# Tuning Parameters In Heuristics By Using Design Of Experiments Methods

## Arif Arin; Ghaith Rabadi; Resit Unal

*Department of Engineering Management and Systems Engineering, Old Dominion University*
*aarin@odu.edu*

**Abstract.** With the growing complexity of today's large scale problems, it has become more difficult to find optimal solutions by using exact mathematical methods. The need to find near-optimal solutions in an acceptable time frame requires heuristic approaches. In many cases, however, most heuristics have several parameters that need to be "tuned" before they can reach good results. The problem then turns into "finding best parameter setting" for the heuristics to solve the problems efficiently and timely. One-Factor-At-a-Time (OFAT) approach for parameter tuning neglects the interactions between parameters. Design of Experiments (DOE) tools can be instead employed to tune the parameters more effectively. In this paper, we seek the best parameter setting for a Genetic Algorithm (GA) to solve the single machine total weighted tardiness problem in which n jobs must be scheduled on a single machine without preemption, and the objective is to minimize the total weighted tardiness. Benchmark instances for the problem are available in the literature. To fine tune the GA parameters in the most efficient way, we compare multiple DOE models including 2-level ($2^k$) full factorial design, orthogonal array design, central composite design, D-optimal design and signal-to-noise (S/N) ratios. In each DOE method, a mathematical model is created using regression analysis, and solved to obtain the best parameter setting. After verification runs using the tuned parameter setting, the preliminary results for optimal solutions of multiple instances were found efficiently.

## 1. INTRODUCTION

One of the most important effects of the improving modern sciences and technologies is to enable us understand and model real life problems realistically and in more details. The natural outcome of this fact is the rapid increase of dimensions and complexity of the problems. With the growing complexity of today's large scale problems, it has become more difficult to find optimal solutions by using only exact mathematical methods. Due to the concern of efficiency in terms of the solution quality, the need to find near-optimal solutions in an acceptable time frame requires using heuristic approaches.

Heuristics are quite new approaches in the field of combinatorial optimization. A heuristic can be defined as "a generic algorithmic template that can be used for finding high quality solutions of hard combinatorial optimization problems" [1]. Heuristic approaches have already proved themselves in many large scale optimization problems by offering near-optimal solutions where there is no optimal solution found by other approaches. In many cases, however, most heuristics have several parameters that need to be "tuned" before they can reach good results. The accepted values of the parameters to be employed in the heuristics have considerably significant impact on both solution process and the solution itself. To obtain the best results, the problem then turns into "finding the best parameter setting" for the heuristics to solve the problems efficiently and timely, which becomes an optimization problem by itself.

There are various methods used to find the best parameter setting in the literature. One-Factor-At-a-Time (OFAT) approach for parameter tuning is one of them; however, it neglects the interactions between the parameters that might change the whole solution process and quality of solution. Particularly, in terms of the interactions, Design of Experiments (DOE) methods are promising approaches and can be easily employed to tune the parameters more effectively.

In this paper, we seek the best parameter setting for a genetic algorithm to solve the single machine total weighted tardiness problem in which n jobs must be scheduled on a single machine without preemption, and the objective is to minimize the total weighted tardiness. Benchmark instances for the single machine total weighted tardiness problem are available in the literature.

## 2. DESIGN OF EXPERIMENTS (DOE)

To fine tune the genetic algorithm parameters in the most efficient way, we compare multiple DOE tools including 2-level ($2^k$) full factorial design, orthogonal array design, central composite design, D-optimal design and signal-to-noise (S/N) ratios method. In each DOE method, a mathematical model is created using regression analysis, and solved to obtain the best parameter setting. After verification runs for other benchmark instances by using the tuned parameter setting,

DOE methods presented will be compared in terms of their solution qualities.

The single machine total weighted tardiness problem is used in this paper as a difficult problem to demonstrate the use of DOE for setting the optimization Genetic Algorithm (GA) parameters. In this problem, n jobs must be scheduled on a single machine where each job j has a given processing time $P_j$ and a due date $d_j$. The tardiness $T_j$ is defined as max $(0, C_j-d_j)$ where $C_j$ is the job's completion time – a decision variable that is based on the job sequence. The objective function then becomes to minimize $\sum_{j=1}^{n} wjTj$. This is a well known problem to which benchmark problems are available. In seeking best parameter setting for the GA, we will be using a MS-excel Add-in called Evolver from Palisade [6].

