Quality Engineering, 12(3), 387–393 (2000)

EXPERIMENTAL SEQUENCE: A DECISION STRATEGY

V. Roshan Joseph*
SQC & OR Unit
Indian Statistical Institute
8th Mile Mysore Road
Bangalore–560059, India

Key Words

Experimental runs; Randomization; Design of experiment; Control factor; Noise factor; Orthogonal array; Optimal sequence; Traveling Salesman Problem.

Introduction

In an experiment, the sequence in which the experimental runs are to be conducted is usually decided in a purely chance manner. This random sequence increases the likelihood that the effect of uncontrollable variables not considered in the experiment will balance out. It also helps in the application of statistical tests of significance for the effects of factors under study (1). However, in many experiments, it becomes difficult to conduct the experiment using a random sequence, as it can increase the level changes of some factors that are difficult to adjust. In such cases, the experimental sequence is obtained using restricted randomization (2). The restrictions on randomization can cause some of the uncontrollable variables to get confounded with factors of interest. Still, restrictions on randomization are being used when the reduction in the cost of conducting the experiments becomes worthier than a little loss of information.

In split-plot designs, the hard-to-vary factors form the whole plots, whereas the easy-to-vary factors are run in subplots (3). This helps in reducing the changeovers necessary for the difficult factors during the experimentation. Using the concepts of split-plot designs, Taguchi (4) has grouped the columns of orthogonal arrays into different groups where hard-to-vary factors are assigned to the lower-order groups and easy-to-vary factors are assigned to the higher-order groups. The above methods can be applied easily when the number of hard-to-vary factors in the experiment is small. When there are many hard-to-vary factors and fractionated designs are used for experiments, the above methods will not always give optimum results.

In this article, we develop a methodology to obtain the experimental sequence which minimizes the cost of adjusting the factor levels during experimentation. The methodology is explained using an industrial example. The optimal sequence is not a random sequence and, hence, it has to be used carefully. The general guidelines for selecting between an optimal sequence and a random sequence for a given experiment are also discussed.

An Example

Printed circuit boards (PCBs) are used to mount electronic components and to interconnect them electrically. In the hot-air leveling (HAL) stage of PCB manufacturing, a thin layer of solder coating is given to the holes and the pads of the PCB. Nonuniformity of solder-coating thickness on the pads was the main quality problem faced at this stage. It was decided to conduct a designed experiment to optimize the parameter levels at the HAL process to improve solder-thickness distribution. In this article, we explain the approach adopted in obtaining the sequence of experimental runs which will minimize the cost of conducting the experiment.

The HAL Process

The panel is loaded to a Vertical HAL machine by clamping at one edge. The panel is then dipped into the solder bath for a specified period of time and then raised out of the bath. During the upward motion, hot air is impinged on the panel to clear the holes and to level the solder.

There were eight control factors and a noise factor that were included in the experiment. Control factors are those variables in the process that engineers can specify by nominal values and maintain cost effectively. Noise factors are those variables in the process that are either uncontrollable or too expensive to control. Noise factors are the causes of variation in the output of the process. The objective of the experiment is to find out the control factor levels that will make the process robust against the noise factors.

Control Factors and Levels

The eight control factors selected for the HAL process experimentation with their levels are given in Table 1. The existing levels of these factors are those marked with an asterisk.

Noise Factor and Levels

The variation of copper level in the solder bath is the major noise factor in the HAL process. As the panels are processed, the copper level increases gradually in the solder bath. The copper level is reduced once a day by removing a part of the copper from the bath and adding fresh solder to it. It is too expensive to control the copper level at a constant value and, hence, its variation becomes a noise to the process. From the usual variation of the copper level in the bath, a low value and a high value were selected for experimentation.

Experimental Layout

The L_{18} ($2^1 \times 3^7$) orthogonal array (4) was selected as the control array. Because the air knife angle in the back side has two levels, it was assigned to the 2-level column and the other seven factors were assigned to the 3-level columns. The experiment is to be conducted with low and high copper levels in the solder bath. The layout of the experiment is given in Table 2. The columns of L_{18} to which the control factors are assigned are also shown in parentheses.

