In fact replication assists experimenters to **observe the difference in two data which are being analysed.**

- **Blocking:** Blocking is a tool **used in experimental design to improve the precision with which comparisons among the factors of interest are made.** The variability transmitted by nuisance factors **(the factor which influences the experiment)** are reduced or eliminated by blocking.

STEPS IN DESIGNING EXPERIMENTS

1. Recognition of and statement of problem
2. Selection of the response variable
3. Choice of factors, level and range
4. Choice of experimental design
5. Performing the experiment
6. Statistical analysis of the data
7. Conclusions and recommendations
8. Feedback from the users who uses the concluding result of an experiment for further analysis and improvement

Factorial design

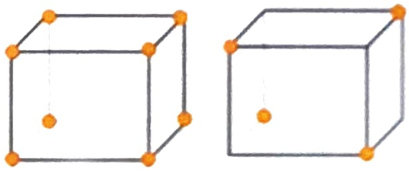
Definition

- A factorial design is defined as, "a design in which two or more variables (factors) are employed in a most efficient way for the experiments that all the possible combinations of selected values of each variable are used".

- In factorial design, selected values of two or more independent variables are manipulated in all possible combinations so that their independent as well as interactive effects on the dependent variable are studied.

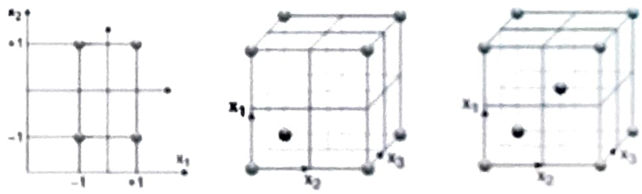
Objectives of Factorial Design

The objective of the factorial designs in experiments is to determine where to set conditions:



1. That has most influence on the response.
2. That has response close to desired nominal value.
3. That has least variability in the response.
4. To minimize effects of the uncontrollable variables.

The 2^2 Design



- ✓ Factorial experiments with two-level factors are most widely used because they are easy to design, efficient to run, straight forward to analyze, and full of information.
- ✓ The full factorial design supports both continuous and categorical factors with up to 9 levels.

Example :

- **A2^2-factorialdesign: Pharmacy pellets.**

- Consider a process for **producing caffeine pellets** with a certain size (0.71-1.4 mm).
- The aim was to **obtain a robust process** giving a yield higher than **95% of this fraction.**
- **Two variables were investigated**: amount of **water in the granulation (Granv)** and **spheronizer speed (Sfhast).**
- A 2^2 **full factorial design** was used to **study the robustness of the process.**

Table- Level of the variables

Experimental domain

	Level	Level
X_1 (Granv)	25,37	25,93
X_2 (Sfhast)	650	950

Table- Design and yield response

X_1	X_2	Yield (0.71-1.4 mm.) (%)
-	-	97.4
+	-	98.1
-	+	97.1
+	+	97.8

The 2^3 Design

- ✓ A 2-factor factorial design is an experimental design in which **data is collected for all possible combinations of the levels of the 2 factors of interest.**
- ✓ **A two by three or (2 x 3) or 2^3 design consists of 2 levels and 3 factors.** If design consists of 2 levels for one variable and 3 levels of another variable then also it is a 23 factorial design.

Example :

- **A2³-factorial design: Pharmacy, formulation of tablets**

➤ In the formulation of a certain tablet **three variables were considered to be important for the thickness of the tablets.**

➤ These variables were **investigated by a factorial design.**

➤ The different variables were the **amount of stearate lubricant.**

➤ The amount of **active substance** and the amount of **starch disintegrate.**

Table: Variables and Experimental Domain of the Formulation

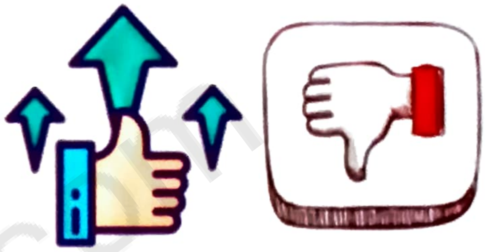
	Experimental domain		
Variables	**(-) Level**	**0-Level**	**(+) Level**
X_1 : Amount of stearate (mg)	0.5	1	1.5
X_2 : Amount of active substance (mg)	60	90	120
X_3 : Amount of starch (mg)	30	40	50