We first implemented the problem in Excel spreadsheet, and used the first instance of 40-job benchmark problem to compare different DOE methods that are discussed below. The upper and lower levels for the GA parameters are given in the Table 1.

**Table 1:** Upper and lower levels for parameters

| Level | Crossover Prob. (A) | Mutation Prob. (B) | Population Size (C) |
|---|---|---|---|
| Lower | 0.01 | 0.06 | 30 |
| Upper | 1 | 0.2 | 100 |

The GA stopping criteria are to run for 10 minutes or to stop whenever the percent deviation of the solution from the optimal solution/best solution found so far becomes 0. In the following sections, we discuss and compare five DOE methods to see which method performs best.

### 2.1. 2-Level ($2^k$) Full Factorial Design

2-Level ($2^k$) full factorial design is the one of the most widely used DOE tools. In $2^k$ full factorial design, k is the number of factors. After the lower and upper levels of the factors are determined, all combinations of these factor levels are studied simultaneously. In order to analyze the design, each factor should be linearly independent, which means the covariance of the factors should be equal to zero. The covariance is a measure of linear relationship between two random variables [5], and can be calculated by using the following equation where E(x) stands for the expected value of x.

$$Cov(x,y) = E(x,y) - E(x)E(y)$$

To calculate the covariance of the design, a transformation is needed from the lower and upper levels to (-1) and (+1), respectively. After these substitutions, because E(x,y) = 0, E(x) = 0, and E(y) = 0, Cov(x,y) is equal to zero. In orthogonal designs, the covariance is always equal to zero.

The $2^k$ full factorial design is generated by using Yates algorithm. According to this algorithm; for the first factor, a column of (-1) and (+1) is written down with the signs alternating each time. For the second factor, the signs alternate in pairs, for the third factor they alternate in triple, and so on. To create the interactions columns, the levels of the each factor forming the interactions are simply multiplied.

In an experimental design, the number of experiments (rows) must at least be equal to the total degrees of freedom (DF) required for the study, as shown in Table 2.

**Table 2:** DF for $2^k$ full factorial design with k=3

| Factors/Interactions | DF |
|---|---|
| Overall Mean | 1 |
| A,B,C | 3 (2-1) |
| AB, AC, BC | 3(2-1)(2-1) |
| ABC | 1(2-1)(2-1)(2-1) |
| Total | 8 |

One drawback of $2^k$ full factorial design is rapid increase of the number of experiments while increasing the number of the factors ($2^5$=32, $2^8$=256, $2^{10}$=1024). In 1940's, Fisher showed that meaningful results can be obtained by conducting a selected fraction of full factorial design which is called fractional factorial design, $2^{k-p}$, where p stands for the fraction portion.

Since there are 3 factors (k=3) in our problem, $2^3$=8 experiments are needed to run for 2-level full factorial design in Table 3.

**Table 3:** The $2^k$ full factorial design with k = 3

| | A | B | C | AB | AC | BC | ABC |
|---|---|---|---|---|---|---|---|
| 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 |

In each experiment, the factors, or parameters, are set and run according to the design. After the solutions Y obtained from the experiments are analyzed by implementing regression analysis, the mathematical model is derived. However, because $R^2$ value of the model is 1.00, the term that has minimum effect (AB) is removed, and after running the regression analysis again, the following model with $R^2$ = 0.96 is obtained.

Y = 954.37 + 1.63A + 4.88B + 9.13C - 9.13AC + 15.63BC - 15.63ABC

When this model is solved by employing Excel Solver to minimize Y, the parameter setting is found by using $2^k$ full factorial design as

"Crossover = 0.01, Mutation = 0.2, Population = 30".

