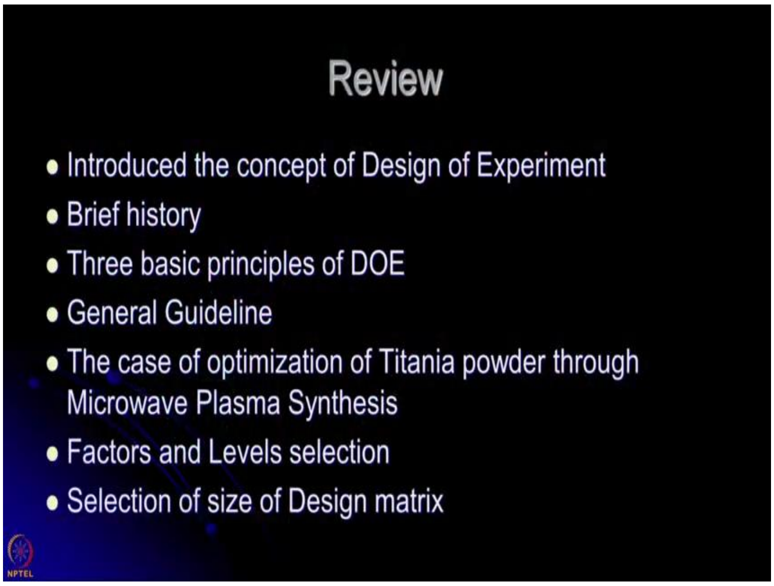
Dealing with Materials Data: Collection, Analysis and Interpretation
Professor M P Gururajan, Professor Hina A Gokhale
Department of Metallurgical Engineering and Materials Science
Indian Institute of Technology, Bombay
Lecture 88
Design of Experiment II

Hello and welcome to the course on Dealing with Materials Data. From the previous session, we have we are learning more about how to conduct a design of experiment and its analysis. In this lecture, let us recall what we went through, we what we studied in the previous session.

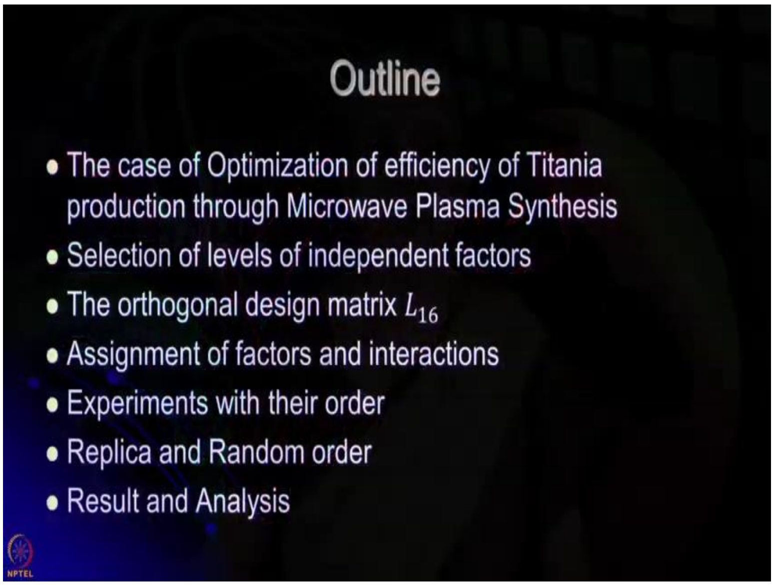
(Refer Slide Time: 0:43)



We introduced the concept of design of experiment and we gave a brief history of it. We explained three basic principles of design of experiment, namely randomization, replication and blocking.

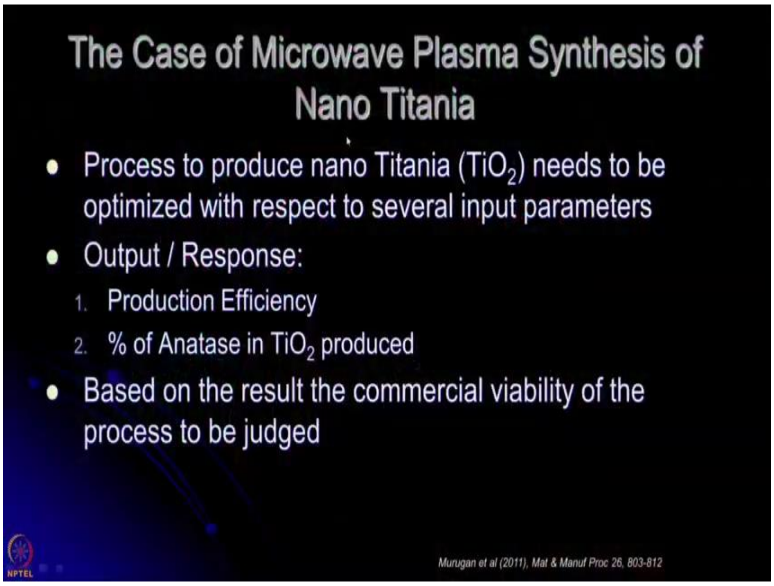
Then we followed, we gave a certain general guideline and we followed this general guideline for the case of optimization on Nano Titania powder through micro plasma synthesis. We identified factors, levels. We selected their levels and we selected a size of design matrix to be 16. What we wish to do is, we will briefly go through what we did in the past, so it comes out very clear.

(Refer Slide Time: 1:33)



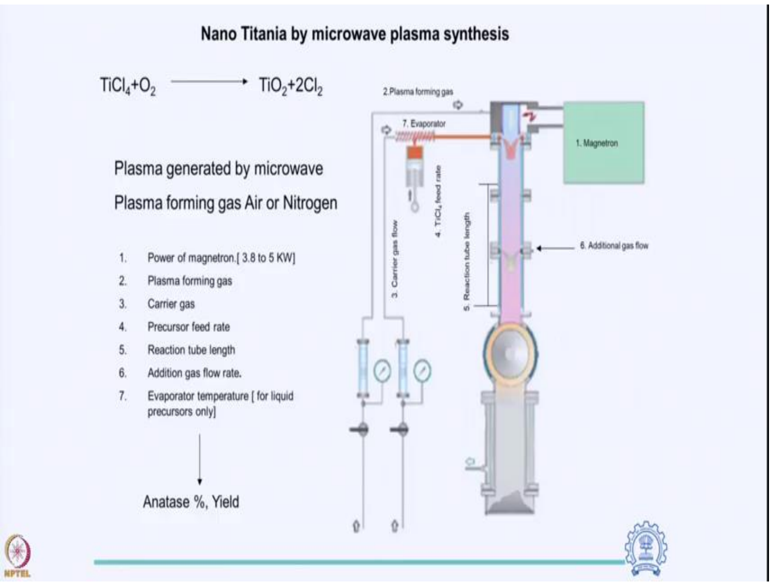
Then we will define the orthogonal design matrix, we will do the assignment of factors and interaction then we will have the experiment in their order, with the, then we will introduce replication and random order and we will discuss result and analysis.

(Refer Slide Time: 2:02)



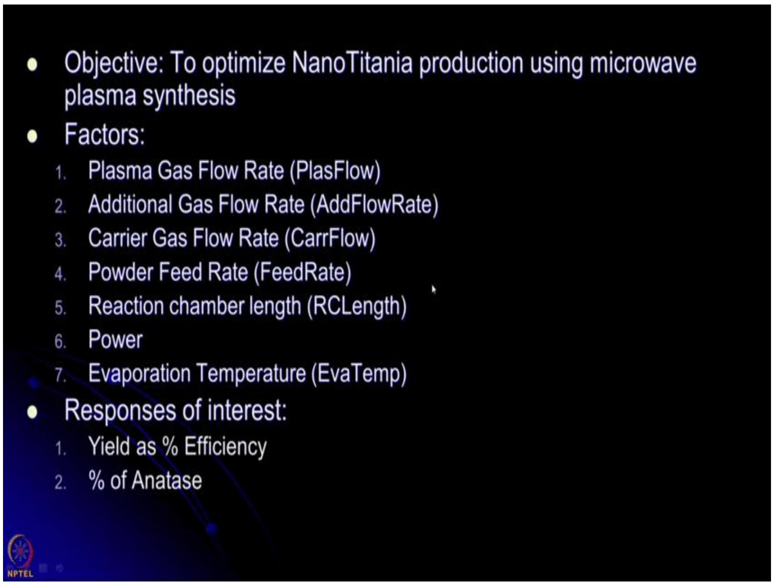
So, to recall we want to we are going through the case of micro plasma synthesis of Nano Titania powder. The process to produce Nano Titania needs to be optimized with respect to several input and output input parameters and the responses, production efficiency in percentage and percentage of anatase in Nano Titania powder. Based on the results, the commercial viability of the process we need to judge.