Cost of Adjusting Factor Levels

The cost of adjusting the level of a factor includes the costs due to the loss of production during the adjustment period and also the cost for doing the adjustment. The last four control factors, Pressure Front and Back, Up Speed, and Dwell Time, are very easy to adjust, as they require only turning some knobs. Thus, their adjustment costs are taken to be zero. The first four control factors are difficult to adjust and take about 5–15 min to change from one level to another. It was difficult to obtain an exact estimate of the cost of adjusting each one of the control factors. Discussions with the operators resulted in arranging these factors in an ascending order of difficulty of adjustment: Distance—Front, Distance—Back, Angle—Front, and Angle—Back. Adjusting the air knife angle in the back side was felt to be

Table 1. Control Factors and Levels

SERIAL NO.	CONTROL FACTORS	NOTATION	LEVEL 1	LEVEL 2	LEVEL 3
1	Air knife angle—front (deg)	A	0	3*	6
2	Air knife angle—back (deg)	B	6*	9	—
3	Air knife distance—front (mm)	C	1.6	2.4*	3.2
4	Air knife distance—back (mm)	D	1.6	2.4*	3.2
5	Air pressure—front (bar)	E	2.0	2.5*	3.0
6	Air pressure—back (bar)	F	2.5	3.0*	3.5
7	Up speed (time) (s)	G	1.2	1.7*	2.2
8	Dwell time (s)	H	3.5	4.0*	4.5

Note: Asterisk denotes existing levels.

Table 2. Layout of the Experiment

EXP. NO.	B (1)	A (2)	C (3)	D (4)	E (5)	F (6)	G (7)	H (8)	LOW Cu LEVEL	HIGH Cu LEVEL
1	1	1	1	1	1	1	1	1		
2	1	1	2	2	2	2	2	2		
3	1	1	3	3	3	3	3	3		
4	1	2	1	1	2	2	3	3		
5	1	2	2	2	3	3	1	1		
6	1	2	3	3	1	1	2	2		
7	1	3	1	2	1	3	2	3		
8	1	3	2	3	2	1	3	1		
9	1	3	3	1	3	2	1	2		
10	2	1	1	3	3	2	2	1		
11	2	1	2	1	1	3	3	2		
12	2	1	3	2	2	1	1	3		
13	2	2	1	2	3	1	3	2		
14	2	2	2	3	1	2	1	3		
15	2	2	3	1	2	3	2	1		
16	2	3	1	3	2	3	1	2		
17	2	3	2	1	3	1	2	3		
18	2	3	3	2	1	2	3	1		

Note: Number in parentheses indicates the columns of L_{18} to which the control factors are assigned.

twice as difficult as that of the distance of air knife in the front side. Hence, a cost of 1 is assigned to the Distance—Front and a cost of 2 is assigned to the Angle—Back. The costs for Distance—Back and Angle—Front were taken as 1.5 each. Thus, the costs for changing a level of the eight control factors are $C_A = 1.5$, $C_B = 2$, $C_C = 1$, $C_D = 1.5$, $C_E = 0$, $C_F = 0$, $C_G = 0$, and $C_H = 0$. Note that in some cases, the cost of adjusting a factor will depend on its level. For example, increasing the concentration in a chemical bath may be easier than decreasing its concentration. Adjusting the copper level in the solder bath takes about 3 h and is extremely difficult to do. Hence, it was decided to conduct the 18 runs of experiment at 1 level of copper and then to repeat it at the other level of copper.

From Table 1, the existing levels of the factors A, B, C, D, E, F, G, H are (2, 1, 2, 2, 2, 2, 2, 2). Denote this by the run number 0. The total cost of changing the levels from the existing setup to experimental run 1 can be calculated as $1.5 + 0 + 1 + 1.5 + 0 + 0 + 0 + 0 = 4$. Similarly, the cost of changing from one run to another run can be evaluated. Note that if two or more factors can be adjusted simultaneously or if adjusting one of them disturbs the setting of the other, then the total cost will not be the sum of individual costs of adjustment. Let c_{ij} be the cost of changing the setup from the ith run to the jth run. The c_{ij}'s are calculated and are given in Figure 1.