## 2.2. Orthogonal Array Design

The fact that effects of 3 or higher interactions tend to be insignificant, and therefore may be ignored, bring us to the fractional factorial design type named orthogonal array (OA) design where only main factors and 2-factor interactions are considered. A typical OA tabulation is in the form of $L_a(b^c)$, where a is the number of experiments, b is the number of levels, and c is the number of columns. Taguchi has formulated 18 standard OA designs [7], however they can also be modified by using various methods. To select the appropriate OA, first, number of factors and levels for each factor, and 2-factor interactions to be estimated must be defined. After calculating the DF, the OA with the closest number of the experiments to DF is selected. Interaction tables, or linear graphs developed by Taguchi are then utilized to follow the confounding pattern.

The DF of our problem for OA is 7 due to the absence of 3-factor interactions. The most appropriate OA for 3 factors, 2 levels and 7 experiments is $L_8(2^7)$ which is created in Table 4.

**Table 4:** The OA design with k = 3

|   | C | B | BC | A | AC | AB |
|---|---|---|----|---|----|----|
| 1 | -1 | -1 | 1 | -1 | 1 | 1 |
| 2 | -1 | -1 | 1 | 1 | -1 | -1 |
| 3 | -1 | 1 | -1 | -1 | 1 | -1 |
| 4 | -1 | 1 | -1 | 1 | -1 | 1 |
| 5 | 1 | -1 | -1 | -1 | -1 | 1 |
| 6 | 1 | -1 | -1 | 1 | 1 | -1 |
| 7 | 1 | 1 | 1 | -1 | -1 | -1 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 |

Because there are only 3 factors in the problem, all 2-factor interactions are included. As you notice, the $2^k$ full factorial and OA designs with k=3 are about the same. The reason is that the number of factors is quite small, and increasing this number will clearly bring out the advantages of OA designs in terms of the number of experiments needed to study.

After implementing regression analysis for the OA design, the same mathematical model with $2^k$ full factorial design is derived, except for the ABC term. This model has $R^2$ value of 0.65. As in $2^k$ full factorial design, Excel Solver gives the same solution set for A, B, and C, respectively, namely, the parameter setting for the OA design is again "Crossover = 0.01, Mutation = 0.2, Population = 30".

## 2.3. Central Composite Design

In $2^k$ full factorial and OA designs it is assumed that the relationship between the 2-level factors is linear. It is possible to increase the number of levels to 3 to capture the nonlinearity, however, it would be a bit controversial and none of the rules for the 2-levels would apply in those designs. Also, this would not be the best candidate for continuous factors like parameters used in heuristics. A better approach to cope with the nonlinearity and continuous factors could be Response Surface Method using the Central Composite Design (CCD) developed by Box & Wilson in 1950's [4].

CCD is a first-order design augmented by additional points that allow the estimation of the second-order mathematical model. CCD consists of a full factorial or fractional factorial design ($2^k$ or $2^{k-p}$), a center point (a row of zero's), and two points on axes for each factor at a distance α from the design center which result $2^k+2k+1$ or $2^{k-p}+2k+1$ experiments in total. The distance α is calculated as (number of experiments in fractional portion)$^{1/4}$. It is possible to choose α = ∓1, which is then called face-centered design.

In our problem, $2^3$ = 8 experiments for the fractional portion, 2(3) = 6 experiments for axial portion, and 1 experiment for center portion, total 15 experiments are needed. The distance α is equal to $(8)^{1/4} \cong 1.4$. To be able to set the parameters for each experiment, the levels of the parameters must be coded for the values (-1.4, -1, 0, 1, 1.4). The complete CCD with k = 3 is shown in Table 5.

