AN ABSTRACT OF THE PROJECT OF

Student Yue Liu for the degree of Master of Science in Mechanical Engineering presented on March 21st 2016.

Title: Design of Experiments to Support Automated Assembly Planning.

Abstract approved:

____

Christopher J. Hoyle

Assembly time estimation is a key factor in evaluating the performance of the assembly process in an automated assembly process. The overall goal of this study is to develop an efficient assembly time estimation method by generating the prediction model from an experimental design. In order to estimate these times, this paper proposes a way to divide the assembly into four actions which consist of a) part movement, b) part installation, c) secure operations, and d) subassembly rotations. The focus of this paper is a model for the secure operation; however, the methodology can be applied to the other three times of interest. To model secure times, a design of experiments is applied to collect experimental data based on the physical assembly experiments performed on products that are representative of common assembly processes. The Box-Behnken design (BBD) is an experiment design to support response surface methodology to interpret and estimate a prediction model for the securing operations. The goal is to use a quadratic model, which contains squared terms and variable interactions, to study the effects of different engineering parameters of securing time. The experiment is focused on individual-operator assembly operations. Various participants perform the experiment on representative product types, including a chainsaw, a lawn mower engine, and an airplane seat. In order to optimize the assembly time with different influence factors, mathematical model were estimated by applying the stepwise regression method in MATLAB. The second-order equations representing the securing time are expressed as functions with six input parameters. The models are trained using all combination data required by the BBD method and predict the hold back data within a 95% confidence interval. Overall, the results indicate that the predicted value found was in good agreement with experimental data, with an Adjusted R-Squared value of 0.769 for estimated securing time. This study also shows that the BBD could be efficiently applied for the assembly time modeling, and provides an economical way to build an assembly time model with a minimum numbers of experiments.

Design of Experiments to Support Automated Assembly Planning

By
Yue Liu

A PROJECT

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

March 21st 2016
Commencement June 2016

Master of Science Project of Yue Liu presented on March 21st 2016

APPROVED:

Major Professor, representing Mechanical Engineering

Head of the College of Mechanical, Industrial and Manufacturing Engineering

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Student Name, Author

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my project advisor Dr. Christopher J. Hoyle and Dr. Matthew I. Campbell. This project could not have been finished without your support and guidance, and I really appreciate all of the help that you have been providing to my research and study in the past few months. I also would like to thank Dr. Demirel Onan for willing to be my committee member and review my thesis. I am also thankful to Weifeng Huang and Nima Rafibakhsh for participating in my experiment and for cheering me up along the way.

Finally, I cannot go this far without the support and love from my family and the love of my life Sihan Li. Thank you so much.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

LIST OF TABLES

1 Introduction

Assembly time is a key point in evaluating the performance of an assembly process in the early stage of manufacturing. In order to support the automated assembly planning, the necessary operation time during the assembly need a more accurate and efficient method to optimize. Therefore, a method of time predicting in assembly need to be developed efficiently.

For the assembly time, the prediction model could be a mathematical model that contains various factors with individual weighted coefficient, it describes the critical effect of each factor on the response in the model. In this report, a design of experiment (DOE) method is used to collect experimental data and estimate the relationships between factors and response in the mathematical model. Three assembly experiments were designed for obtaining quality data of all input factors and output response during the assembly.

In order to estimate these times, this paper divides the assembly into four actions, which consist of a) part movement, b) part installation, c) secure operation, and d) subassembly rotation [1]. The focus of this paper is a model for the secure operation. Securing model reflects the time necessary to secure two or more parts after installation. There are several kinds of method to secure parts in assemblies, such as snap fitted, tight fitted and welding etc. Due to the particularity and variety of securing methods, the combination among different securing methods would influence the result of securing time. Fastener securing is considered as the only securing method in this experimental design, since it is commonly used in industry for securing.

The assembly experiments were conducted by each participant individually. The experimental objects from various types of products were selected, and sequence for each assembly step were designed and instructed to the participants. In the experiments, a chain saw, a lawn mower engine and an airplane seat were chosen for assembly experiments. The assembly working stations were designed by the authors to simulate manufacturing process. All components were placed in a designed arrangement with necessary tools provided. The participants followed the instructions to install and secure the components, and the data for building securing prediction model was collected during the experiment.

The multiple linear regression model based on stepwise method was used to analyze the relationships between factors and establish a good prediction model. The stepwise regression method is well developed in the research of evaluating the significance of variables and improving the model by removing the insignificant factors [2]. Seven variables were selected as input parameters for model predication, and six of them were identified as significant variables to securing time by the analysis of stepwise regression.

2 Related Work

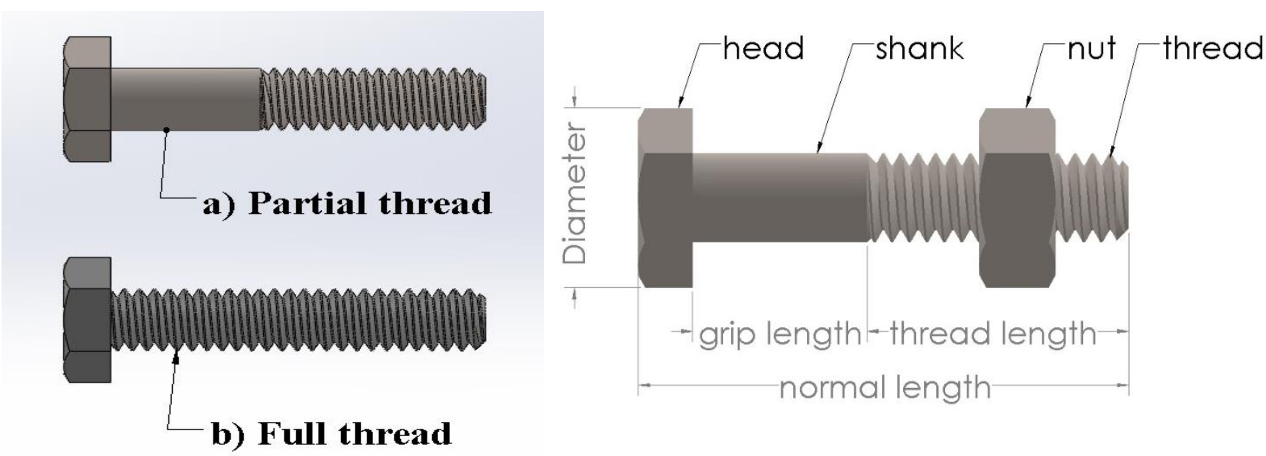
Many research have been done to build assembly time prediction model, and several methods to estimate the assembly time are well established [3]. The Interference Detection Method (IDM) is one of the tools built to predict assembly time and reduce the required input from designers. In this method, the assembly time can be estimated from the part location represented by bi-partite graphs in the assembly space, where the components connectivity information is obtained from a computer-aided design (CAD) model. Different from the IDM method, the Assembly Mate Method (AMM) focus on the assembly mate information based on connectivity graph [3]. Although the two methods estimate the assembly time within 45% of the target time [4], both methods are limited to only CAD assembly files. Most methods using geometric and assembly connectivity are highly related to the CAD model, but some of the securing difficulties haven't been specified by these research. Different types of fasteners are major input parameters in the estimation of assembly time. The securing operation is a key section that needs to be considered in the assembly estimation. For example in the SolidWorks, the components and fasteners are only fitted with standard or mechanical mate, so the securing difficulty could be under-estimated. For some small size fasteners, more time is possibly necessary due to the handling or insertion difficulty. Hence, the simulation based on CAD model may not comprehensively include the securing difficulty in the real-time assembly. Therefore, more input parameters related to the fasteners are required to estimate the securing time.