(Refer Slide Time: 2:28)



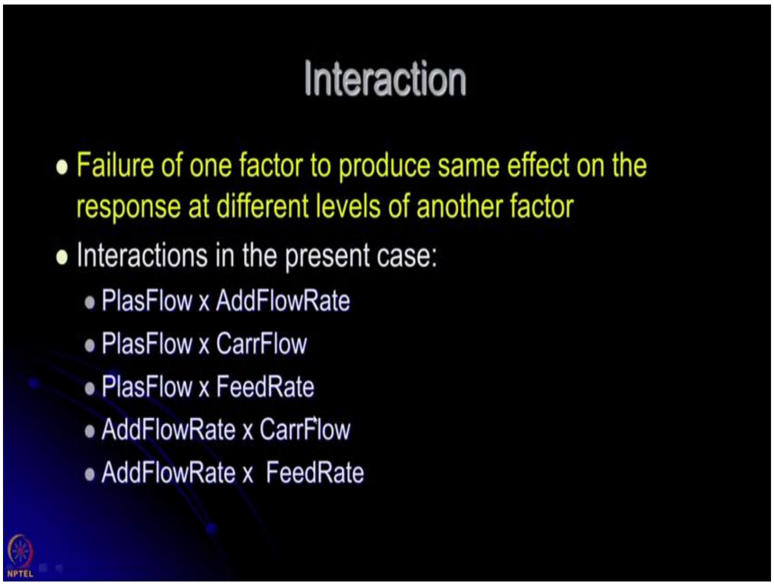
This is again a schematic of the microwave plasma synthesis. This gives the seven factors which we discussed in the previous session, which effect the production and the quality of Nano Titania produced in this particular system.

(Refer Slide Time: 2:48)



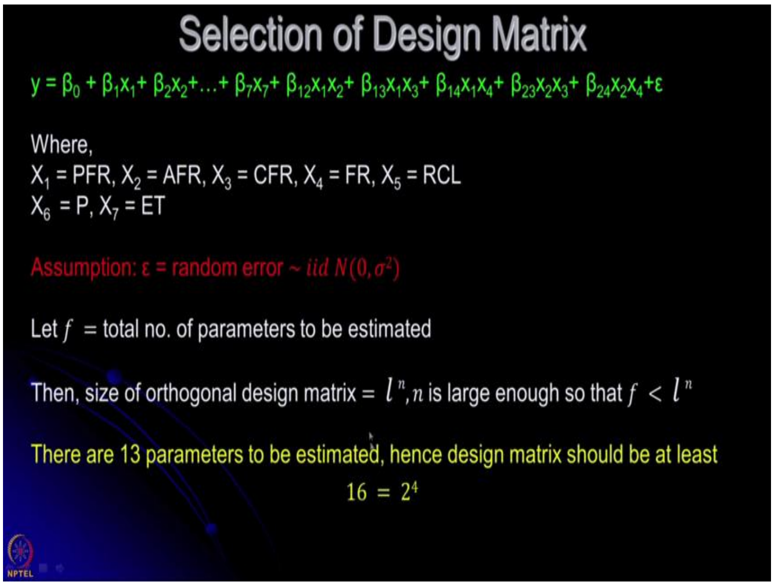
These factors are plasma gas flow rate, additional gas flow rate, carrier gas flow rate, powder feed rate, reaction chamber lengths, power of the magnetron and evaporating temperature of the precursor. The response of interest as we said is a yield as a percentage efficiency and percentage of anatase in the Nano Titania powder.

(Refer Slide Time: 3:24)



We also introduced interaction, we said that over and about these 7 factors there are 5 interactions which play a role. And we repeat interaction means that effect of one factor, the effect of one factor if it remains same, even though you change the another factor from level 1 to level 2 then you say that there is no interaction, but the effect of one factor does not remain same by changing another factor then there is an interaction between the two factors. So, these are the 5 interactions that have been identified.

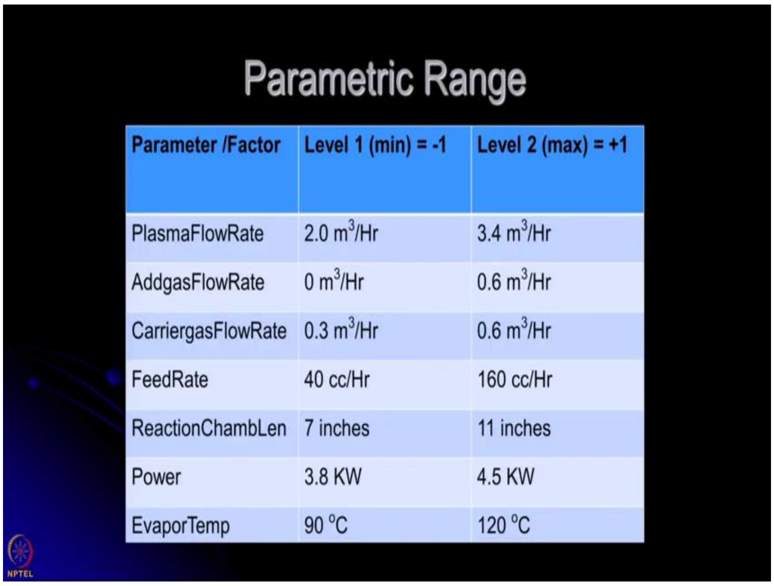
(Refer Slide Time: 4:03)



Then, we said that this is the model with the 5 interactions and 7 factors, we would like to estimate, this is a regression, multiple regression model and our assumption on the error epsilon is that it is a random error, it is independent and identically distributed with mean 0 and variance sigma square. If F is the total number of parameters to be estimated, then we said that the orthogonal design matrix is of the size l to the power n, so that we take n large enough to have F less than l to the power n.

So, here we have 13 parameters, the levels, l is the level at which the experiment is being conducted, so our level is only 2. So, we have to find a 2 to the power something, so that it is just above 13 which is 2 to the power 4. And therefore, our number of experimentation, I mean number of experiments we have to carry out, the size of design matrix is 16.

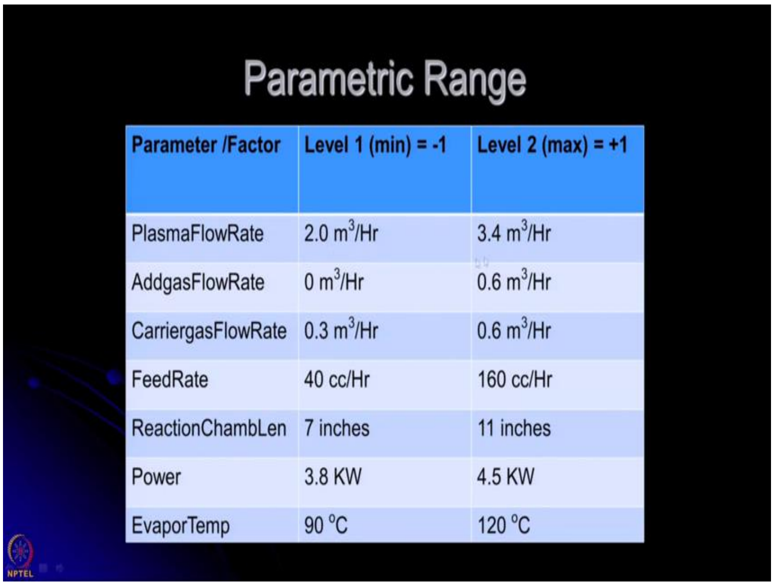
(Refer Slide Time: 5:16)



This is what we have done in the past. Now, we have 2 we said that we will have the experiment conducted at two levels. So, the question comes, what should be the parametric range? Generally, this parametric range is to be selected in such a way that as very beautifully said by Doctor Taguchi, “The experiment must perform”. Well, sounds very simple, but when you go through the experimentation, we have to realise that the kind of arrangements that we are going to, the kind of combinations we are going to have with different factors and their levels.

We must make sure that each experiment is going to happen. It is going to result into something, it should not be an obvious failure because then you already know that it is not going to perform, then why to conduct the experiment?