Optimal Sequence

To conduct the experiment, the existing setup has to be changed to a setup in an experimental run and then complete the runs one by one. Because conducting the experiment under both levels of copper on the same day can affect seriously the day's production target, it was decided to carry out the experiment with a low level of copper on one day and with a high level of copper on another day. Hence, after the completion of 18 runs, the setup has to be reverted to the existing setup. Thus, the problem is to find out a sequence of experimental runs starting with 0 and moving through 1, 2, . . . , 18 once and only once and return to 0 with minimum cost. This problem is similar to the Traveling Salesman Problem (TSP) in the field of combinatorial optimization. Here, the experimental runs are the "cities" and c_{ij} is the "distance" between the ith and jth cities. Standard algorithms and heuristics are available for solving TSP. Refer to Ref. 5 for more details on TSP. The optimal sequence of the experimental runs, which minimizes the total cost is 0-2-1-4-9-17-11-10-16-14-15-13-12-18-7-8-3-6-5-0. The cost of conducting the experiment with this sequence is 40.

The only major noise factor in the HAL process is the copper level in the solder bath. In 1 day, about 1000 panels are processed in HAL; hence, the change in copper level during the processing of 18 panels (1 panel per experiment)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	*	4	1.5	4	2.5	0	2.5	2.5	3	4	6	5	4.5	3	3.5	4.5	6	5	4.5
1	4	*	2.5	2.5	1.5	4	4	3	4	2.5	3.5	3	4.5	5	6	4.5	5	4.5	6
2	1.5	2.5	*	2.5	4	1.5	4	2.5	3	4	4.5	3.5	3	4.5	5	6	6	5	4.5
3	4	2.5	2.5	*	4	4	1.5	4	2.5	3	3	4.5	3.5	6	4.5	5	4.5	6	5
4	2.5	1.5	4	4	*	2.5	2.5	3	4	2.5	5	4.5	6	3.5	4.5	3	5	4.5	6
5	0	4	1.5	4	2.5	*	2.5	2.5	3	4	6	5	4.5	3	3.5	4.5	6	5	4.5
6	2.5	4	4	1.5	2.5	2.5	*	4	2.5	3	4.5	6	5	4.5	3	3.5	4.5	6	5
7	2.5	3	2.5	4	3	2.5	4	*	2.5	2.5	5	6	4.5	3.5	6	6	3.5	4.5	3
8	3	4	3	2.5	4	3	2.5	2.5	*	2.5	4.5	5	6	6	3.5	6	3	3.5	4.5
9	4	2.5	4	3	2.5	4	3	2.5	2.5	*	6	4.5	5	6	6	3.5	4.5	3	3.5
10	6	3.5	4.5	3	5	6	4.5	5	4.5	6	*	2.5	2.5	3	2.5	4	1.5	4	4
11	5	3	3.5	4.5	4.5	5	6	6	5	4.5	2.5	*	2.5	4	3	2.5	4	1.5	4
12	4.5	4.5	3	3.5	6	4.5	5	4.5	6	5	2.5	2.5	*	2.5	4	3	4	4	1.5
13	3	5	4.5	6	3.5	3	4.5	3.5	6	6	3	4	2.5	*	2.5	2.5	3	4	2.5
14	3.5	6	5	4.5	4.5	3.5	3	6	3.5	6	2.5	3	4	2.5	*	2.5	2.5	3	4
15	4.5	4.5	6	5	3	4.5	3.5	6	6	3.5	4	2.5	3	2.5	2.5	*	4	2.5	3
16	6	5	6	4.5	5	6	4.5	3.5	3	4.5	1.5	4	4	3	2.5	4	*	2.5	2.5
17	5	4.5	5	6	4.5	5	6	4.5	3.5	3	4	1.5	4	4	3	2.5	2.5	*	2.5
18	4.5	6	4.5	5	6	4.5	5	3	4.5	3.5	4	4	1.5	2.5	4	3	2.5	2.5	*

Figure 1. (c_{ij}) matrix.

can be neglected. Hence, it is not required to use a random sequence for the runs of this experiment. Suppose a random sequence 0-10-5-16-2-15-7-11-6-17-3-13-8-12-4-18-1-14-9-0 were used; then, the cost would have been 112, which is about three times more than the cost with an optimal sequence.

The general formulation for obtaining the optimal sequence of experimental runs in a design of experiment is given in the Appendix.