Meanwhile, the Design of Experiment (DOF) has also been applied to the model prediction. In Arbizu and Perez's study [5], a factorial design using the response surface methodology is used to predict the components' surface quality and dimensional precision. The results show that a factorial design with regression analysis can successfully be applied for modeling, and significantly optimize the time and cost by reducing number of experiments.

Also, stepwise regression method is one of the most successful methods to analyze and improve the prediction. The regression analysis method is compared with decision tree and neural network in Tso and Yau's study [6] in predicting energy consumption. According to their results, the decision tree model and neural network model are slightly better in model prediction with simpler structure and fewer variables in energy field compared to the regression method. More factors and complex interaction effects are considered in this paper, making the regression method with statistical analysis a good estimator for model prediction.

3 Input parameters selection

Assembly time can be affected by many factors. Design for assembly (DOA) indicates several factors that influence the assembly operations, such as part handling, size and geometrical features [7]. Since this paper is only focused on the securing time, and the number of input parameters is limited in data collecting in order to reduce cost and time; thus, the securing related factors are selected as the input parameters. Two types of screws were used in the experiments: a) partial threaded screw, b) full threaded screw, as shown in the Fig.1a. The selected input parameters need to represent the engineering features for these two types of screws and could be physically measurable. Moreover, the parameters also need be read from the CAD model, which requires less input value from users.



(a) (b)

Figure 1: (a) Two types of screws and (b) Detailed features for screw

As a result of the aforementioned reasons, normal length, diameter, thread number and insertion distance were selected as controllable variables and coded in three levels; insertion difficulty, tool effect and nut were chosen as covariates, which were coded as 0 and 1 as shown in the Tab. 1, where 0 represents non-existent and 1 existent for insertion difficulty and nut; and for tool effect, 0 represents manual tool, such as screwdriver and wrench and 1 represents drill driver. The covariates are the factors during the measurements which might affect the response, so the covariates would be useful to be included as design factors. Covariates in this design are defined as uncontrollable variables that influence the response value but do not need to be included in the DOE; therefore, it can be shown as an experiment or measurement error, and only need to be recorded in the experiments.

	Input variables	Coded symbol	Coded level		
Controllable variables	Normal length	x1	-1(low)	0(mid)	1(high)
	Diameter	x2	-1(low)	0(mid)	1(high)
	Thread number	x3	-1(low)	0(mid)	1(high)
	Insertion distance	x4	-1(low)	0(mid)	1(high)
Covariates	Insertion difficulty	x5	0(no)		1(yes)
	Tool effect	x6	0(no)		1(yes)
	Nut	x7	0(no)		1(yes)

Table 1: Input parameters chosen for securing operation model.

3.1 Normal length and Diameter

Oriented bounding box volume (OBB) is a major factor in securing model. It is defined as the smallest box that encloses the fastener, and its volume affects the time in handling. For example, if the fasteners are very small or very large, it would require more handling time during the assembly. A small fastener have the most possibility to be released before it is located [7], so it is reasonable to assume OBB volume would influence the securing time in handling and installing. For most of the fasteners, the

OBB volume can be approximately calculated by the normal length and diameter of a fastener because of its rotational symmetry. Moreover, normal length and diameter may have interaction effects on each other or on other variables; thus, these two were selected to represent OBB volume as variables for securing model.

3.2 Insertion distance

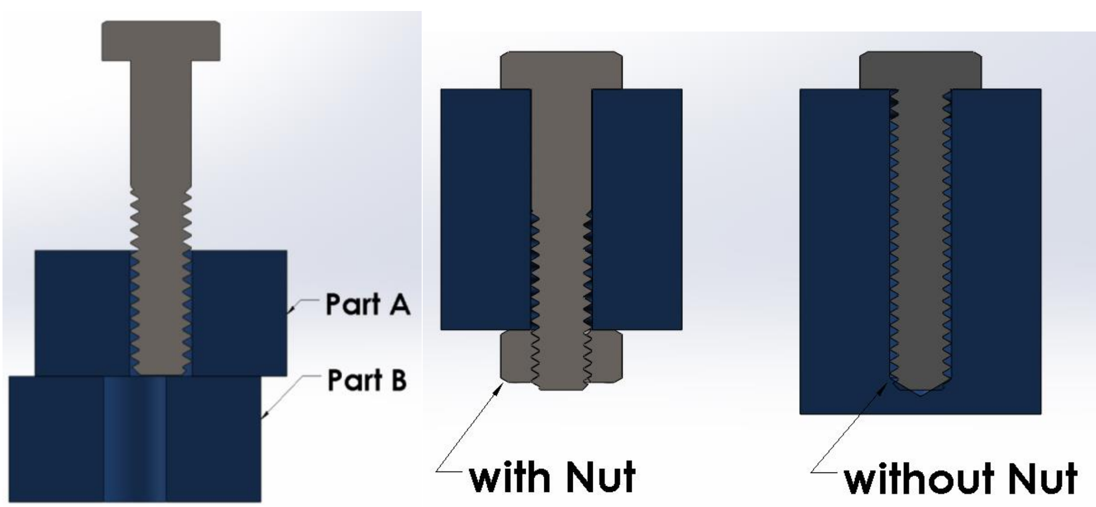
The insertion distance is used to estimate the distance for fastener insertion by approximating the depth that a fastener needs to be thread-inserted into the parts. Usually, this parameter can be represent by the thread length. However, the thread length cannot accurately describe the actual insertion distance for fastening in some situation. Meanwhile, the actual insertion distance inside the part is very difficult to collect from experiments. In order to save time and cost, approximately half of the thread length is estimated as the insertion distance for the situations where thread length could not represent insertion distance.