(Refer Slide Time: 6:28)



Parametric Range

Parameter /Factor	Level 1 (min) = -1	Level 2 (max) = +1
PlasmaFlowRate	2.0 m^3/Hr	3.4 m^3/Hr
AddgasFlowRate	0 m^3/Hr	0.6 m^3/Hr
CarriergasFlowRate	0.3 m^3/Hr	0.6 m^3/Hr
FeedRate	40 cc/Hr	160 cc/Hr
ReactionChambLen	7 inches	11 inches
Power	3.8 KW	4.5 KW
EvaporTemp	90 °C	120 °C

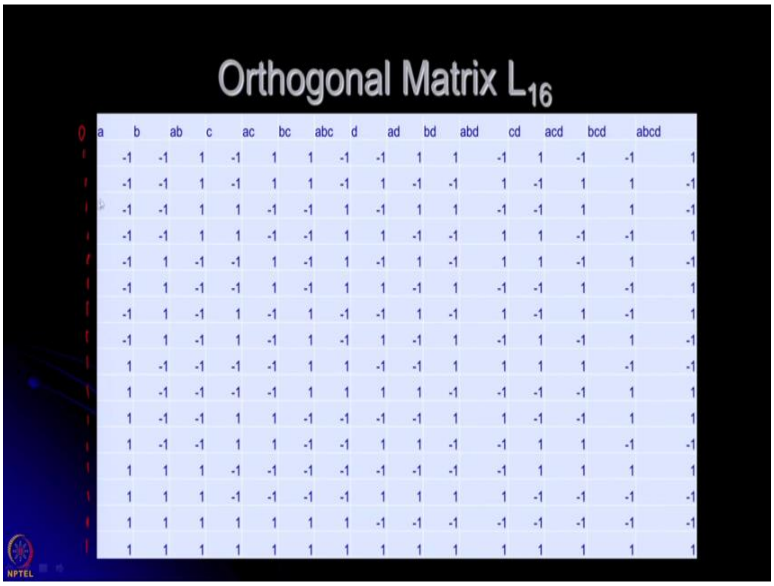
So, here this is one aspect that plasma flow rate, the range has to be large enough to cover the difference between the level 1, level minus 1 and level low and level high, which we are going to call level minus1 and level maximum, plus1, minimum and maximum. So, this levels have to be selected in such a way that it will show the difference in the effect of plasma glass flow rate.

It should not be so close that it has no effect on the plasma gas flow rate. And it should not be large that at some level the experiment may not perform or obviously it will give you result which you are not expecting. You already know that this is not going to give you the result of to your acceptance and therefore, there is no point selecting a range that will give you either no change because of the change in level from minimum to maximum.

It should not give you the very far away levels that is minimum to maximum distance not be so far away that you know that one of them is not going to give you the result that will be acceptable to you. So, it has to be optimum range in which you have to conduct the experiment.

And as I said, this comes only by doing some preliminary experimentation. So, accordingly the range has been selected for each of these 7 factors, so this is a parameter of factors, 7 of them, the level, minimum level is given here, maximum level is given here and I am going to denote this level here as minus1, minimum level as minus1 and maximum level as plus1.

(Refer Slide Time: 8:29)



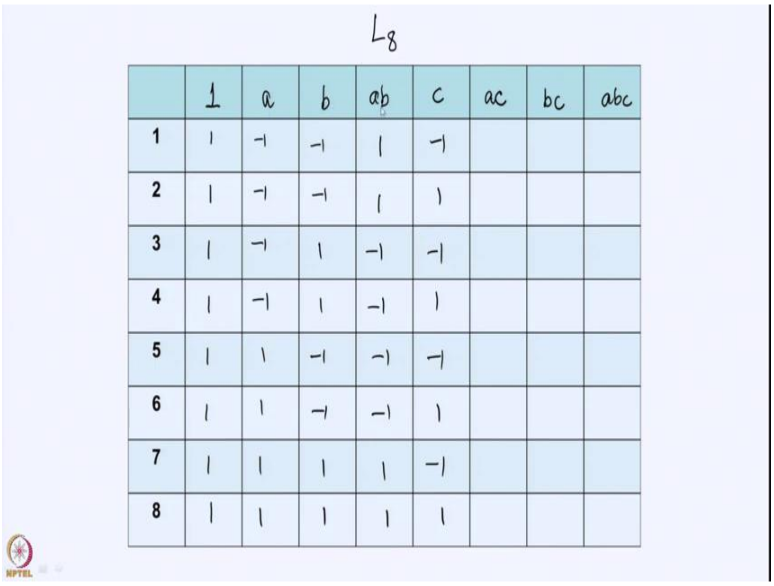
a	b	ab	c	ac	bc	abc	d	ad	bd	abd	cd	acd	bcd	abcd
-1	-1	1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	1
-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1
-1	-1	1	1	-1	-1	1	-1	1	1	-1	-1	1	1	-1
-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1
-1	1	-1	-1	1	-1	1	-1	1	-1	1	1	-1	1	-1
-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1
-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1
-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1
1	-1	-1	-1	-1	1	1	-1	-1	1	1	1	1	-1	-1
1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1
1	-1	-1	1	1	-1	-1	-1	-1	1	1	-1	-1	1	1
1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1
1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1
1	1	1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1
1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

So, my design matrix is, I am going to call it l 16 because it is a orthogonal matrix of size 16 and it looks something like this. It is all minus1’s or plus1’s which represents the minimum value, minimum level of the parameter and maximum level of the parameter and because it is an orthogonal matrix, these are, there are 15 columns to it, actually 1 column is missing and that is 1, 1, 1, 1, 1 that is called 0th column, so that makes it total 16 column, it is an orthogonal matrix.

Please remember, orthogonal matrix says that the columns are orthogonal to each other, so if you take a dot product of any column with any other column here, it should be 0, if you add the first column of 1, 1, 1, 1, 1 here, if you add the first column which I call a 0 column as 1, 1, 1, 1, 1, 1, 1 all 16 1’s, you will find that that is also orthogonal to e, all the other columns.

So, if you take the dot product of 0 with A, 0 with any column or you take, pick any two columns and you take a dot product then the dot product is 0 and therefore, this is called an orthogonal matrix of size 16 because it has a 16 column. The matrix I have shown as 15 column, the first column is obviously 1, 1, 1, 1, 1.

(Refer Slide Time: 10:13)



L_8

	1	a	b	ab	c	ac	bc	abc
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			

How do you make this? I thought I will just give you the trick, there is, it is nothing very great. Suppose, we want to make l18, we want to design l8. So, there are 8 rows to it and there are 1,2,3,4,5,6,7,8 columns to it. So, I begin, the first one I start is that I have 1,1,1,1,1,1,1,1 simple, this is my 0th column. Then I start with -1,-1,-1,-1, four -1's and four plus1's, this I call my column A.

Then I build a column B in which I have first two minus1's then plus1 then minus1 and then plus1. Then I build a column ab which is a multiplication of two columns a and b, so, I have -1,-1,-1,-1,1,1. Then I introduce another column c, which I say now, here I have done -1,-1,1,1, so I go -1,1,-1,1,-1,1,-1,1. Then, I introduce a column ac, I introduce a column a, sorry, let us correct.

Then I call, introduce a column bc and then I introduce a column abc. I think I will leave this column for you to fill up, but this is how you can construct an l8 or l16 or whatever you want, it is very easy to do. The first column is 1,1,1 which I call 0 or it is actually known as column 1, so let me put the right name to it, it is called a column 1, this is call column 1. Then you have a which is the, half of it I make -1, rest of it I make plus1. Then I introduce a column b in which, in which I have other half of it that is at, here there are four minus1's, so I make two minus1's and two plus1 then again two plus1 and plus2, plus1's.

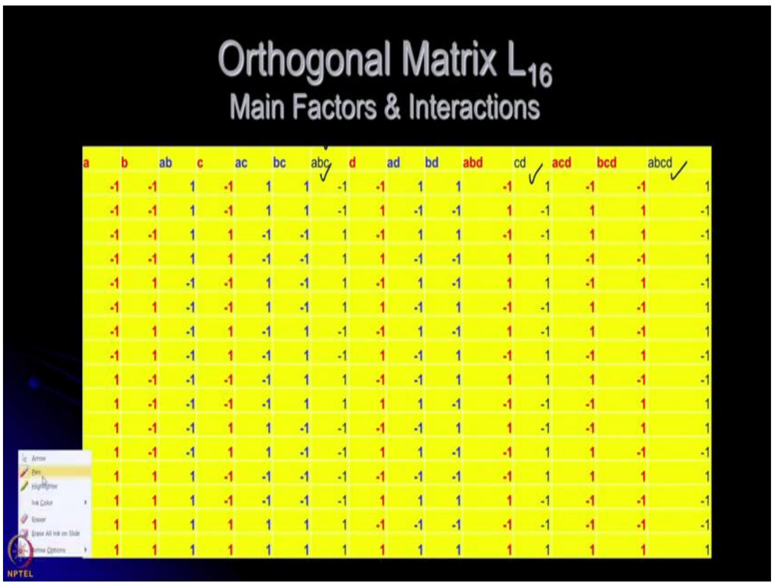
Then I take the multiplication of the two in the next column. Then I introduced another column which is minus1, 1, minus 1, 1 alternately. So, I am putting this minus1 and 1 in alternate fashion in different patterns and this pattern is very fixed. First, half of it is minus1, the rest half of it is 1 then half of half is minus1 and then 1,1,1 then again minus1,minus1 and 1,1 and then once my two such columns have come, I put multiplication, I introduce one column then I put the multiplication and that is how I create the l8. So, you can see that l16 is also done in the same way, I have put 8 of them as minus1 and rest of them are plus1.