Assignment of Factors to Columns

The optimal sequence for the HAL experiment was obtained for the design of experiment shown in Table 2. Because it is possible to make several designs for the same problem, it is important to know whether the design has any effect on the optimum cost or not. The factor Angle–Back (B) is a 2-level factor; hence, it is assigned to the first column of L_{18}. The other factors can be assigned to the remaining columns in an arbitrary manner. Because only three other factors, A, C, and D, have an impact on the cost of conducting the experiment, we can study the designs by assigning these factors to any of the seven 3-level columns. It is possible to make $7 \times 6 \times 5 = 210$ different designs. Because $C_A = C_D$, we need to study only 105 of them. Of the 105 possible designs, the assignment A → 7, C → 5, and D → 8 gives the minimum cost for conducting the experiment. The optimal sequence is 0-2-16-12-14-17-10-15-18-11-13-9-5-1-8-4-3-7-6-0 and the minimum cost is 32. We see that there is a saving of eight units compared to the minimum cost of the experimental design considered earlier. Suppose that the factors A, C, and D were assigned to the columns 3, 7, and 8, respectively; then, the minimum cost would have been 76.5.

In Taguchi's approach (4), the first column is in group1, the second column is in group2, and columns 3–7 are in group3. Now, the factors are assigned based on the degree of difficulty in changing their levels. The hard-to-vary factors are assigned to the lower-order groups and the easy-to-vary factors are assigned to the higher-order groups. The design of experiment in Table 2 is one such design. Carrying out this design with a sequence 0-1-2-3-4-·····-18-0 will result in a cost of 59. We see that Taguchi's approach led to a cost almost twice the optimum cost of the problem.

These examples show that the design of experiment has an effect on the minimum cost of conducting an experiment. Thus, when a fractionated design is to be used, one has to study the minimum cost of conducting the experiment at the design stage of the experiment itself to get maximum benefit. This can be done by finding the minimum cost for all possible designs and choosing the design that gives the minimum of minimum costs.

Some Special Cases

In experiments, particularly in the manufacturing stage, factors might be selected from different subsystems of the process. Consider, for example, a process with two subsystems (say solder mask developing and HAL) where three factors are selected from subsystem I and four factors are selected from subsystem II, as shown in Figure 2. If the process is not continuous, then a different sequence for the experimental runs may be used in subsystem II than in subsystem I. Hence, two optimal sequences are to be found; one using the columns of F_1, F_2, and F_3 in the design of experiment for subsystem I and another using the columns of F_4, F_5, F_6, and F_7 in the design of experiment for subsystem II. Thus, if there are many subsystems, an optimal sequence for each subsystem must be found using the factors in the subsystem and the respective columns in the design of experiment.

In some experimental design problems, some of the factors may depend on the level taken by another factor. For example, an experimenter wishes to study two methods of processing (say Vertical HAL and Horizontal HAL) where three factors in each method are selected for experimentation as in Figure 3. The processing method and the other six factors can be assigned to an orthogonal array using branching design (4). Here, an optimal sequence for the experimental runs corresponding to method 1 and another optimal sequence for the experimental runs corresponding to method 2 can be obtained by considering the respective columns of the factors in the design of experiment.

Suppose two machines are available for doing an experiment where there is no machine-to-machine difference. Then, some of the experimental runs may be performed on one of the machines and the rest on the other machine. Here, the experimenter wishes to find out which of the runs are to be performed in which machine and in what sequence so that the total cost of conducting the experiment is a minimum. This problem is similar to a TSP with two salesmen (6).

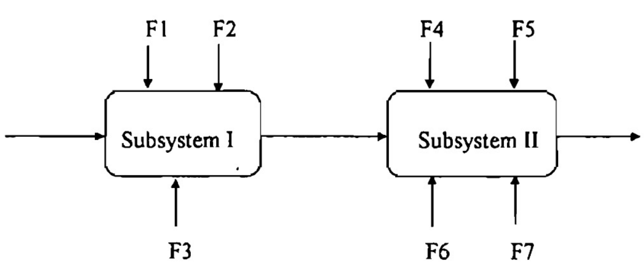

Figure 2. Case with two subsystems.

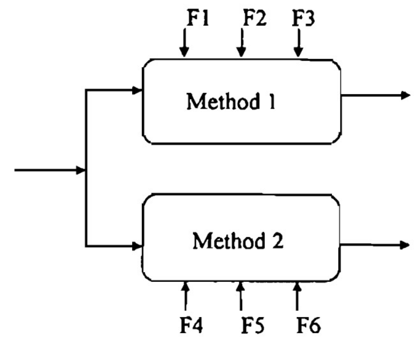

Figure 3. Case with branching designs.