3.3 Thread number

Thread number is another input parameter for securing model. It is defined as the how many turns are needed for fastening. There are two reasons that thread number was selected. First, if two types of fasteners have the same normal length and assuming fasten speed is constant, then the one with more thread number obviously need relatively more time than the one with less thread number. Second, it is a measurable feature related to the fasten speed, which is very hard to measure and control in the experiment. Again, approximately half of the thread number is used as input parameter when a nut is involved in the fastening or the actual thread number is difficult to measure.

3.4 Insertion difficulty

Insertion difficulty is used to describe the extra time for inserting fasteners before fastening, which happens quite often during the experiment. As shown in Fig. 2a, the screws may be blocked by the bad alignment from two components. The value 0 and 1 are assigned to represent the insertion difficulty. If there is no insertion difficulty, 0 is assigned, which means there is no insertion effect on the experiment. And a value of 1 will be assigned if more time is needed during the fastener insertion due to bad alignment of two connecting parts.



(a) (b)

Figure 2: (a) Insertion difficulty for two connecting parts; (b) Nuts option for two types of fastener.

3.5 Tool effect

Two fixture methods are designed in this research. Drill driver and manual tools, such as screw driver, wrench and torque wrench are used in the assembly process. Fig.3 shows the tools used during the assembly process. This parameter is defined as: 0 represents the manual tools and 1 represents the drill driver.

Figure 3: Manual tools and drill driver.

3.6 Nuts

Two nut options are considered in this research, screws with nuts, coded as 1 and screws without nuts, coded as 0, and the example is shown in Fig.2b. This parameter is included because screws with nuts would require relatively more time and more tools than the alternative.

4. Design of experiments

In recent years, design of experiments has been applied to optimize and predict multiple variable systems in various fields [2][5]. The advantage of DOE is to use a statistical approach to avoid studying all possible combinations, and a minimum number of experiments is required to analyze the relationships between the response and multiple-variables.

4.1 Response surface methodology and Box-Behnken design

Response surface methodology (RSM) is a statistical analysis for model building and predicting. By carefully designing experiments, RSM can identify the relationship between response and predictors or input variables [8]. Therefore, the goal of RSM in this study is to find an approximation for the true relationship between securing time and input parameters. 7 variables for securing model have been determined, but the relationship between the securing time and variables is unknown. The next step for RSM is to use a special experimental design to perform a response surface regression. If a full factorial design is constructed for seven variables, all possible variable interactions must be considered to estimate the fitting error of the model. Usually, it would require a large number of repeated testing which is time consuming and less efficient. The Box-Behnken design (BBD) was used to efficiently optimize the response of seven input parameters. Each factor was divided into three levels and a reasonable number of experimental runs were proposed at the same time for all seven variables. Thus by studying the designed combinations of factors, the coefficients for each factor can be estimated and fitted in a mathematical model that best fits the experimental data. Moreover, this method also involves predicting the response of the fitted model and checking the accuracy of the model by comparing the predicted value to the experimental value [8].

4.2 Determination of levels and number of experiments

The controllable variables, normal length, insertion distance, diameter and thread number were chosen and assigned three range levels. Three coded range levels were defined to describe the different levels for each factor: low (-1), center (0) and high (1) as shown in Tab. 2.

Range levels of variables chosen for the BBD

Variables	Coded Symbols	Uncoded range levels		
		-1	0	1
Normal length (mm)	x1	0~20	20~40	40~60
Diameter(mm)	x2	5~7	7~10	10~15
Thread length (mm)	x3	5~14	14~16	16~18
Thread number	x4	5~7	10~13	13~22

Table 2: Range level of variables chosen for the design

The required combinations of four variables in Box-Behnken design are shown in Tab. 3. In this table, each (±1, ±1) combination within a row represents a full 2^2 factorial design. In each block (each experiment), two factors are in all combinations for the factorial design, while the other factors are kept at the central values [9]. 24 is required as the minimum number of experiments in this BBD, and appropriate replications on experimental run 7 is used to decrease the experimental error. Therefore, 29 experiments (with 5 central replications) were determined for the four factors BBD.

Coded factor range levels for Box-Behnken design				
Experiment run	Four-factor			
	x1	x2	x3	x4
1	±1	±1	0	0
2	0	0	±1	±1
3	±1	0	0	±1
4	0	±1	±1	0
5	±1	0	±1	0
6	0	±1	0	±1
7	0	0	0	0

Table 3: Coded factor levels for Box-Behnken designs for four factors

4.3Mathematical model for prediction

The mathematical relationship between the seven variables and the response can be approximated by the following quadratic polynomial equation:

$$
\begin{aligned}
T_s = {} & w_0 + w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + w_5x_5 + w_6x_6 + w_7x_7 \\
& +w_{11}x_1^2 + w_{22}x_2^2 + w_{33}x_3^2 + w_{44}x_4^2 + w_{55}x_5^2 + w_{66}x_6^2 + w_{77}x_7^2 \\
& +w_{12}x_1x_2 + w_{13}x_1x_3 + w_{14}x_1x_4 + w_{15}x_1x_5 + w_{16}x_1x_6 + w_{17}x_1x_7 \\
& +w_{23}x_2x_3 + +w_{24}x_2x_4 + w_{25}x_2x_5 + w_{26}x_2x_6 + w_{27}x_2x_7 + w_{34}x_3x_4 \\
& +w_{35}x_3x_5 + w_{36}x_3x_6 + w_{37}x_3x_7 + w_{45}x_4x_5 + w_{46}x_4x_6 + w_{47}x_4x_7 \\
& +w_{56}x_5x_6 + w_{57}x_5x_7 + w_{67}x_6x_7
\end{aligned}
$$

where T_s is response variable; $w_0$is model constant; w_i (i = 1,2,3,4,5,6,7) are linear coefficients; w_{ii} are quadratic coefficients; w_{ij} (j = 1,2,3,4,5,6,7) are cross product coefficients. The experimental process and operations are presented in section five.