**Table 5:** Central Composite Design with k = 3

|    | A | B | C | AB | AC | BC | $A^2$ | $B^2$ | $C^2$ |
|----|---|---|---|----|----|----|----|----|----|
| 1  | -1 | -1 | -1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2  | 1 | -1 | -1 | -1 | -1 | 1 | 1 | 1 | 1 |
| 3  | -1 | 1 | -1 | -1 | 1 | -1 | 1 | 1 | 1 |
| 4  | 1 | 1 | -1 | 1 | -1 | -1 | 1 | 1 | 1 |
| 5  | -1 | -1 | 1 | 1 | -1 | -1 | 1 | 1 | 1 |
| 6  | 1 | -1 | 1 | -1 | 1 | -1 | 1 | 1 | 1 |
| 7  | -1 | 1 | 1 | -1 | -1 | 1 | 1 | 1 | 1 |
| 8  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 9  | -1.4 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| 10 | 1.4 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 |
| 11 | 0 | -1.4 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| 12 | 0 | 1.4 | 0 | 0 | 0 | 0 | 0 | 2 | 0 |
| 13 | 0 | 0 | -1.4 | 0 | 0 | 0 | 0 | 0 | 2 |
| 14 | 0 | 0 | 1.4 | 0 | 0 | 0 | 0 | 0 | 2 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

After implementing regression analysis for outcomes of the experiments, the following mathematical model with $R^2 = 0.90$ is derived:

$$Y = 939.42 + 3.58A + 0.75B - 5.58C - 1.13AB + 5.38AC - 11.88BC + 5.94A^2 + 5.94B^2 - 11.31C^2$$

The solution set produced by Excel Solver is back coded to their real values, and the parameter setting found by CCD is "Crossover = 0.218, Mutation = 0.193, Population = 100".

## 2.4. D-Optimal Design

CCD is quite an efficient design especially due to adding the second-order nonlinearity; however, in some cases it may not be enough to understand the relationships between factors. And also, the number of experiences must be kept to an absolute minimum. If a design has an absolute minimum number of experiments, such design is called "saturated design". The minimum number of experiments can be calculated as $(n+1)(n+2)/2$ where n is number of factors. Besides these advantages, if some experiments are infeasible, saturated designs can be still used by extracting these experiments from the design.

As some of the interesting features of saturated designs, unlike the previous DOE methods, they are not orthogonal and there are no degrees of freedom to test the accuracy of the model.

Saturated designs are constructed by applied D-optimality criterion. The following equation is the estimator of simple linear regression:

$$\hat{Y} = b_0 + \sum b_i x_i$$

where $b_0$ is the intercept, $b_i$ are the slopes. If this equation is written in matrix form, we have:

$$Y = XB + \varepsilon .$$

The set of design B can be estimated in the following form by applying the Least Square Regression method.

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

A statistical measure of accuracy of B is the variance-covariance matrix:

$$V(B) = \sigma^2 (X^T X)^{-1}$$

where $\sigma^2$ is the variance of the error. V(B) is a function of $(X^T X)^{-1}$ and to increase the accuracy, $(X^T X)^{-1}$ should be minimized. Statistically, minimizing $(X^T X)^{-1}$ is equal to maximizing the determinant of $(X^T X)$. "D" in the term of D-optimal comes from the first letter of the word "determinant". There are some heuristics [2], and software [3] to come up with a design that maximizes the determinant of $(X^T X)$. To obtain more accurate results, D-optimal designs can be augmented by adding more experiments.

The absolute minimum of experiments for our problem is 10 [=(3+1)(3+2)/2], and the D-optimal design displayed in Table 6 is created by augmenting the design by 2 experiments.

Like CCD, the levels of the parameters must be coded for the values (-1, 0, 1). With the help of regression analysis, the following mathematical model is acquired:

$Y = 92.48 - 0.63A + 2.62B + 8.37C - 6.38AB - 9.13AC + 15.63BC + 19.86A^2 + 23.11B^2 - 18.83C^2$

After the solution set given by Excel Solver is back coded to their real values, and the parameter setting found by D-Optimal is "Crossover = 0.420, Mutation = 0.148, Population = 30".