Evaluation of Benefit

To evaluate the benefit of using an optimal sequence for experimental runs, the minimum cost can be studied in a design of experiment with different possible arrangements of factors. In this study, L_{18} orthogonal array is chosen, as it is the most commonly used orthogonal array in robust design projects. Let us not consider any outer array and take the cost of changing a factor level as 1 for all the factors. Let c_{ij} be the sum of individual costs of changing the factor levels. Here, c_{ij} represents the number of adjustments required in the factor levels for changing from the ith run to the jth run. Let Z_{min} be the minimum cost and Z_{max} be the maximum cost for conducting the experiment. Z_{max} will represent the worst possible case in using a random sequence.

The optimal sequence of experimental runs can be obtained by finding the shortest Hamiltonian path in the network of 18 nodes, as explained in the Appendix. Z_{min}, Z_{max}, and one of the optimal sequences for different arrangements of factors in the L_{18} orthogonal array are given in Table 3. We see that significant savings can be obtained by using an optimal sequence compared to a random sequence in a design of experiment.

Randomization or Optimization

Randomization is done on the experimental units and on the experimental runs to minimize the effect of noise factors on the model factor effects. Because the optimal sequence is not a random sequence, we may not be able to average out the effect of noise factors that vary over time. Hence, optimal sequences for the experimental runs are to be used only when the noise factors are adequately taken care of during the experiment. Fowlkes and Creveling (7) discuss in detail developing good noise strategies for experiments.

The choice between randomization and optimization is

Table 3. Results of Some L_{18} Orthogonal Array Designs

NO. OF FACTORS	COLUMNS ASSIGNED	OPTIMAL SEQUENCE	Z_{min}	Z_{max}
3	1, 2, 3	1-2-3-6-5-4-7-8-9-18-17-16-13-14-15-12-11-10	17	47
4	1, 2, 3, 4	1-4-5-2-3-6-8-7-9-15-11-17-14-10-16-13-12-18	28	67
5	1, 2, 3, 4, 5	1-4-15-12-2-5-13-10-3-6-14-11-17-9-7-18-16-8	36	73
6	1, 2, 3, 4, 5, 6	1-11-17-13-12-18-14-10-16-15-4-9-3-6-8-2-5-7	51	82
7	1, 2, 3, 4, 5, 6, 7	1-4-2-5-3-6-7-16-8-17-9-18-11-14-10-13-12-15	68	100
8	1, 2, 3, 4, 5, 6, 7, 8	1-12-18-15-17-14-16-13-11-3-10-2-9-6-8-5-7-4	85	102

to be made depending on the problem faced by the experimenter. The broad guidelines for selecting between an optimal sequence and a random sequence can be formulated as follows.

Case 1: All major noise factors that vary over time are known. In this case, an optimal sequence for the experimental runs shall be applied.

Case 2: All major noise factors that vary over time are not known. In this case, one has to choose randomization for experimental runs.

In Case 2, if some factors are hard to vary, then restricted randomization can be used. This can also be done as follows. Assign a cost zero to those factors that are easy to vary. This will lead to more than one optimal solution to the experimental sequence. The number of optimal sequences will depend on the number of hard-to-vary factors, their levels, and the total number of experimental runs. Then, choose one of the optimal sequences randomly. The restricted randomization is a strategy between the complete randomization and complete optimization. The degree of the randomization/optimization required for a given experiment shall be decided by weighing the advantages that can be gained by a random sequence against the cost benefits of using an optimal sequence.

Data Analysis

The restrictions on randomization imposed by the use of an optimal sequence for experimental runs prevent the usual analysis of variance F-test from being a meaningful test for significance. Taguchi (4) does not advocate the use of the statistical test of significance in the analysis of experiments. In the data analysis technique devised by him, the F ratio is used only to rank the model factor effects in the order of importance; see Ref. 8. Hence, Taguchi's technique should be used for analyzing the data when an optimal sequence is used for the experiment.

Conclusions

There are situations where the cost of conducting the experiment can be reduced significantly by optimizing the experimental runs. The optimal sequence for an experiment is obtained by solving a Traveling Salesman Problem for which standard algorithms are available. It is also demonstrated that the reduction in cost can be improved by studying the optimal sequences at the design stage of the experiment itself.