5. Experiment of product assemblies

The goal of this experimental design is to collect data for model prediction, where the data consist of the input parameters and the time for each step during the assembly. In this experiment, three products were assembled by various participants and the operation processes were recorded. The assembly times were recorded by breaking the video into frames to avoid possible measurement error.

5.1 Product selection for experiment

Various assembly products with significant distinctive features, such as fastener size, type and special features need to be selected carefully to ensure that wide ranges of data can be obtained for estimating a more universal prediction model. In this study, a chain saw, an engine and an airplane seat were selected to increase the variety of the data. For each object, an overall arrangement were created in one white and blank paper for each component individually. All parts were displaced in a random location on the paper without any connection to the next part to which they will be fastened. The authors traced the profile of all components for each object, and illustrated the assembly order by labeling number beside each component, and it is required that each participant follows identical instructions. The securing information, like location and type of fasteners, were provided to the participants before the experimental assembly.

5.1.1 Chain Saw

Chain saw is a relative small and light-weight product primarily consists of twenty two plastic components, and only the chain supporter and the motor are made of metal. This product also contains 16 threaded fasteners which are of 6 different types. This product is chosen because it has more complex alignment features than other products that make securing more difficult. And most of fasteners have less thread number than those from the other two products. Therefore, choosing this product can test whether it is suitable to select insertion difficulty and thread number as input parameters.

5.1.2 Engine

The second product is a lawnmower engine. In order to save time in the experiment, the repeated components such as valves, valve rods and cams were removed to reduce the installation time. After the reduction, this engine has 21 parts, and most of them are made of metal. Nine fasteners from three different types are used for fastening the product. The majority of the parts is small and needs relatively more time to assembly. Meanwhile, the fasteners for this product consist of cap socket screws, which required different tools and all fasteners are full threaded screws with large thread numbers. Thus it provides significant diversity to the experimental data compared to the chain saw.

5.1.3 Airplane Seat

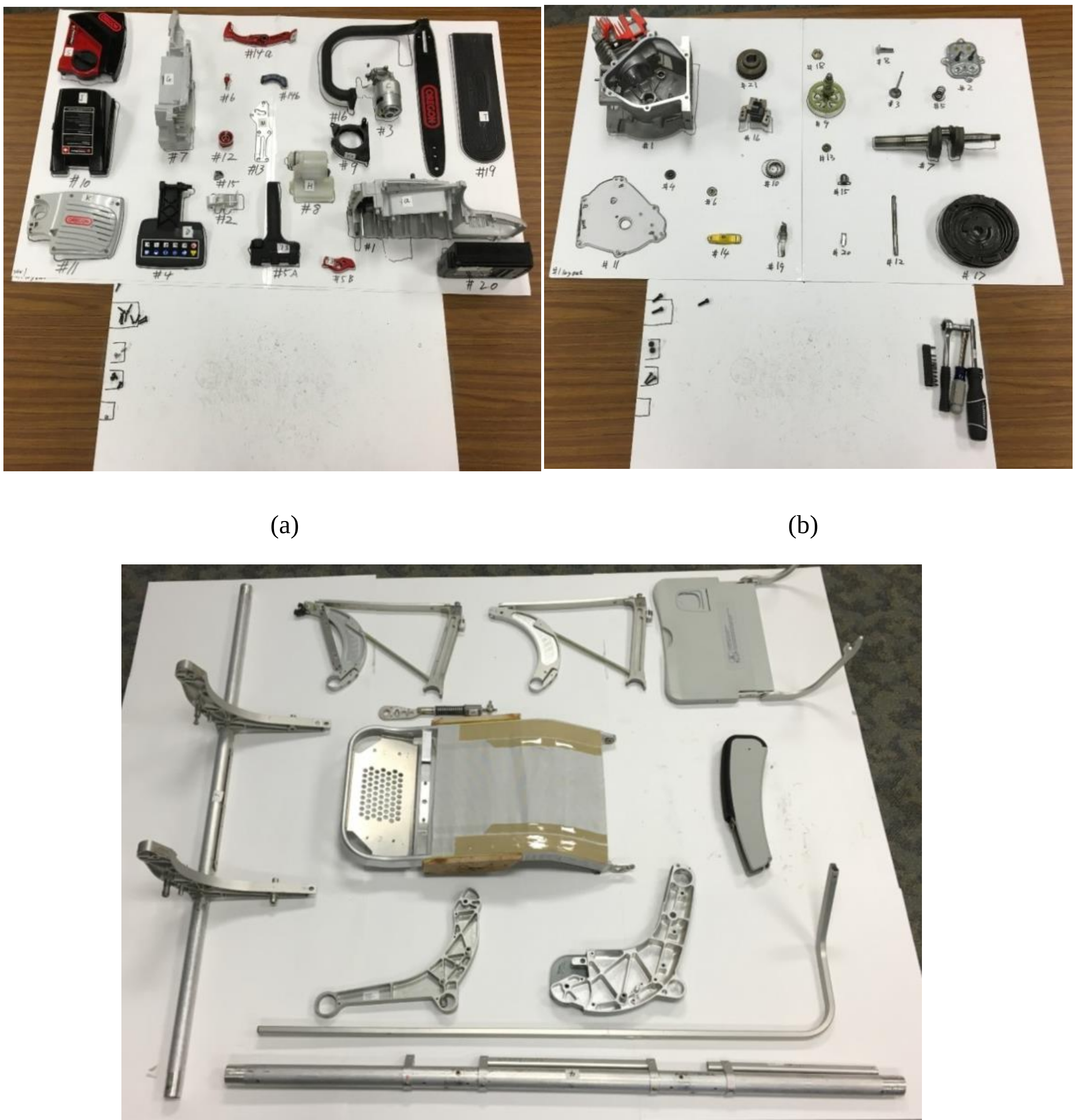
The third assembly is an airplane seat. Overall 32 parts could be disassembled and 48 fasteners are used for the assembly. There are three reasons to select this product. First, it has both the largest and smallest size of fasteners of three products. Moreover, some fasteners have the longest normal length but a very small diameter, so it is a good decision that normal length and diameter were selected to represent OBB volume. Second, compared to the first two objects, the airplane seat has a more complex structure and larger size, which can be used both for developing installing and securing model. Several screws came with nuts in the experiment. Selecting this product can obtain more experimental data related to the different size of fasteners, and also explore the nut effect for fastening time.