Then b I have put four of them as minus1, four of them as plus1, again four of them as minus1, again this. And then I take a multiplication of a and b. Then I introduce C as minus1,minus 1,1,1,minus 1,minus 1,1,1 etc. Then I introduce ac, ab, abc. Then I introduce a d column which is again minus 1, 1 minus 1,1, minus alternate, so I get ad, bd, abd, cd, acd, bcd and abcd, so this is how it is done.

And in this, all the fresh factors that I introduced without multiplication are straight your assignment of main factor. Now, we have 7 main factors. So, we have assignment of only 4 . We still have to assign 3 more main factors, so this 3 more main factors , I am going to assign by asking myself a question that leave out the interactions of your interest and then assign this columns which do not have those interactions. Now, we are going to consider only two level interactions, we are not considering the abcd, abc, abd, acd, bcd, etc as interaction.

So, at 10 point, I have chosen the left most, the right most interactions to be there. So, I am taking abd, acd and bcd as also an independent factor, main factor of further our design and that is how make seven factor, main factors are assigned to this column.

(Refer Slide Time: 16:15)

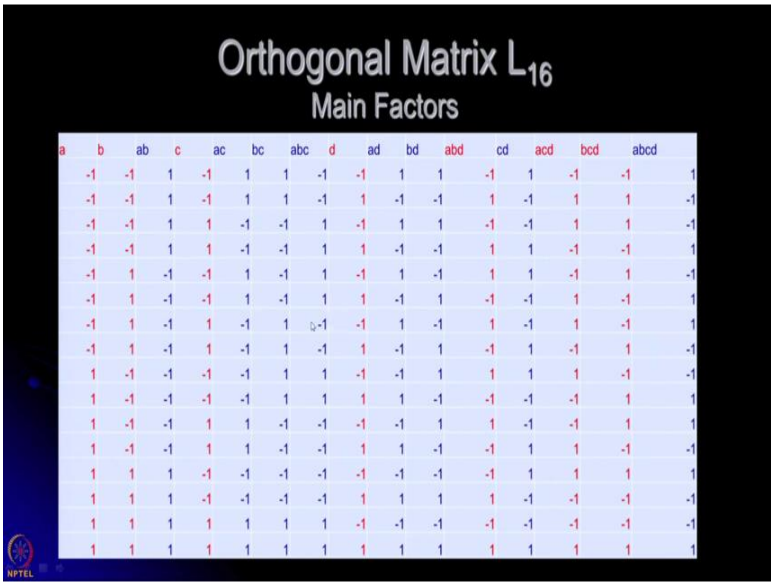


a	b	ab	c	ac	bc	abc	d	ad	bd	abd	cd	acd	bcd	abcd
-1	-1	1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	1
-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1
-1	-1	1	1	-1	-1	1	-1	1	1	-1	-1	1	1	-1
-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1
-1	1	-1	-1	1	-1	1	-1	1	-1	1	1	-1	1	-1
-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1
-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1
-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1
1	-1	-1	-1	-1	1	1	-1	-1	1	1	1	1	-1	-1
1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1
1	-1	-1	1	1	-1	-1	-1	-1	1	1	-1	-1	1	1
1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1
1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1
1	1	1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1
1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

So, according to this assignment, then I assign, I actually take the multiplication, assign those columns to the interaction. So that is how I find the interaction ab, ac, bc, ad and bd as my interaction, this dark blue ones that I have shown here are my 5 interactions of interest. Red ones are my main factors, dark blue ones are my interactions and this light blue three columns are unassigned columns, so, these are my, sorry, these are my unassigned columns, three of them.

They also have a role to play in future, so that is why I am indicating them to you. So, this is how I do my factor assignment.

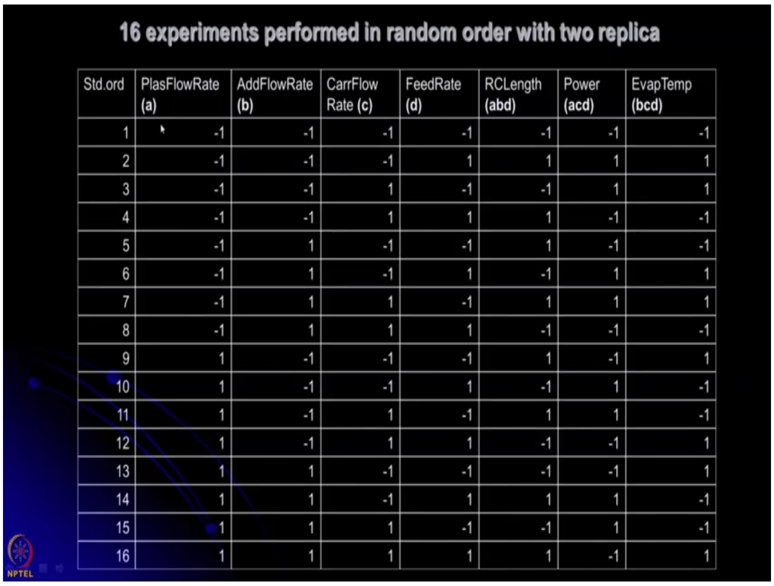
(Refer Slide Time: 17:15)



a	b	ab	c	ac	bc	abc	d	ad	bd	abd	cd	acd	bcd	abcd
-1	-1	1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	1
-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1
-1	-1	1	1	-1	-1	1	-1	1	1	-1	-1	1	1	-1
-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1
-1	1	-1	-1	1	-1	1	-1	1	-1	1	1	-1	1	-1
-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1
-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1
-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1
1	-1	-1	-1	-1	1	1	-1	-1	1	1	1	1	-1	-1
1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1
1	-1	-1	1	1	-1	-1	-1	-1	1	1	-1	-1	1	1
1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1
1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1
1	1	1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1
1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

So, if you look at my orthogonal matrix, I have done the assignment of main factors and interactions.

(Refer Slide Time: 17:28)



16 experiments performed in random order with two replica

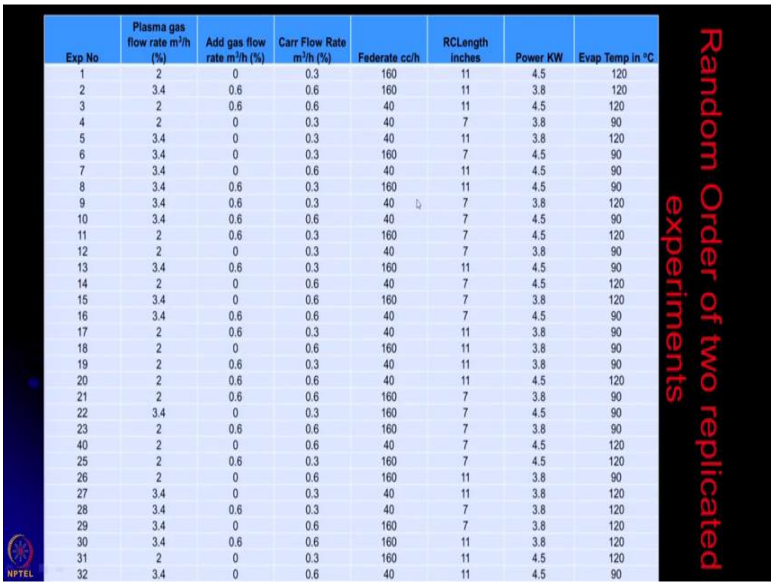
Std.ord	PlasFlowRate (a)	AddFlowRate (b)	CarrFlow Rate (c)	FeedRate (d)	RCLength (abd)	Power (acd)	EvapTemp (bcd)
1	-1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	1	1	1	1
3	-1	-1	1	-1	-1	1	1
4	-1	-1	1	1	1	-1	-1
5	-1	1	-1	-1	1	-1	-1
6	-1	1	-1	1	-1	1	1
7	-1	1	1	-1	1	1	1
8	-1	1	1	1	-1	-1	-1
9	1	-1	-1	-1	1	-1	1
10	1	-1	-1	1	-1	1	-1
11	1	-1	1	-1	1	1	-1
12	1	-1	1	1	-1	-1	1
13	1	1	-1	-1	-1	-1	1
14	1	1	-1	1	1	1	-1
15	1	1	1	-1	-1	1	-1
16	1	1	1	1	1	-1	1

And now if you look at it, interactions come automatically. Remember, the interactions are coming by multiplication, so they will come automatically. So, the main factors which I have assigned are kept here and this is actually my 16 experiments that I need to perform in two replica, in two replica.