Table Design and Responses

Experiment No.	Variables			Thickness (mm)
	X_1	X_2	X_3	
1.	-	-	-	4.75
2.	+	-	-	4.87
3.	-	+	-	4.21
4.	+	+	-	4.26
5.	-	-	+	5.25
6.	+	-	+	5.46
7.	-	+	+	4.72
8.	+	+	+	5.22
9.	0	0	0	4.86

ADVANTAGE OF FACTORIAL DESIGN

- A **"main effect"** is the effect of **one of independent variables on the dependent variable**, ignoring the effects of all other independent variables.
- FDs are extremely **useful to psychologists** and **field scientists as a preliminary study**, allowing them to judge whether there is a **link between variables, whilst reducing the possibility of experimental error and confounding variables.**
- The FD, as well as **simplifying the process** and **making research cheaper**, allows many levels of analysis.



DISADVANTAGES OF FACTORIAL DESIGN

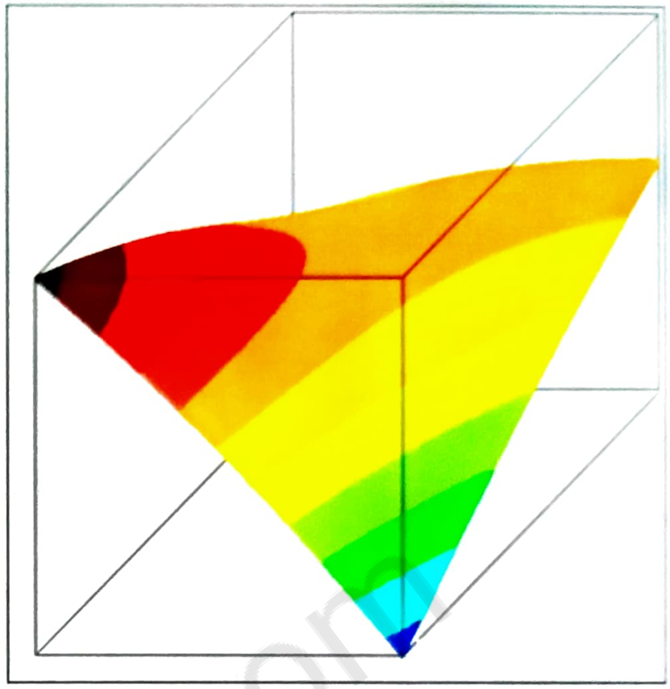
- The main disadvantage of full experimental design is the **exponential increase of experiments because of the increase of factors.**
- The disadvantage is the **difficulty of experimenting with more than two factors, or many levels.**
- A factorial designs has to be planned meticulously, as an **error in one of the levels, or in the general operationalization, will jeopardize a great amount of work.**

Applications of Factorial Designs

- In product development factorial design is an **effective tool for formulation scientists starting from pre-formulation stage to various stages of clinical trial.**
- Some of the applications of factorial design at various stages of formulation development at R&D include **compatibility studies between drug-drug and drug-excipients**, **granulation**, **pre-tablet granulation**, **oral-controlled release formulation**, **modeling of properties of powder**, **dissolution testing**, **tablet formulation**, **coating of tablets**, **extrusion-spheronization**, **inhalation formulation** etc.

RESPONSE SURFACE METHODOLOGY

- Response Surface Methodology (RSM) is a collection of mathematical, graphical, and statistical techniques for modelling and analysis of problem.
- By careful design of experiments, the objective is to optimize a process which is influenced by several independent variables.



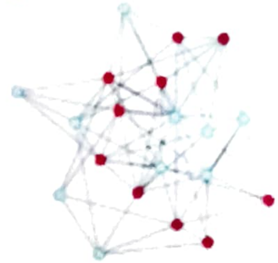
Central Composite Design

- Central composite designs are two level full factorial (2^K) or fractional factorial (2^K) designs augmented by a number of centre points and other chosen runs.
- These designs are such that they allow the estimation of all the regression parameters required to fit a second order model to a given response.