5.2 Operating participants

In order to obtain more data from experiments, the assembly operations for each of the three products were conducted by five individuals, and all fasteners were repeated twice by each participant. All participants were properly trained and instructed before the experimental assembly, thus it is assumed all participants have qualified capability to accomplish the assembly operation.

5.3 Experimental procedure

During the assembly, the experiments were designed and simulated as the manufacturing scenario, with the assembly layout shown in Fig. 4. All components were placed in a randomly designed location with a number that indicates the assembly sequence. The goal of this assembly layout is to collect the experimental data for the installing and securing operation at the same time. Each fastener and tool should be taken one per time during the experimental procedure. The majority of the fasteners were fastened by a drill driver, some fasteners were fixed manually by screw driver or wrench. All fixture tools and fasteners were placed within the working area storage, and brought back after the installing or securing.



(a) (b)

(c)

Figure 4: Three assembly layouts for product assembly experiments: (a) chain saw, (b) lawnmower engine (c) airplane seat

3.5 Time recording

The securing time for each step need to be recorded precisely in the experimental design. The experiments basically focus on two major times, inserting and fastening time. In order to eliminate the confounding factors, the authors defined a criterion that inserting time starts as the fastener touch the part, and ends when it is completed inserted in the hole. And then the fastening time is recorded as the fixture time by tools.

To minimize the error during the time recording, a High Definition (HD) camera was used to record every step of experiment for each participant. All videos were replayed on Video Studio X Pro, in which each second was broken down in 30 frames. By applying this method, the authors could precisely read the assembly time by simply calculating how many frames were used during each step of assembly.

6. Results and discussion

The experimental data and analysis of the stepwise regression model for BBD are discussed in this section. The relationships between the securing time and selected variables were shown in the plots. And experimental data was used to construct predictive model for securing time. The analysis and variance (ANOVA) was used to determine whether this predictive model is significant or not [9]. The model was improved and finalized through the stepwise regression analysis.

6.1 Data processing

Each experimental run (each fastener) was done by five participants with two replications. That means, for each type of fastener, ten datasets were collected from five participants. The experimental time were transformed with a natural log logarithm. Using all of data instead of the mean response is to eliminate the experimental error and consider the person to person variance in the further work. Fig. 5 shows the log-transformed securing time with 4 controllable variables. The plots don't show a monotonously increase for each variables, which is possibly due to the fact that the experimental sample is not large enough or there are interaction effects, so the securing time trend cannot be clearly seen from these plots. Considering this situation, a multiple regression method is used to estimate the securing time by the seven variables.

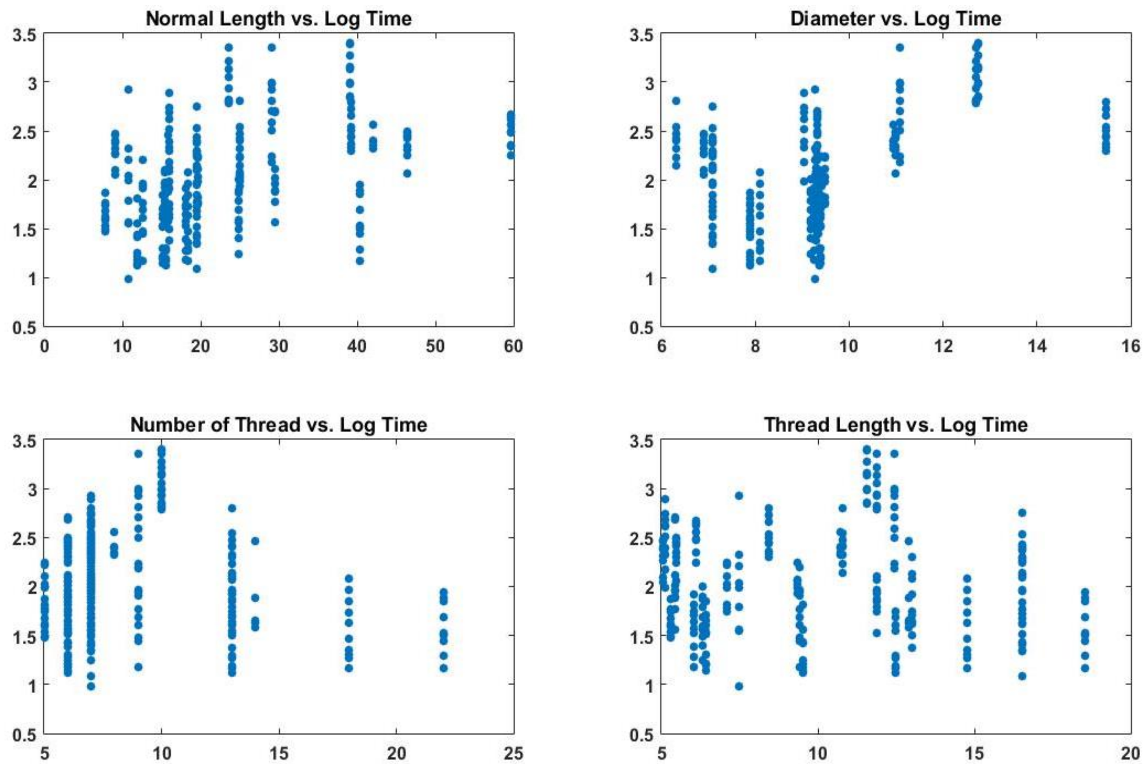


Figure 5: Plots of experimental time with input parameters: (a) Normal length; (b) Diameter; (c) Thread length; (d) Thread number

6.2 ANOVA results and regression model

A quadratic model using standard multiple regression was built and the result shows that this model is highly significant regardless of the significance of factors. The model has a F-value of 37.87 with a low probability value ($P = < 0.05$). The adjusted R-square is 0.775. This value suggests that the model can predict 77.5% of the variability in the response [10]. The p-values were used to identify the significant variables, and they also showed whether interactions between different factors are significant. Given the p-value of the F-test for each factor and interactions shown in Tab. 4, several terms in this model suggest an insignificant effect on the response time.

Including these terms may provide a well-fitted model for the training data, but could result in an imprecise prediction for the new data.