So, here this is my plasma flow rate which will vary in this fashion, additional gas flow rate will vary in this fashion, etc. So, if I look at the standard order of my experiment, my first experiment will have plasma gas flow rate at minimum, additional gas flow rate, everything will be at minimum. The second experiment we will have plasma flow rate, additional gas flow rate, carrier gas flow rate at minimum and rest of them will be at a maximum level or you pick up any other number.

Tenth experiment is going to be plasma flow rate, plasma gas flow rate at maximum, additional gas flow rate at minimum, carrier gas flow rate at minimum, feed rate at maximum, my reaction chamber length as the smallest one, the power is going to be maximum and the evaporating temperature is going to be the lower limit. So, this is how my 16 experiments are now designed.

(Refer Slide Time: 19:02)



Exp No	Plasma gas flow rate m³/h (%)	Add gas flow rate m³/h (%)	Carr Flow Rate m³/h (%)	Federate cc/h	RCLength inches	Power KW	Evap Temp in °C
1	2	0	0.3	160	11	4.5	120
2	3.4	0.6	0.6	160	11	3.8	120
3	2	0.6	0.6	40	11	4.5	120
4	2	0	0.3	40	7	3.8	90
5	3.4	0	0.3	40	11	3.8	120
6	3.4	0	0.3	160	7	4.5	90
7	3.4	0	0.6	40	11	4.5	90
8	3.4	0.6	0.3	160	11	4.5	90
9	3.4	0.6	0.3	40	7	3.8	120
10	3.4	0.6	0.6	40	7	4.5	90
11	2	0.6	0.3	160	7	4.5	120
12	2	0	0.3	40	7	3.8	90
13	3.4	0.6	0.3	160	11	4.5	90
14	2	0	0.6	40	7	4.5	120
15	3.4	0	0.6	160	7	3.8	120
16	3.4	0.6	0.6	40	7	4.5	90
17	2	0.6	0.3	40	11	3.8	90
18	2	0	0.6	160	11	3.8	90
19	2	0.6	0.3	40	11	3.8	90
20	2	0.6	0.6	40	11	4.5	120
21	2	0.6	0.6	160	7	3.8	90
22	3.4	0	0.3	160	7	4.5	90
23	2	0.6	0.6	160	7	3.8	90
40	2	0	0.6	40	7	4.5	120
25	2	0.6	0.3	160	7	4.5	120
26	2	0	0.6	160	11	3.8	90
27	3.4	0	0.3	40	11	3.8	120
28	3.4	0.6	0.3	40	7	3.8	120
29	3.4	0	0.6	160	7	3.8	120
30	3.4	0.6	0.6	160	11	3.8	120
31	2	0	0.3	160	11	4.5	120
32	3.4	0	0.6	40	11	4.5	90

I do the replication, so I actually perform 32 experiments, 16 times 16, I do twice. And this is not my standard order, this is my random order. So, you can see that the actual experiment that I am going to perform first is going to have these values of these parameters. So, plasma gas flow rate, we have given it in the units, so it is going to be 2, the additional gas flow rate to be kept at 0, carrier gas flow rate is to be kept at 0.3 m cube by that meter cube by hour feed rate by cc by hour 160, RC length has to 11 inches, power has to be 4.5 kilowatt and evaporating temperature has to be 120 degree centigrade, this is my first experiment to be performed.

Remember, what we gave is a standard order. Standard order is what you get from the designed experiment. Experimental order is a random order, it is a random order as shown as written here, it is a random order and the 32 experiments will be performed in this random order.

(Refer Slide Time: 20:28)

Exp order vs. Std Order

Exp No	Std.ord	Exp No	Std.ord
12	1	4	17
1	2	31	18
40	3	14	19
18	4	26	20
19	5	17	21
25	6	11	22
3	7	20	23
21	8	23	24
27	9	5	25
6	10	22	26
32	11	7	27
29	12	15	28
9	13	28	29
8	14	13	30
16	15	10	31
30	16	2	32

So, if you want to look at what is the experimental order and standard order, this is how it is. The standard order first experiment will be performed 12th, the second order experiment will be, second standard order experiment will be performed first, etc, etc.

So, like this the 32 of them have been given the, the this is randomized order and this is the standard order and the experiment will be performed in this randomized order.

(Refer Slide Time: 21:02)

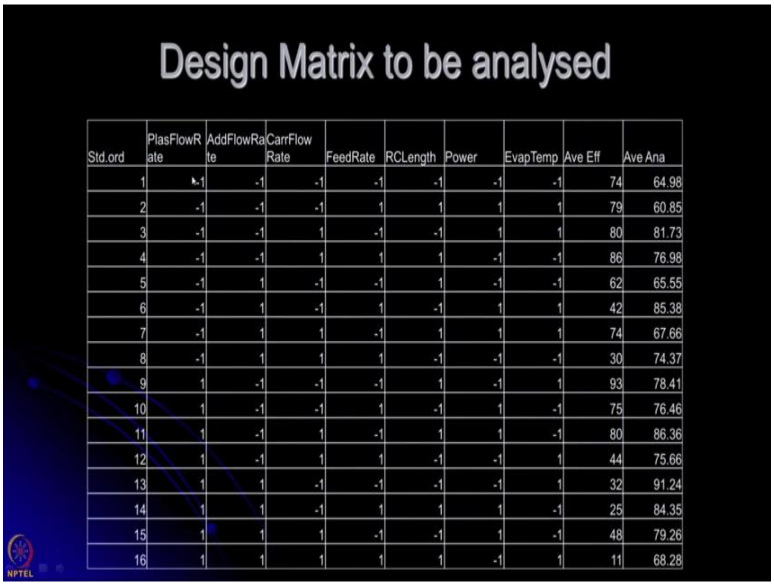
Results

Std.ord	Eff (%)	% Ana	Std.ord	Eff (%)	% Ana	Ave Eff	Ave Ana
1	74	68.46	17	73	61.49	74	64.98
2	86	65.08	18	72	56.62	79	60.85
3	84	83.5	19	76	79.96	80	81.73
4	83	78.9	20	88	75.05	86	76.98
5	60	82.47	21	64	48.62	62	65.55
6	46	86.96	22	37	83.79	42	85.38
7	59	65.66	23	89	69.65	74	67.66
8	28	73.51	24	32	75.22	30	74.37
9	97	89.77	25	89	67.04	93	78.41
10	73	73.66	26	76	79.26	75	76.46
11	81	86.58	27	79	86.13	80	86.36
12	32	78.93	28	56	72.38	44	75.66
13	35	93.98	29	29	88.5	32	91.24
14	23	85.61	30	27	83.09	25	84.35
15	56	82.89	31	39	75.63	48	79.26
16	10	77.79	32	11	58.76	11	68.28

Suppose the experiments have been performed and we have got the results, so in standard order we put them down. So, this is percentage efficiency and percentage anatase. The first 16 experiment and this is replicated, another 16 experiment. And therefore, we take effect as an average effect of percentage efficiency of the first standard order and then the 17th standard order is actually a first standard order, so that is going to be 73, so we have taken average, we have taken care of the, there is no decimal point there, so we cannot have decimal points recording them, suitably they have been rounded of, so these are the average value.

So, let us repeat. These are 32 experiments, the first 16 are first 16 experiment. This is the next 16 replicated experiment, for each one of them we are finding the results. They are actually the same experiments and therefore, we are taking average of it as a final experimental result as a average effect and average anatase percentage.