**Table 6:** D-Optimal Design with k = 3

|    | A  | B  | C  | AB | AC | BC | $A^2$ | $B^2$ | $C^2$ |
|----|----|----|----|----|----|----|----|----|----|
| 1  | -1 | -1 | -1 | 1  | 1  | 1  | 1  | 1  | 1  |
| 2  | -1 | -1 | 1  | 1  | -1 | -1 | 1  | 1  | 1  |
| 3  | -1 | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  |
| 4  | -1 | 1  | -1 | -1 | 1  | -1 | 1  | 1  | 1  |
| 5  | -1 | 1  | 1  | -1 | -1 | 1  | 1  | 1  | 1  |
| 6  | 0  | -1 | 0  | 0  | 0  | 0  | 0  | 1  | 0  |
| 7  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 1  |
| 8  | 1  | -1 | -1 | -1 | -1 | 1  | 1  | 1  | 1  |
| 9  | 1  | -1 | 1  | -1 | 1  | -1 | 1  | 1  | 1  |
| 10 | 1  | 1  | -1 | 1  | -1 | -1 | 1  | 1  | 1  |
| 11 | 1  | 1  | 0  | 1  | 0  | 0  | 1  | 1  | 0  |
| 12 | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  | 1  |

## 2.5. Signal-To-Noise (S/N) Ratio

DOE methods until this section are only based on one instance of our problem, and do not consider any information of other instances. The method of signal-to-noise (S/N) ratio can be defined as a performance measure that takes the mean and the variability into account, and give the ability to use information of other instances in seeking the best parameter setting. It involves two types of factors: control factors and noise factors. Noise factors cause variability which leads to loss of quality. There are three kinds of noise; outer noise, inner noise, and between product noise, or here can be defined as "between instance noise" is the main reason in applying S/N ratio method in our problem.

Generally, data analysis using S/N ratio ($\eta$) can be performed to achieve three types of purposes: smaller-the-better, larger-the-better and nominal-the-best. Since our target is to minimize the total weighted tardiness for the single machine, the appropriate type of $\eta$ is smaller-the-better. To minimize the sensitivity to noise factors, we maximize $\eta$ which is calculated by the following equation [4].

$$\eta = -10\log_{10} (\overline{Y}^2 + \sigma^2)$$

In addition to first instance, fourth and ninth instance are randomly selected as different "products". Unlike in other methods, instead of OA, D-optimal design in Table 6 is used in creating the experiments for each instance because of its advantages, and $\eta$ is calculated as the outcome for each experiment. Three replications of D-optimal design for three instances increase the total number of experiments by 36 (=3x12).

76

After applying the steps of D-optimal design for each instance, the regression analysis is run for to obtain the following mathematical model:

$$Y = 2.36 - 0.83A + 1.17B - 1.62C - 0.36AB + 0.52AC - 4.49BC - 9.27A^2 - 7.30B^2 + 6.07C^2$$

After back coding the findings in Excel Solver to their real values, the parameter setting found by D-Optimal are "Crossover = 0.465, Mutation = 0.157, Population = 30".

## 3. COMPARISON OF DOE RESULTS

After applying five DOE methods to find the best parameter setting for the single machine total weighted tardiness problem, the findings are summarized in Table 7. To test which method is most effective with this problem, these parameter settings are used in solving the first 20 instances for both 40-job problems in Table 8 and 50-job problems in Table 9 respectively [8].

**Table 7:** Parameter settings of DOE methods

| DOE Type | Crossover Prob.(A) | Mutation Prob. (B) | Population Size (C) |
|---|---|---|---|
| $2^k$ FF | 0.010 | 0.200 | 30 |
| OA | 0.010 | 0.200 | 30 |
| CCD | 0.218 | 0.193 | 100 |
| D-Opt. | 0.420 | 0.148 | 30 |
| S/N | 0.465 | 0.157 | 30 |

To be able to compare the solutions for different instances, the percent deviation of the solution from the optimal solution/best known solution is used instead of the real outcomes of the experiments.

Because $2^k$ full factorial and orthogonal array designs give same parameter settings for 3 factors, their common results share the first three columns.