Analysis of Variance Table

Response: Totaltime

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
x1	1	1323.07	1323.07	216.9518	< 2.2e-16
x2	1	840.01	840.01	137.7410	< 2.2e-16
x3	1	68.99	68.99	11.3133	0.0008846
x4	1	113.88	113.88	18.6729	2.205e-05
x5	1	732.74	732.74	120.1525	< 2.2e-16
x6	1	0.03	0.03	0.0053	0.9419659
x7	1	810.24	810.24	132.8601	< 2.2e-16
x12	1	259.85	259.85	42.6099	3.443e-10
x22	1	21.69	21.69	3.5563	0.0604244
x32	1	306.88	306.88	50.3212	1.217e-11
x42	1	31.45	31.45	5.1567	0.0239686
x1:x2	1	8.60	8.60	1.4105	0.2360468
x1:x3	1	289.02	289.02	47.3919	4.285e-11
x1:x4	1	38.58	38.58	6.3270	0.0124905
x1:x5	1	91.44	91.44	14.9947	0.0001362
x1:x6	1	632.83	632.83	103.7687	< 2.2e-16
x1:x7	1	50.89	50.89	8.3444	0.0041921
x2:x3	1	215.24	215.24	35.2946	8.987e-09
x2:x4	1	86.74	86.74	14.2227	0.0002008
x2:x5	1	0.97	0.97	0.1593	0.6900850
x2:x6	1	198.90	198.90	32.6140	3.042e-08
x2:x7	1	0.36	0.36	0.0591	0.8081711
x3:x4	1	30.53	30.53	5.0056	0.0261072
x3:x5	1	0.60	0.60	0.0977	0.7549000
x3:x6	1	3.89	3.89	0.6372	0.4254411
x3:x7	1	2.03	2.03	0.3336	0.5640571
x4:x5	1	76.75	76.75	12.5856	0.0004605
Residuals	262	1597.79	6.10		

Table 4: ANOVA table for the standard regression model

6.3 Stepwise regression method

In view of this full term quadratic model, some variables or terms have an insignificant effect on the predicted response. To improve the model, a robust method is to use stepwise regression to add terms that contribute to the prediction or remove terms that don't. In general, the stepwise regression model is to select the best combination of variables to predict the response, not all variables or interactions are kept in the equation [2]. During the analysis, one predictor is removed or added at each time, where the predictor could be a variable or interaction; the process of selecting significant factors stops when the predicted model cannot be improved in terms of increasing the correlation, adj-R^2 value [10]. All variables and interactions are tested by a F-test to fit the model, and a significant factor will be kept if the P-value given by F-test is less than 0.05.

The first step for stepwise regression is to choose a constant model as the initial equation. In this research, the initial equation is comprised of the first-order and second-order terms of all factors. Then, the significant interaction terms are added and insignificant variables are removed. The finalized prediction model through stepwise regression analysis is shown in the following equation (in term of coded symbols).

$$\begin{aligned}\text{Log(Predicted time)} = {} & 9.170 - 0.049x_1 - 1.662x_2 + 0.053x_3 + 1.543x_5 \\ & + 2.774x_6 - 0.904x_7 + 0.001x_1^2 + 0.101x_2^2 - 0.004x_3^2 \\ & - 0.188x_2x_5 - 0.651x_2x_6 + 0.563x_3x_6 + 0.055x_3x_7 + 0.356x_5x_7\end{aligned}$$

The above equation is based on the quadratic model improved by the stepwise regression method. The p-value of each factor was used to examine the significance level, which also show the interaction effect, as shown in the Tab. 5. The model's F-value of 69.6 implies that this reduced model is significant. The significant variables and interactions are indicated as the value of "Prob>F" less than 0.05 [10]. Thus, $x_1, x_2, x_3, x_5, x_6, x_7, x_1^2, x_2^2, x_3^2, x_2x_5, x_2x_6, x_3x_6, x_3x_7, x_5x_7$ are significant model terms. The adjusted R-square value of this quadratic model is 0.769, which indicates that the fitness of the chosen model of this process is good.

	SumSq	DF	MeanSq	F	pValue
x1	2.4244	1	2.4244	37.869	2.6753e-09
x2	1.4996	1	1.4996	23.423	2.1737e-06
x3	0.48541	1	0.48541	7.582	0.0062897
x5	1.1785	1	1.1785	18.409	2.4742e-05
x6	0.37169	1	0.37169	5.8058	0.016633
x7	4.1214	1	4.1214	64.375	3.025e-14
x2:x5	1.215	1	1.215	18.977	1.8715e-05
x2:x6	8.7321	1	8.7321	136.39	7.6224e-26
x3:x6	3.642	1	3.642	56.888	6.8114e-13
x3:x7	1.0966	1	1.0966	17.129	4.652e-05
x5:x7	1.0323	1	1.0323	16.124	7.6653e-05
x1^2	3.7137	1	3.7137	58.007	4.2558e-13
x2^2	5.917	1	5.917	92.422	4.8253e-19
x3^2	0.82402	1	0.82402	12.871	0.00039514
Error	17.542	274	0.064021		

Table 5: ANOVA for stepwise regression model

6.4 Verification and predication testing

A normal probability plot of residuals is shown in Fig. 6. The plot shows the error terms of the regression model are approximately normally distributed along a least-square line. Thus, it is reasonable to assume that no serious assumptions are violated under the analysis [9]. The histogram of residuals show there is no obvious outlier in the modified model.