Appendix: Mathematical Formulation

Consider an experiment with n runs. In general, the setup before the experiment need not be any one of the experimental setups. Denote the existing setup by 0. The experimenter has to change the existing setup to one of the experimental setups and then complete the experimental runs one by one. After the completion of the experiment, one may have to return to the original setup, as there is a time lag associated with data collection and analysis for finding the optimum levels. Thus, the problem faced by the experimenter is to find the sequence starting with 0 and moving through $1, 2, \ldots, n$ once and only once and return to 0, with minimum cost. This is a Traveling Salesman Problem (TSP) with $n + 1$ cities. Note that visiting a node once is equivalent of completing all the replicates in the experimental run. Also, if an experimental setup is repeated many times in the design of experiment (like in central composite designs), all of them can be grouped as one node in the network of experimental runs.

Let

$$x_{ij} = \begin{cases} 1 & \text{if the } j\text{th run succeeds the } i\text{th run} \\ 0 & \text{otherwise} \end{cases}$$

and c_{ij} is the cost of changing from the ith run to the jth run, $i, j = 0, 1, 2, \ldots, n$

An integer programming formulation of the problem is given as follows (refer to Ref. 5):

$$\text{Minimize} \quad Z = \sum_{i=0}^{n} \sum_{j=0}^{n} c_{ij} x_{ij}$$

$$\text{Subject to} \quad \sum_{j=0}^{n} x_{ij} = 1 \qquad \text{for all } i = 0, 1, \ldots, n$$

$$\sum_{i=0}^{n} x_{ij} = 1 \qquad \text{for all } j = 0, 1, \ldots, n$$

$$\sum_{i \in S} \sum_{j \in S} x_{ij} \leq |S| - 1 \qquad \text{for every proper nonempty subset of } S \text{ of } \{0, 1, \ldots, n\}$$

$$x_{ij} = 0, 1$$

where $|S|$ is the cardinality of the set S.

If the existing setup is one of the experimental runs, then the above problem reduces to an n-city TSP. In case of large experiments, it may not be possible to run the whole experiment at a stretch and the setup may have to be changed to the existing setup more than once. This situation can be modeled as a Vehicle Routing Problem with one depot (5). The need for changing the existing setup to an experimental run does not arise in many experiments at the design stage, in lab experiments, and so forth. Here, the problem is equivalent to finding the shortest Hamiltonian path (5) in a network with nodes 1, 2, . . . , n. This problem can be converted to an $(n + 1)$-city TSP by adding a node 0 to the network with $c_{0j} = c_{j0} = 0$ for all $j = 1, 2, \ldots, n$.

Acknowledgments

I am grateful to Professor B. K. Pal, Indian Statistical Institute, for his suggestions during the preparation of this article. I am thankful to Dr. Donald S. Holmes for the valuable comments, which helped in improving the content of this article.

References

1. Cox, D. R., *Planning of Experiments*, John Wiley & Sons, New York, 1958.
2. Lorenzen, T. J. and Anderson, V. L., *Design of Experiments: A No Name Approach*, Marcel Dekker, Inc., New York, 1993.
3. Montgomery, D. C., *Design and Analysis of Experiments*, 4th ed., John Wiley & Sons, New York, 1997.
4. Taguchi, G., *System of Experimental Design*, UNIPUB/Kraus International Publications and American Supplier Institute, Inc., White Plains, NY, 1987, Vols. 1 and 2.
5. Lawler, E. L., Lenstra, J. K., Rinnooy Kan, A. H. G., and Shmoys, D. B. (eds.), *The Traveling Salesman Problem*, John Wiley & Sons, Chichester, 1985.
6. Rao, M. R., A Note on the Multiple Traveling Salesman Problem, *Oper. Res.*, 28, 628–632 (1980).
7. Fowlkes, W. Y. and Creveling, C. M., *Engineering Methods for Robust Product Design*, Addison-Wesley, Reading, MA, 1998.
8. Phadke, M. S., *Quality Engineering Using Robust Design*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

About the Author: V. Roshan Joseph is a faculty member in the Statistical Quality Control and Operations Research Unit of Indian Statistical Institute, Bangalore. He holds a master's degree in quality, reliability, and operations research and a bachelor's degree in production engineering and management. He is a consultant to various industries in the field of statistical process control, design of experiments, and optimization techniques.