- ❑ **Numeric Factors:** The number of numeric factors involved in the experiment.
- ❑ **Category Factors:** The number of category factors involved in the experiment.
- ❑ **Alphas:** To ensure that even the extreme axial runs are within the area of operability.
- ❑ **Type:** There will be different options available in this pull down based on how many factors are to be included in the experiment.
- ❑ **Blocks:** Central composite designs can be split into blocks.
- ❑ **Options button:** Click on the options button to change the axial (alpha) distance which is how far the star points will be from the centre in coded units

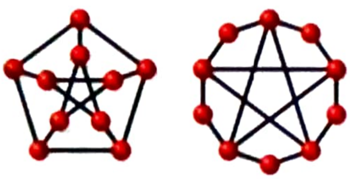
D-OPTIMAL DESIGNS

- The D-optimality criterion enables a more **efficient construction of a quadratic model.** The objective is to select **P design points from a larger set of candidate points.**
- Equation below can be expressed in matrix notation as:

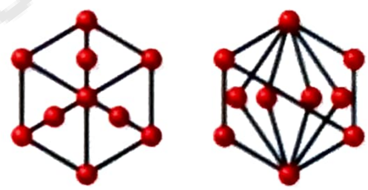


$$Y = X \times B + e$$

- Where **Y is a vector of observations, e is a vector of errors, X is the matrix of the values of the design variables at plan points** and **B is the vector of tuning parameters.**
- B can be estimated using the **least-squares method** as:



$$B = (X^T \times X)^{-1} X^T Y$$

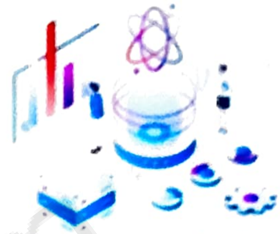


THE SEQUENTIAL NATURE OF THE RESPONSE SURFACE METHODOLOGY

- **Most applications of RSM are sequential in nature and can be carried out based on the following phases.**
    - **Phase 0:** The objective of factor screening is to **reduce the list of candidate variables to a relatively few** so that subsequent experiments will be **more efficient** and require fewer runs or tests.
    - **Phase 1:** The experimenter's objective is to determine if the **current settings of the independent variables result in a value of the response that is near the optimum.**
    - **Phase 2:** Begins when the **process is near the optimum.** At this point the experimenter **usually wants a model that will accurately approximate the true response function** within a relatively small region around the optimum

HISTORICAL DESIGN

- The purpose of a historical research design is to collect, verify, and synthesize evidence from the past to establish facts that defend or refute a hypothesis.
- It is the study of objects of design in their historical and stylistic contexts. With a broad definition, the contexts of design history include the social, the cultural, the economic, the political, the technical and the aesthetic.

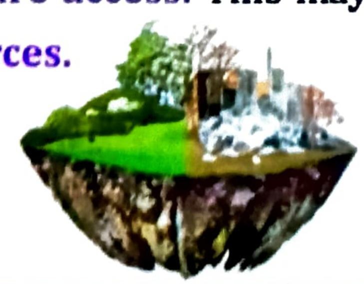


❖ What do we can observe from these studies

- The historical research design is unobtrusive; the act of research does not affect the results of the study.
- The historical approach is well suited for trend analysis.
- Historical records can add important contextual background required to more fully understand and interpret a research problem. There is often no possibility of researcher subject interaction that could affect the findings.

❖ What are drawbacks of these studies

- The ability to fulfil the aims of our research are directly related to the amount and quality of documentation available to understand the research problem.
- Since historical research relies on data from the past, there is no way to manipulate it to control for contemporary contexts. Interpreting historical sources can be very time consuming.
- The sources of historical materials must be archived consistently to ensure access. This may especially challenge for digital or online-only sources.



OPTIMIZATION TECHNIQUES

- **Optimization** is a process of **finding the best way of using the existing resources while taking in to the account of all the factors that influences decisions in any experiment.**
- Optimization techniques consisting of **statistically valid experimental design were originated to provide an economical** way to obtain efficiently the most information while expending the least amount of experimental effort.