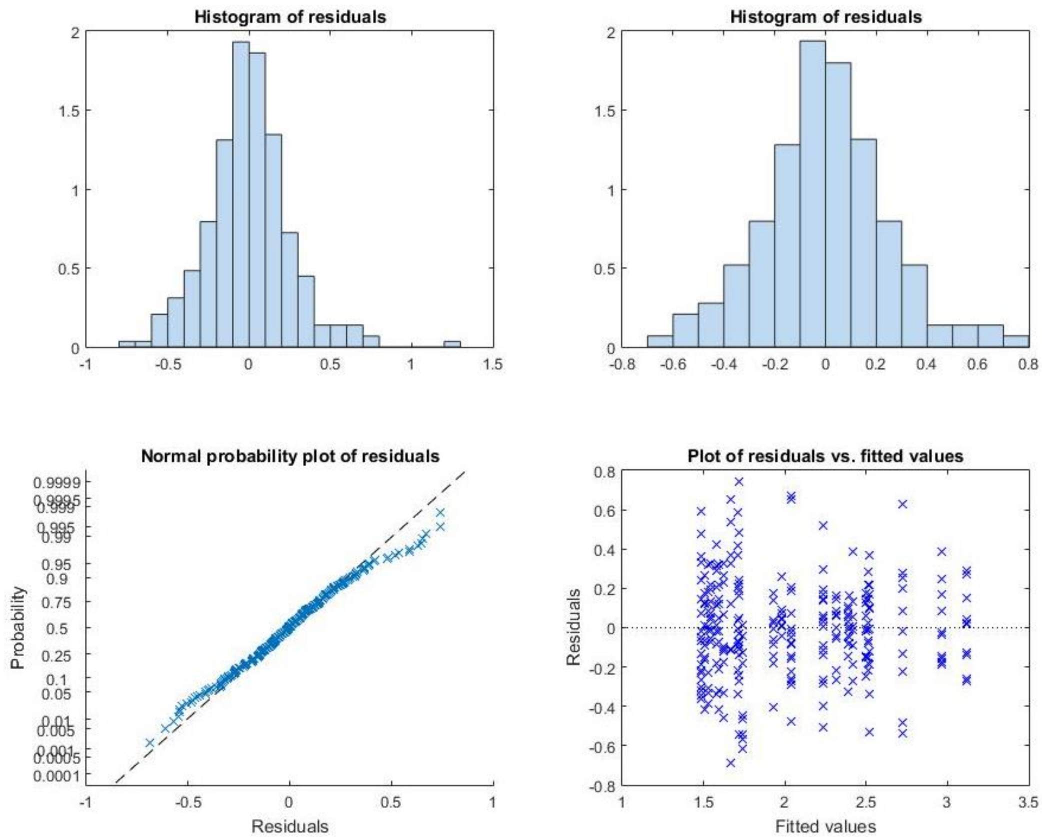


Figure 6: Plot of histogram of residuals and normal probability of residual for securing time model

This model was built using the training data, and three testing data were used to test the prediction model. Tab. 6 shows the comparison of prediction value and experimental value. For the 29 training data sets, mean experimental response was used to examine the model accuracy. According to the results, 75.9% of the experimental points distribute between the 95% confidence interval, which is basically close to the fitness correlation of 0.769 (adjusted R-square). Three new testing data that were also collected from the experiment but didn’t get use in the model building were chosen for the verification. Fig. 7 shows the 95% confidence interval (CI) of predictions, each data point represents the log-transformed securing time from five operating participants. The results show the mean response (dash line) falls between the predicted confidence

Training data					Testing data					
Experimental run	Mean experimental response	Predicted response	Predicted response CI		Experimental run	Operating participants	Experimental response	Predicted response	Predicted response CI	
1	3.88	4.41	4.09	4.75	30	1	6.6	4.50	3.96	5.11
2	10.44	9.36	8.51	10.30		1	5.26			
3	8.47	9.36	8.51	10.30		2	4.57			
4	5.18	4.50	3.96	5.11		2	3.77			
5	5.18	4.70	4.24	5.21		3	5.06			
6	7.63	7.67	6.86	8.58		3	5.66			
7	7.71	7.24	6.57	7.99		4	5.33			
8	11.81	11.20	9.81	12.80		4	4.16			
9	7.87	5.31	4.92	5.73		5	5.23			
10	12.48	12.39	10.59	14.49		5	6.14			
11	10.16	10.13	8.83	11.62	31	1	9.3	5.59	5.18	6.04
12	4.27	5.68	5.26	6.14		1	6.11			
13	10.78	10.91	9.75	12.19		2	4.47			
14	5.08	4.42	4.02	4.86		2	9.7			
15	5.36	4.86	4.34	5.44		3	7.61			
16	16.20	15.24	13.31	17.46		3	6.98			
17	23.01	22.64	20.22	25.36		4	3.57			
18	20.13	19.49	17.43	21.78		4	3.8			
19	12.35	12.17	10.48	14.14		5	5.9			
20	6.31	5.55	5.14	6.00		5	4.57			
21	12.52	12.43	10.86	14.22	32	1	17.03	17.41	15.02	20.18
22	5.00	5.06	4.36	5.86		1	21.73			
23	6.69	5.60	5.17	6.05		2	21.31			
24	4.47	4.61	4.20	5.06		2	18.97			
25	11.11	11.01	9.85	12.31		3	16.47			
26	5.80	5.58	5.23	5.95		3	13.83			
27	6.64	6.90	6.00	7.94		4	17.25			
28	8.35	7.69	6.68	8.84		4	15.26			
29	4.97	4.94	4.45	5.48		5	13.17			
-	-	-	-	-		5	14.77			

Table 6: Three experimental runs with their experimental and predicted response (s)

interval, but only 43% of experimental data falls in-between confidence bounds, which is much lower compared to the adjusted R-square value. One possible reason for this is that all experiments were done by students, even though they were trained before the experiments, there is still large variance in some operations, as shown in Fig. 8. Moreover, the variable combinations requirements for BBD couldn't be meet perfectly in the experiment, some of the experiment runs are missing because of the selection of product and experimental cost. Therefore, a better method need to be developed to eliminate the person-to-person variance effect and enhance the model accuracy in the further work.