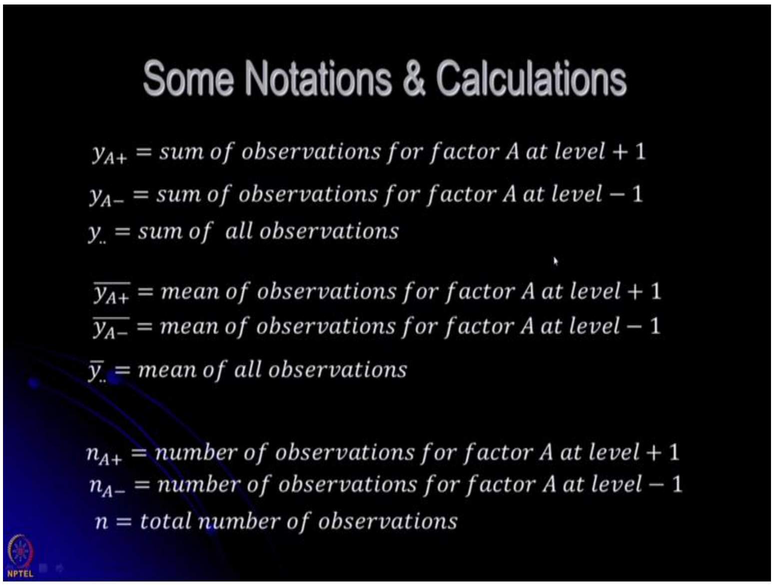
(Refer Slide Time: 22:22)

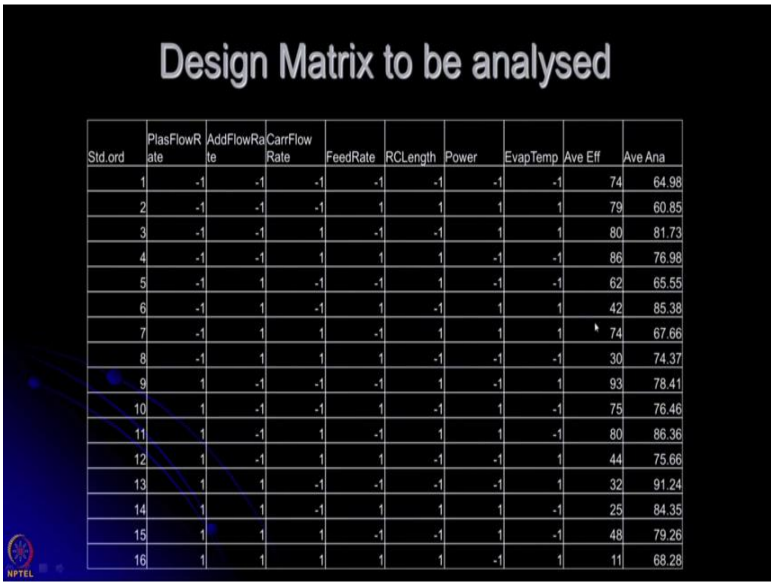

Design Matrix to be analysed

Std.ord	PlasFlowRate	AddFlowRate	CarrFlow Rate	FeedRate	RCLength	Power	EvapTemp	Ave Eff	Ave Ana
1	-1	-1	-1	-1	-1	-1	-1	74	64.98
2	-1	-1	-1	1	1	1	1	79	60.85
3	-1	-1	1	-1	-1	1	1	80	81.73
4	-1	-1	1	1	1	-1	-1	86	76.98
5	-1	1	-1	-1	1	-1	-1	62	65.55
6	-1	1	-1	1	-1	1	1	42	85.38
7	-1	1	1	-1	1	1	1	74	67.66
8	-1	1	1	1	-1	-1	-1	30	74.37
9	1	-1	-1	-1	1	-1	1	93	78.41
10	1	-1	-1	1	-1	1	-1	75	76.46
11	1	-1	1	-1	1	1	-1	80	86.36
12	1	-1	1	1	-1	-1	1	44	75.66
13	1	1	-1	-1	-1	-1	1	32	91.24
14	1	1	-1	1	1	1	-1	25	84.35
15	1	1	1	-1	-1	1	-1	48	79.26
16	1	1	1	1	1	-1	1	11	68.28

So, the design matrix that we have to analyse looks like this, this is a standard order, this are the levels at which your different 7 parameters have been organized and these are the results that you have obtained.

(Refer Slide Time: 22:34)





Design Matrix to be analysed

Std.ord	PlasFlowRate	AddFlowRate	CarrFlow Rate	FeedRate	RCLength	Power	EvapTemp	Ave Eff	Ave Ana
1	-1	-1	-1	-1	-1	-1	-1	74	64.98
2	-1	-1	-1	1	1	1	1	79	60.85
3	-1	-1	1	-1	-1	1	1	80	81.73
4	-1	-1	1	1	1	-1	-1	86	76.98
5	-1	1	-1	-1	1	-1	-1	62	65.55
6	-1	1	-1	1	-1	1	1	42	85.38
7	-1	1	1	-1	1	1	1	74	67.66
8	-1	1	1	1	-1	-1	-1	30	74.37
9	1	-1	-1	-1	1	-1	1	93	78.41
10	1	-1	-1	1	-1	1	-1	75	76.46
11	1	-1	1	-1	1	1	-1	80	86.36
12	1	-1	1	1	-1	-1	1	44	75.66
13	1	1	-1	-1	-1	-1	1	32	91.24
14	1	1	-1	1	1	1	-1	25	84.35
15	1	1	1	-1	-1	1	-1	48	79.26
16	1	1	1	1	1	-1	1	11	68.28

Let us work out some notations and calculations. We say that, this is y is your response. So, y sub A plus means that sum of all observations for factor A at level plus1. So, if you look at here, I look at all factor A, suppose plasma flow rate, it is kept at level 1 then I have to take the average of 9 to 16 values of the efficiency and that average is what I call y sub A plus then I have y sub A minus, it means that for example, if I take additional flow rate then the first 4 and the experiments 9 to 12 refer to the additional gas flow rate kept at a minimum value or at minus1.

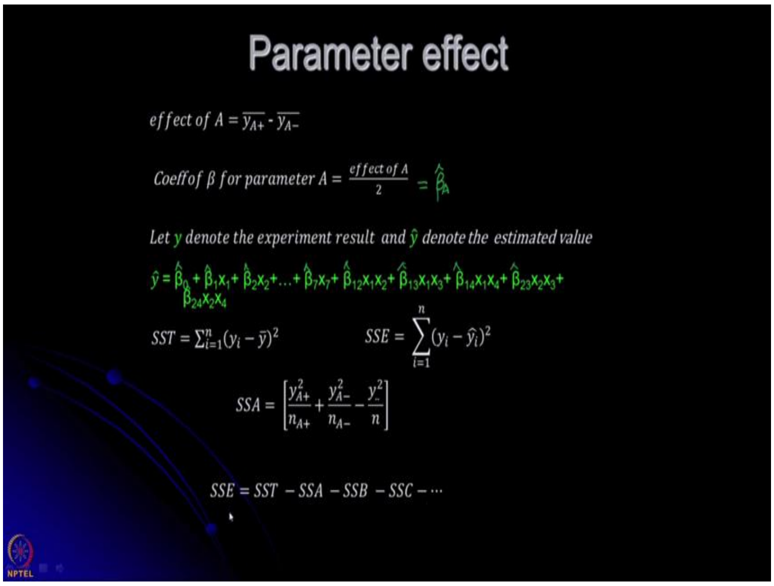
So, you accordingly take the first, average of, first 4 and the 9 to 12th experiment, that is going to be your y sub A minus, which is the, A is the factor of your interest.

At factor level minus1 this is going to be y sub A minus1 and y dot dot, please recall the notations we have done in the past, sum of all observations, so sum of all of this is y dot dot, you remember that this is y, this two are actually y, your response y1 which is efficiency, percentage efficiency and this is percentage anatase.

You take an average because there are 2 levels, sorry, at each level you take an average, so you divide it by number of observations you have added, so in this case you would have added 8, 8 experiments because there are 8 experiment conducted at level minus1 and 8 experiment conducted at level plus1. And therefore, this will be average of 8 observations. So, average of yA plus is mean observation for factor A at level plus1, so it is this divided by 8, and why dot dot bar is mean of all observation, so this will be divided by 16.

Similarly, you have number of observations at factor at level plus1 and A minus is the number of factors at level minus1 and n is the total number of observations.