❖ **Following are some steps that represent the use of latest optimization techniques for pharmaceutical formulations:**

I. **Defining objectives and planning the experiment accordingly.**
II. **Screening of factors those influence study.**
III. **Use of RSM for experimental designs.**
IV. **Formulation and evaluation of drug delivery systems as per design. Computer-aided modeling and searching for an optimum.**
V. **Validation of design selected for experiments.**
VI. **Scale-up and implementation to large scale production.**

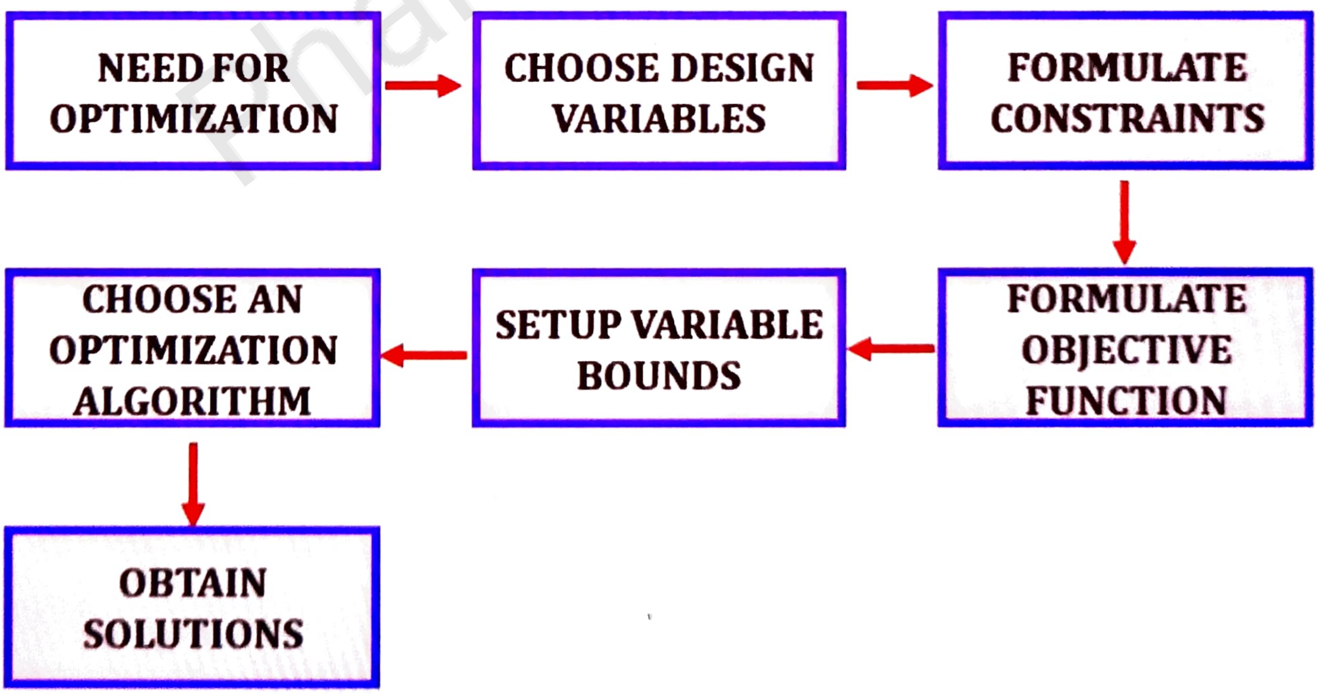


Fig : Steps in Formulating an Optimal Design

OPTIMIZATION OF FACTORS

(a) Model Development:

- A model is an expression defining the **quantitative dependence of a response variable on the independent variables**. It is a set of polynomials of a given order or degree.
- From this polynomial equation, researcher calculates the **coefficient with the help of principle of Multiple Linear Regression Analysis (MLRA).**

(i) Graphical optimization:

- Graphical optimization deals with **selecting the best possible formulation out of a feasible factor space region.**
- To do this, the desirable limits of response variables are set, and the **factor levels are screened accordingly with the help of overlay plot.**

(ii) Brute-force search (Feasibility and Grid search):

- Brute-force search technique is the **simple and exhaustive search optimization technique.**
- It checks each and every **single point in the function space.**
- Herein, the formulations that can be **prepared by almost every possible combination of independent factors** and **screened for their response variables.**