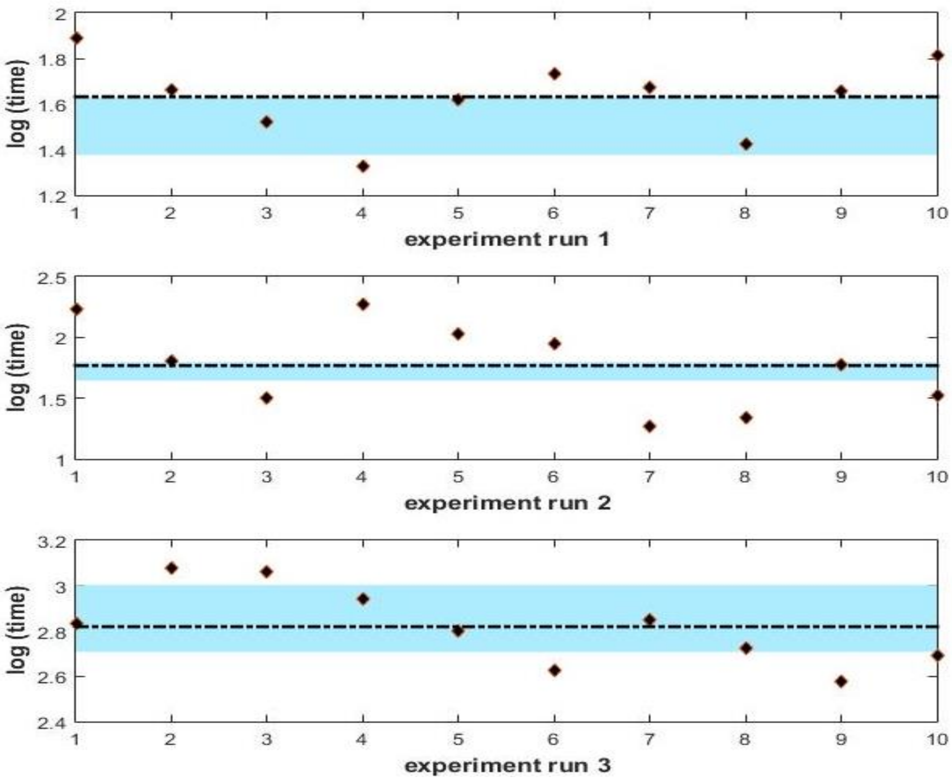


Figure 7: Predicted log time vs. experimental value. The blue section define the 95% confidence bound for predicted log time; the diamond points are experimental value from 5 participants; dash line is the mean response for the experimental value.

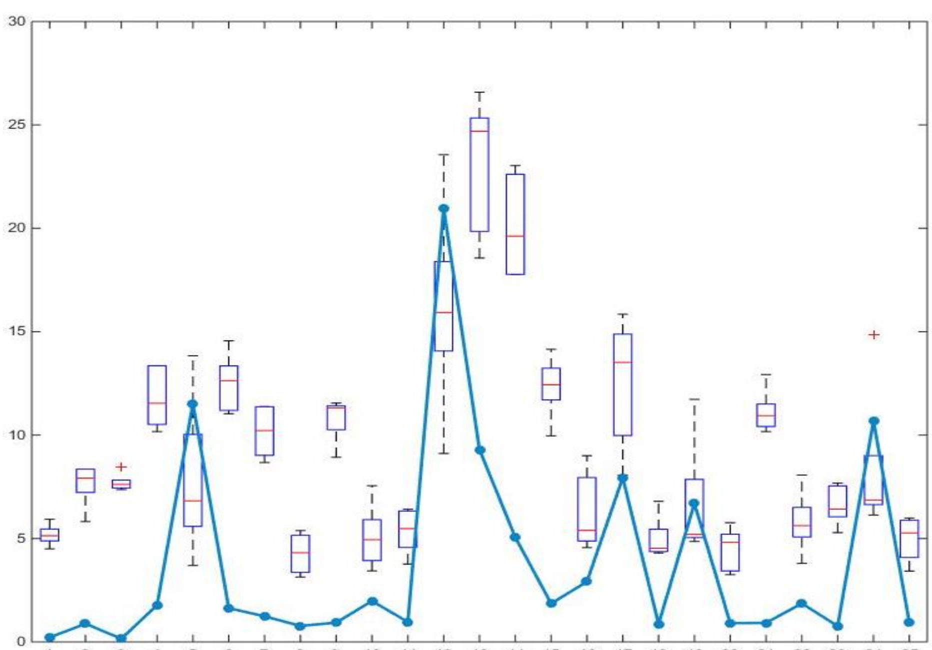


Figure 8: Person to person variance in the experiments. Box plot defines the range and mean response for each experiment run; the blue point and solid line are used to represent the corresponding variance.

7. Conclusion

Securing operation time for various type of fasteners was optimized using Box-Behnken design of experiments based on the regression analysis of experimental data. The securing prediction model could be a support to develop an automated time prediction tool for evaluating assembly times for sequence planning. Three assembly experiments were conducted to obtain the training data to ensure product diversity. A HD camera was used to accurately record time data. The Box-Behnken experimental design is an efficient method to predict securing time from limited number of experimental runs with the help of stepwise regression analysis. Out of seven different variables chosen for the study, six were found to be significant for predicting a securing operation time. Given the experimental data for building model, the predicted value had an adjusted R-square value of 0.769, which could be considered as a good fit of the chosen model in analyzing the experimental data.

The ANOVA results indicate that normal length, diameter, thread number, insertion difficulty, tool effect, nut, normal length squared, diameter squared and thread number squared are the main parameters, which have significant influence on securing time. The interactions between diameter and insertion difficulty; diameter and tool effect; thread number and tool effect; thread number and nut option; insertion difficulty and nut option also have great effects on securing time.

Based on the experimental design and analysis, the BBD approach was used to develop a predict model with low cost and high efficiency. The results indicate that the prediction model can fit the training data with 76.9% variability in the response; but only 43% for testing data because of the larger experiment variance from participants.

For the further work, more experimental data need be obtained from more complex assemblies to achieve higher accuracy. More analyses need to be conducted to eliminate the person-to-person variance during the experiments.

8. Reference

[1] Huang, W., Liu, Y., Hoyle, C., & Campbell, M. (n.d.). Automated assembly operation time prediction using the Gaussian process method.

[2] Liao, X., Li, Q., Yang, X., Zhang, W., & Li, W. (2008). Multiobjective optimization for crash safety design of vehicles using stepwise regression model. *Structural and Multidisciplinary Optimization*, *35*(6), 561-569.

[3] Namouz, E. (2013). Automated Complexity Based Assembly Time Estimation Method.

[4] Namouz, E. Z., &Summers, J. D. (2014). Comparison of graph generation methods for structural complexity based assembly time estimation. *Journal of Computing and Information Science in Engineering*, *14*(2), 021003.

[5] Arbizu, I. P., & Perez, C. L. (2003). Surface roughness prediction by factorial design of experiments in turning processes. *Journal of Materials Processing Technology*, *143*, 390-396.

[6] Tso, G. K., &Yau, K. K. (2007). Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks.*Energy*, *32*(9), 1761-1768.

[7] G. Boothroyd, P. Dewhurst, and W. Knight, Product Design for Manufacture Marcel Dekker, 1994.

[8] Muthukumar, M., Mohan, D., &Rajendran, M. (2003). Optimization of mix proportions of mineral aggregates using Box Behnken design of experiments. Cement and Concrete Composites, 25(7), 751-758.

[9] Wu, L., Yick, K. L., Ng, S. P., & Yip, J. (2012). Application of the Box–Behnken design to the optimization of process parameters in foam cup molding. Expert Systems with Applications, 39(9), 8059-8065.

[10] Huang, C., & Townshend, J. R. G. (2003). A stepwise regression tree for nonlinear approximation: applications to estimating subpixel land cover.*International Journal of Remote Sensing*, *24*(1), 75-90.