(Refer Slide Time: 25:25)



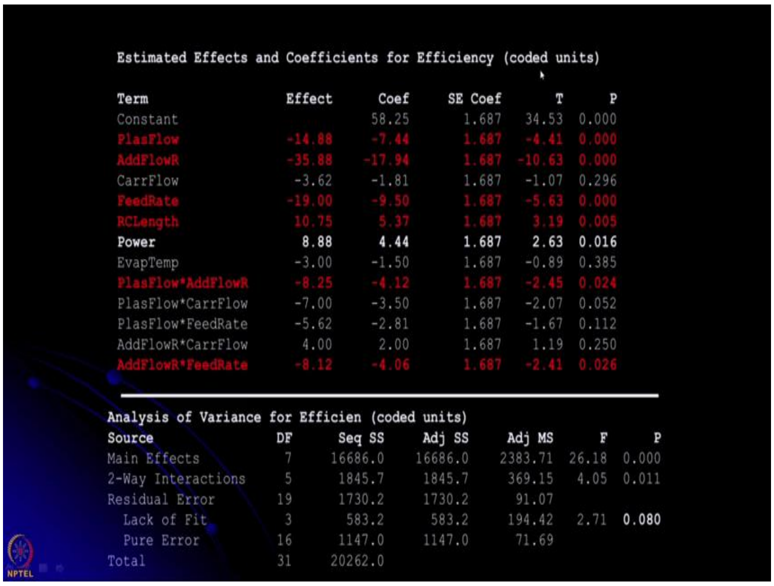
So, effect of any parameter you can find out if you take average that is mean of observations at factor level A plus minus the mean of observation at factor level A minus, it is called effect of A. Coefficient of B of a parameter A is nothing but an average because there are two levels, so, it is averaged out at 2. This is called coefficient beta, so this is the estimator, estimate of beta hat or we used to call it b in the past. I must make correction here, ink colour I make it green, this should be all hats, these are all estimators that we are calculating.

So, if y denotes the experiment result and y hat denotes the estimated value, then y hat is this. You remember, you please see that I have removed plus epsilon because now we are calculating the estimated value by putting in the estimated value of beta hat. Coefficient of beta hat is estimated here, so as such this is actually your beta hat A, beta hat for factor A is given here, so this is what it is.

Recall what we did in analysis of variance, so we have a total sums of squares which is summation of yi minus y bar whole square, this is an average of all the observations. And summation, sum of squares of errors is summation of yi minus yi hat whole square, yi hat is here, so it is the actual value minus the estimated value whole square.

Sum of squares due to factor A, sum of squares due to this factor A is given by this formula, it is given by this formula and you can show that total sums of squares is actually sum of squares due to each factor plus sum of squares due to error or this formula we have already worked out to describe the chi square distribution, it goes in the same way.

(Refer Slide Time: 28:11)



Estimated Effects and Coefficients for Efficiency (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		58.25	1.687	34.53	0.000
PlasFlow	-14.88	-7.44	1.687	-4.41	0.000
AddFlowR	-35.88	-17.94	1.687	-10.63	0.000
CarrFlow	-3.62	-1.81	1.687	-1.07	0.296
FeedRate	-19.00	-9.50	1.687	-5.63	0.000
RCLength	10.75	5.37	1.687	3.19	0.005
Power	8.88	4.44	1.687	2.63	0.016
EvapTemp	-3.00	-1.50	1.687	-0.89	0.385
PlasFlow*AddFlowR	-8.25	-4.12	1.687	-2.45	0.024
PlasFlow*CarrFlow	-7.00	-3.50	1.687	-2.07	0.052
PlasFlow*FeedRate	-5.62	-2.81	1.687	-1.67	0.112
AddFlowR*CarrFlow	4.00	2.00	1.687	1.19	0.250
AddFlowR*FeedRate	-8.12	-4.06	1.687	-2.41	0.026

Analysis of Variance for Efficien (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	7	16686.0	16686.0	2383.71	26.18	0.000
2-Way Interactions	5	1845.7	1845.7	369.15	4.05	0.011
Residual Error	19	1730.2	1730.2	91.07		
Lack of Fit	3	583.2	583.2	194.42	2.71	0.080
Pure Error	16	1147.0	1147.0	71.69		
Total	31	20262.0				

Now, this whole thing is put in a table, so let us read this table. These are the estimated effects of the parameters in coded unit, by what do I mean by coded unit? It is not the actual values with which it is calculated, it is calculated with level minimum as minus1 and maximum as plus1. So, our design matrix does not have the actual value of the plasma gas flow rate or additional gas flow rate, but it has an coded value of plasma gas flow rate as minus1 if it takes a minimum level and it is plus1 if it takes a maximum level for each factor.

So, this shows the effect as we have calculated, this shows the coefficient. This is standard error of coefficient, remember standard error of coefficient is residual error divided by total number of experiments that is, that will be, it is a mean error, so it is divided by the total number of experiment we have performed and its square root, so this is what it is. So, this is what is standard error of coefficient you can just confirm it. This is T statistic to test that this beta coefficient is equal to 0 or not.

The beta coefficient of plasma gas flow rate that is this beta 1 is great, is 0 or not. The null hypothesis is that beta 1 is equal to 0 and to test that null hypothesis this is the T statistic and this is the P value, you remember when we did the hypothesis testing, we also calculated the P value of the test, so these are the P value. In other words, if we want to test the hypothesis at 95 percent confidence or 5 percent significance, then these values have to be, if they are larger than 0.05 then the null hypothesis that the coefficient of that particular factor is 0 is rejected.

So, this shows that these are all values which are smaller than that, so these are all the significant values. I think I made a mistake in what I said, if these value, if P values are smaller than 0.05 then you are going to reject the null hypothesis that they are 0, if it is greater than 0.05 then you are going to accept it as a null hypothesis.

So here for example, we find that evaporating temperature is, has a beta value, has a P value which is larger than 0.05 and therefore, the we accept the null hypothesis that the coefficient for evaporating temperature is 0. So, it is not having any effect on your efficiency while the red ones which are clearly shown, we can say it has a effect on the, it has an effect on the efficiency of the process, microwave plasma synthesis process for producing Nano Titania.

So, plasma gas flow rate, additional gas flow rate, feed rate, reaction chamber length then interaction plasma flow rate with additional gas flow rate and interaction of additional gas flow rate with feed rate, these are all significant factors which are having an effect on efficiency of the system. All others have their beta value close to 0, so they can be ignored. When beta is not equal to, when null hypothesis that beta i is equal to 0 is rejected, it means that those factors are playing an important role.

(Refer Slide Time: 32:45)

Estimated Effects and Coefficients for Efficiency (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		58.25	1.687	34.53	0.000
PlasFlow	-14.88	-7.44	1.687	-4.41	0.000
AddFlowR	-35.88	-17.94	1.687	-10.63	0.000
CarrFlow	-3.62	-1.81	1.687	-1.07	0.296
FeedRate	-19.00	-9.50	1.687	-5.63	0.000
RCLength	10.75	5.37	1.687	3.19	0.005
Power	8.88	4.44	1.687	2.63	0.016
EvapTemp	-3.00	-1.50	1.687	-0.89	0.385
PlasFlow*AddFlowR	-8.25	-4.12	1.687	-2.45	0.024
PlasFlow*CarrFlow	-7.00	-3.50	1.687	-2.07	0.052
PlasFlow*FeedRate	-5.62	-2.81	1.687	-1.67	0.112
AddFlowR*CarrFlow	4.00	2.00	1.687	1.19	0.250
AddFlowR*FeedRate	-8.12	-4.06	1.687	-2.41	0.026

Analysis of Variance for Efficien (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	7	16686.0	16686.0	2383.71	26.18	0.000
2-Way Interactions	5	1845.7	1845.7	369.15	4.05	0.011
Residual Error	19	1730.2	1730.2	91.07		
Lack of Fit	3	583.2	583.2	194.42	2.71	0.080
Pure Error	16	1147.0	1147.0	71.69		
Total	31	20262.0				

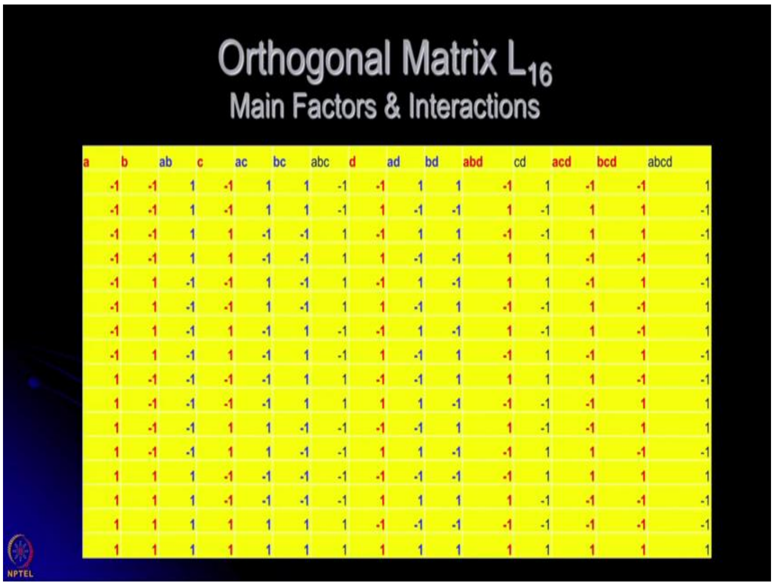
This is a table, sorry, this is a table for analysis of variance for efficiency. These are the 7 main effects, so this all 7 of them are added, all the 7 main effects are added, it has a 7 degrees of freedom. This is sums of squares, sequential, we have not gone through this study, but this is an adjusted sums of squares. The sums of squares divided by degrees of freedom gives you a mean square error.