(iii) Numerical optimization:

- It deals with selecting the **best possible formulation out of a suitable factor.**
- To do this, the desirable **limits of response variables are set,** and the **factor levels are displayed by the software.**

(b) Validation of Model:

- The predicted optimal formulation (check point) is **prepared as per optimum factor level and the responses evaluated.**
- On comparison of results of observed and predicted response **conclusion can be drawn for model validation.**

❑ Advantages of Optimization

1. Reduces cost with return on investment.
2. Optimization may be helpful in shortening the experimenting time.
3. It reduces error and provides safety.
4. Provides best solution with reproducible results within the domain of study.
5. Require fewer experiments to achieve an optimal formulation.
6. Can trace and rectify problems in a remarkably easier manner.
7. Helps better to understand the process and thus less batch failures.
8. More efficient and effective control of change.
9. Provides opportunities for more flexible regulatory approaches reducing number of post-approval submissions.
10. Facilitate incorporation of manufacturing changes within the approved design space without further regulatory review.
11. Better innovation due to the ability to improve processes without resubmission.

OPTIMIZATION PARAMETERS

(a) Unconstrained optimization:

- For unconstrained problems, Fletcher-Reeves (FR) and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) methods are used.
- The FR method makes use of conjugate search directions to reach the optimum.
- The joint search directions are created using information gained from the previous design repetition.
- This method works well and theoretically minimizes a quadratic function in n or fewer repetitions.

(b) Constrained optimization:

- For constrained optimization problems, Sequential Unconstrained Minimization Techniques (SUMT) approach and direct (or constrained) methods are used.

FORMS OF OPTIMIZATION TECHNIQUES

(a) Sequential optimization (Hill climbing method):

In this method initially **small number of experiments is conducted**, and then **research is done by either increase or decrease of response.**

(b) Simultaneous optimization:

In this optimization, the **relation between any dependent variable and one or more independent variables is known.** This technique involves the use of **full range of experiments by an experimental design** and **results are fitted in the mathematical model** to obtain maximum or minimum response.

(c) Classical optimization:

These techniques being **analytical methods, makes use of differential calculus** in locating the optimum solution to the basic problem.

(d) Numerical Methods:

1. Linear method: In this method the **objective function f is linear and the set design variable space is specified** using only linear equalities and inequalities.

2. Integer method: This method **studies linear plan** in which some or all variables are constrained to take on integer values.

3. Quadratic method: This method allows the **objective function to have quadratic terms**, while the design variable space must be specified with linear equalities and inequalities.

4. Non-linear method: This method studies the general case in which the **objective function or the constraints or both contain non-linear parts.**

(e) Applied optimization:

(1) Lagrangian method: Lagrangian method is mathematical optimization technique which is an **extension of classic method.** It is especially applied to a **pharmaceutical formulation and processing.** This technique requires that the **experimentation be completed before optimization** so that the mathematical models can be developed.

(2) Evolutionary operation:

- Evolutionary operation (EVOP) is an **optimization technique for process improvement based on the principle** that, **processes generate products or services together with data useful** for providing guidance for improvement.
- It is a process or technique of systematic experimentation **especially well suited to a production.**

(3) Search Method:

- Search is an attempt to find a **feasible point that satisfies all constraints,** generally using function values only.
- Local search is improving a feasible point by searching among neighbouring points; whereas stochastic search is searching using a **non-deterministic criterion for choosing trial points.**

(4) Simplex Lattice Method:

- Simplex lattice method is an optimization technique most often **used in analytical rather than formulation and process optimization.**
- A simplex is a geometric figure, defined by **number of points or vertices equal to one more than number of factors examined.**

(5) Canonical Analysis:

- Canonical analysis is a technique used to **reduce a second order regression equation.**
- This method allows **immediate interpretation of the regression equation** by including the linear and interaction terms in constant term.

(6) Artificial Neural Network:

- **Artificial neural network (ANN)** has been used in pharmaceutical studies to **forecast the relationship between the response variables and casual factors.**
- This is an **input-output model** that uses **data clusters to transform inputs using weighted transformation linear or non-linear equations.**
- ANN is most successfully used in **multi-objective simultaneous optimization problems.**