**Table 8:** Comparison of Parameter settings for 40-job problem

| Inst. | Orthogonal Array & $2^k$ Full Factorial Designs | | | Central Composite Design | | | D-Optimal Design | | | S/N Ratios Design | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | %Dev. | Iteration | Time | %Dev. | Iteration | Time | %Dev. | Iteration | Time | %Dev. | Iteration | Time |
| 1 | 4.71 | 4918 | 00:01:22 | 1.86 | 18863 | 00:04:44 | 1.86 | 1895 | 00:00:16 | 1.86 | 31781 | 00:17:00 |
| 2 | 4.65 | 23416 | 00:03:55 | 0.08 | 22391 | 00:05:09 | 0.08 | 23344 | 00:02:19 | 4.65 | 3789 | 00:01:16 |
| 3 | 6.70 | 726 | 00:00:09 | 6.70 | 2850 | 00:00:38 | 6.70 | 2260 | 00:00:18 | 0 | 2826 | 00:00:38 |
| 4 | 1.29 | 5016 | 00:01:02 | 0 | 6853 | 00:01:11 | 0 | 4966 | 00:00:37 | 1.29 | 2607 | 00:00:33 |
| 5 | 0 | 5213 | 00:00:51 | 0 | 4780 | 00:00:51 | 0 | 15052 | 00:01:41 | 0 | 961 | 00:00:13 |
| 6 | 0 | 35635 | 00:04:50 | 0 | 34295 | 00:04:23 | 0 | 6226 | 00:00:50 | 0 | 14237 | 00:02:18 |
| 7 | 3.91 | 10038 | 00:01:39 | 0 | 85445 | 00:13:50 | 0 | 15282 | 00:01:37 | 3.91 | 2536 | 00:00:34 |
| 8 | 0 | 12210 | 00:01:48 | 0 | 26492 | 00:03:32 | 0 | 8461 | 00:01:06 | 0 | 4203 | 00:01:03 |
| 9 | 0 | 39811 | 00:05:15 | 0.59 | 46315 | 00:06:28 | 0.65 | 4881 | 00:00:40 | 0 | 35208 | 00:04:39 |
| 10 | 1.40 | 24753 | 00:03:25 | 0.04 | 88492 | 00:12:38 | 0.04 | 51283 | 00:08:53 | 0 | 51208 | 00:06:26 |
| 11 | 1.94 | 72938 | 00:11:37 | 0.01 | 96244 | 00:18:17 | 0 | 23267 | 00:03:58 | 0 | 23501 | 00:03:51 |
| 12 | 1.40 | 93314 | 00:14:41 | 0.68 | 96335 | 00:21:41 | 0 | 26996 | 00:05:14 | 0 | 29814 | 00:05:48 |
| 13 | 1.00 | 99337 | 00:15:36 | 0.74 | 98275 | 00:29:59 | 0.64 | 25285 | 00:03:31 | 0.64 | 29204 | 00:05:11 |
| 14 | 0.35 | 94223 | 00:19:43 | 0.77 | 88256 | 00:17:54 | 0.33 | 24304 | 00:03:11 | 0.33 | 29969 | 00:07:08 |
| 15 | 0.95 | 94728 | 00:34:23 | 0.91 | 89698 | 00:16:42 | 0.84 | 36549 | 00:12:00 | 0.09 | 76972 | 00:14:47 |
| 16 | 2.96 | 85614 | 00:22:42 | 3.55 | 94294 | 00:11:51 | 0.69 | 94079 | 00:30:19 | 0.82 | 77826 | 00:14:46 |
| 17 | 1.96 | 82399 | 00:19:53 | 2.94 | 96468 | 00:30:05 | 0.49 | 82127 | 00:29:51 | 0.37 | 94687 | 00:17:20 |
| 18 | 2.53 | 99367 | 00:28:07 | 2.61 | 95935 | 00:22:10 | 0.85 | 89000 | 00:26:04 | 0.66 | 93455 | 00:15:24 |
| 19 | 2.26 | 93600 | 00:11:32 | 3.25 | 50709 | 00:10:12 | 0.94 | 99959 | 00:31:45 | 0.82 | 97005 | 00:17:52 |
| 20 | 3.29 | 94366 | 00:10:28 | 3.91 | 79570 | 00:40:05 | 1.04 | 92765 | 00:17:28 | 1.54 | 90153 | 00:32:35 |
| $\bar{Y}$ | 2.07 | 53581 | 00:10:39 | 1.43 | 61128 | 00:13:37 | 0.76 | 36399 | 00:09:05 | 0.85 | 39597 | 00:08:28 |
| $\sigma$ | 1.85 | 39953 | 00:10:04 | 1.84 | 36776 | 00:11:03 | 1.49 | 34984 | 00:11:22 | 1.31 | 35674 | 00:08:37 |
| 0% | 4 | | | 5 | | | 7 | | | 8 | | |