Similarly, two way interactions, there are 5 degrees of freedom, you have 5 of them and it has a adjusted mean square as this, the residual error comes out is here which is shown here, this is the error sums of squares, so residual error is error sums of squares divided by its degrees of freedom which comes to this.

And if you, this F statistic, please recall, is the main effect means sums of squares divided by the residual sums of squares which gives you the F statistic which we did in the analysis of variance. And here are the P values calculated in a similar manner and if you look at these P values, it actually tells you that both the factors, they are the P values are smaller than alpha, it means that both the factors are actually significant.

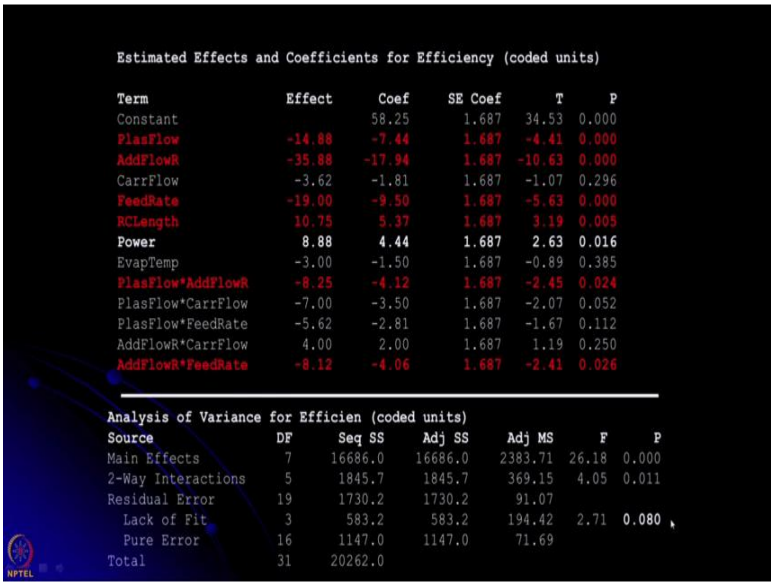
So, main effects in two way interactions are significant as we can see here and then we calculate what is called lack of feed. You remember the 3 columns that we left out, they are calculated here that is a part of the residue, as it comes. So, pure error is if you take out this lack of feed.

(Refer Slide Time: 34:48)



You remember, if you feed whole 16 by 16 matrix then there is no lack of feed. You have fitted the whole equation, but instead you have only fitted 13 of them and you have left out the 3 of them.

(Refer Slide Time: 35:08)

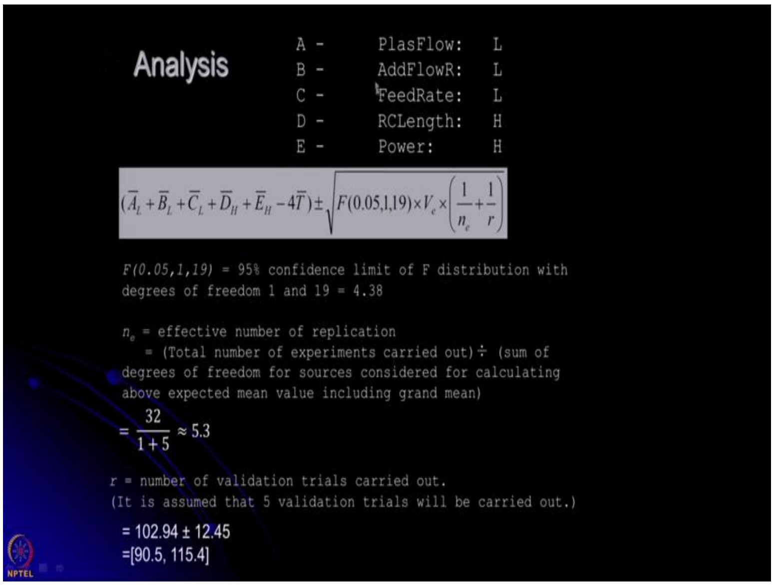


So, if you calculate this the lack of feed, it shows that the lack of feed is significant. It means that the coefficient of lack of feed is not significant, it is 0. So, there is no lack of feed here, this is a pure error and then you have total of this because you have total sums of squares, which you has been shown here, this is total sums of squares.

So, there are total 32 experiments minus 1 degree of freedom, so it comes to 31 degrees of freedom and this is there and this is your analysis of variance table. So, this very clearly says that these are important and there is no lack of fit in your model. Your model completely define, decides and defines, gives you the value of efficiency.

So, then we come to the final analysis, if you look at back here, the effect is negative, it means that the effect is the, lower the plasma flow rate higher the efficiency, lower the additional gas flow rate higher the efficiency, lower the feed rate higher the efficiency, higher the reaction chamber length higher the efficiency.

(Refer Slide Time: 36:38)



So, likewise we find that the plasma flow rate has to be kept at level low, feed rate should be kept at level low, additional gas flow rate should also be at low, but RC length that is reaction chamber length should be high and the power should be high and then this becomes your predictive interval.

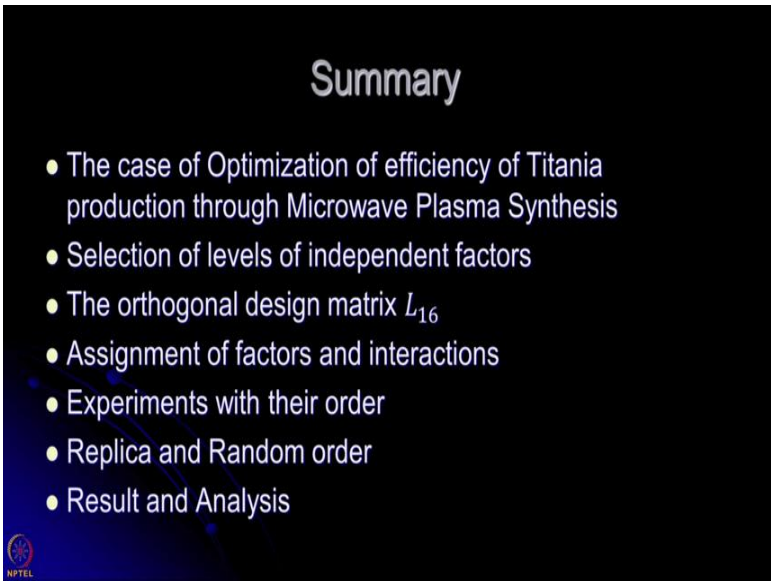
It means that you take the average value of factor plasma flow rate when keeping it at low and then add up all of these, minus you have to subtract 4 times the total average, because you see from each one of them you have to remove the average and then you have to remove one more average, so add the one more average and therefore it comes to minus40 and this becomes your predictive interval.

You remember we had done the, in regression analysis what is your predictive interval. This is F, this shows the 95 degree confidence limit of F distribution with degrees of freedom 1 and

19, why 1 and 19? Because you are taking the 19 degrees of freedom and it is only 1 equation, so it is coming 1 and 19, ne is an effective number of replication which can be calculated as total number of experiment carried out divided by number of degrees of freedom from the sources considered for calculating these means.

So, we calculated these mean by the 5 and 1 total, the grand mean. And therefore, it is 32 by 1 plus 5 which comes to 5.3 and then we say that the average efficiency should lie in the interval 90.5 to 150.4.

(Refer Slide Time: 38:52)



The question is, is everything okay? And I leave the matter here and we go to the next session, but in this, let us quickly summarize. We took the case of optimizing the efficiency of Titania production. We went through process of selection of levels of independent factors, selection of orthogonal matrix l16, assignment of factors and interaction, experiments with their random order then we did the experiments with replica and a random order. Finally, we did the result and analysis. Thank you.