According to data from the 40-job and 50-job problems, the S/N ratios and D-optimal designs seem to be the best two methods of the five DOE methods. While S/N ratios design could reach optimum solutions/best known in 8 instances for 40-job and 6 instances 50-job problems, D-optimal design could obtain them in 7 instances for the 40-job problems, and 6 instances for the 50-job problems. In terms of average percentage deviation, the number of iteration and running time, they are also better than the other three methods. We might accept that S/N ratios design is slightly better than D-optimal design, but it needs three times more experiments than D-optimal design. Even though all DOE methods are completed based on the first instance of 40-job problem, the parameter settings found in these processes produce very close results to the 50-job problems which gives an idea about the robustness of the parameter settings.

**Table 9:** Comparison of Parameter settings for 50-job problem

| Inst. | Orthogonal Array & 2^k Full Factorial Designs | | | Central Composite Design | | | D-Optimal Design | | | S/N Ratios Design | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | %Dev | Iteration | Time | %Dev | Iteration | Time | %Dev | Iteration | Time | %Dev | Iteration | Time |
| 1 | 0 | 2198 | 00:00:43 | 0 | 4320 | 00:00:34 | 0 | 1095 | 00:00:20 | 0 | 3787 | 00:00:35 |
| 2 | 0.10 | 3836 | 00:01:00 | 0.10 | 4320 | 00:00:34 | 0.75 | 1079 | 00:00:14 | 0.75 | 2963 | 00:00:31 |
| 3 | 0 | 55007 | 00:12:09 | 1.39 | 2517 | 00:00:29 | 1.39 | 469 | 00:00:07 | 0 | 3500 | 00:00:34 |
| 4 | 0 | 2977 | 00:00:59 | 0 | 4306 | 00:00:34 | 0 | 1339 | 00:00:19 | 0 | 2802 | 00:00:30 |
| 5 | 5.67 | 7730 | 00:02:51 | 5.67 | 7645 | 00:00:57 | 5.67 | 5460 | 00:01:02 | 5.67 | 2313 | 00:00:27 |
| 6 | 0.57 | 83204 | 00:37:36 | 0.63 | 98371 | 00:09:48 | 0.09 | 8665 | 00:01:39 | 1.9 | 15136 | 00:01:52 |
| 7 | 0.53 | 86284 | 00:19:19 | 0.10 | 98266 | 00:10:33 | 0 | 16295 | 00:03:46 | 0 | 23888 | 00:03:10 |
| 8 | 1.59 | 45109 | 00:12:23 | 1.52 | 89622 | 00:10:42 | 0 | 10580 | 00:02:14 | 0.53 | 18688 | 00:02:14 |
| 9 | 0.34 | 15032 | 00:03:37 | 0.39 | 25102 | 00:02:39 | 0.39 | 8665 | 00:00:25 | 0.39 | 6699 | 00:01:53 |
| 10 | 0 | 69766 | 00:11:03 | 0 | 84602 | 00:12:46 | 0 | 7054 | 00:01:17 | 0 | 19074 | 00:02:38 |
| 11 | 3.37 | 92353 | 00:14:54 | 4.92 | 87583 | 00:13:10 | 0 | 26931 | 00:03:34 | 0.51 | 93839 | 00:13:46 |
| 12 | 4.15 | 85150 | 00:14:19 | 5.74 | 98291 | 00:14:20 | 0.93 | 17253 | 00:02:25 | 0.79 | 93465 | 00:12:35 |
| 13 | 3.49 | 89551 | 00:24:01 | 5.00 | 78979 | 00:11:31 | 0.26 | 23561 | 00:03:44 | 0.16 | 97849 | 00:11:58 |
| 14 | 1.12 | 86613 | 00:21:19 | 2.06 | 92867 | 00:14:09 | 0.52 | 14498 | 00:02:18 | 0.62 | 65048 | 00:10:35 |
| 15 | 1.16 | 96432 | 00:11:46 | 2.59 | 93136 | 00:13:10 | 0.92 | 27107 | 00:04:37 | 0 | 82041 | 00:12:49 |
| 16 | 5.50 | 95136 | 00:16:26 | 7.11 | 71351 | 00:15:37 | 0.02 | 85596 | 00:08:51 | 1.41 | 90719 | 00:14:29 |
| 17 | 6.67 | 97738 | 00:09:44 | 7.61 | 91351 | 00:13:28 | 0.15 | 87644 | 00:09:17 | 1.2 | 99878 | 00:15:23 |
| 18 | 5.58 | 96171 | 00:09:25 | 7.97 | 65694 | 00:10:25 | 0.53 | 97533 | 00:09:33 | 2.16 | 99272 | 00:11:57 |
| 19 | 3.51 | 95654 | 00:10:56 | 5.38 | 80849 | 00:12:35 | 0.19 | 86540 | 00:08:52 | 0.92 | 93995 | 00:18:40 |
| 20 | 5.31 | 78924 | 00:08:26 | 6.60 | 91293 | 00:16:13 | 0.48 | 88186 | 00:08:31 | 1.92 | 99114 | 00:12:08 |
| $\bar{Y}$ | 2.43 | 64243 | 00:12:09 | 3.24 | 63523 | 00:09:13 | 0.61 | 30777 | 00:03:39 | 0.95 | 50704 | 00:07:26 |
| $\sigma$ | 2.37 | 36918 | 00:08:57 | 2.94 | 38429 | 00:05:47 | 1.25 | 35523 | 00:03:26 | 1.31 | 42911 | 00:06:24 |
| 0% | 4 | | | 3 | | | 6 | | | 6 | | |

## 4. CONCLUSIONS

DOE offers a practical way to tune the heuristic parameters. Because the number of parameters, or factors, is not the same for all heuristics, it is important to select the right DOE method. Table 10 shows how fast the number of experiments increases for a small amount of increase in the number of factors with three levels. Other important issues include selecting the number of levels, values of the levels, the type of relationships between factors, and the cost of running of an experiment.

**Table 10:** Number of experiments for 3-levels

| Factor (k) | 3^k FF | OA | CCD | D-Opt. |
|---|---|---|---|---|
| 4 | 81 | 27 | 25 | 15 |
| 5 | 243 | 81 | 43 | 21 |
| 7 | 2187 | - | 143 | 36 |

It should be noted that the same parameter setting produces different solutions for different instances although all instances are created from the same distributions. For the total weighted tardiness problem, the most effective methods turned out to be the D-Optimal and S/N Ratios Design, with the D-Optimal design requiring less runs.

This paper presented a structured framework on using DOE to tune optimization algorithm parameters. The weighted tardiness scheduling problem was used as a vehicle to demonstrate the approach. The same approach can be applied to other problems.

## REFERENCES

1. Birattari, M. (2009) "Tuning Metaheuristics: A Machine Learning Perspective", *Studies in Computational Intelligence*, Volume 197, Springer-Verlag Berlin Heidelberg.

2. Box, J.M., Drapper, N.R. (1974) "On minimum point second order designs", Technometrics, 16, 613-616.

3. JMP, Design of Experiments Guide (2007), SAS Institute Inc, Cary, North Carolina.

4. Montgomery, D.C. (2001) "Design and Analysis of Experiments", Fifth Edition, John Wiley & Sons, Inc., USA.

5. Montgomery, D.C., Runger G.C. (2003) "Applied Statistics Applied Statistics for Engineers", Third Edition, John Wiley & Sons, Inc., USA.

6. Palisade Corporation (2008), "Evolver, The Genetic Algorithm Solver for Microsoft Excel", Version 5.0, Ithaca, NY USA.

7. Phadke, M.S. (1989) "Quality Engineering Using Robust Design", AT&T Bell Laboratories, PTR Prentice-Hall, Inc., New Jersey.

8. http://people.brunel.ac.uk/~mastjjb/jeb/orlib/wtinfo.html.