Int J Adv Manuf Technol (2006) 27: 990–998
DOI 10.1007/s00170-004-2285-2

ORIGINAL ARTICLE

Lee-Ing Tong · Chung-Ho Wang · Lin-Chan Hsiao

Optimizing processes based on censored data obtained in repetitious experiments using grey prediction

Received: 30 October 2003 / Accepted: 3 June 2004 / Published online: 10 August 2005

Abstract The design of experiment (DOE) has been extensively adopted to increase the efficiency of designing new products and developing manufacturing processes in industry. However, some designed experiments cannot be completed for some uncontrollable reasons, such as cost and time restrictions or power damage during the experiment. Under such circumstances, incomplete data obtained in the experiment are referred to as censored data. Conventional approaches to analyzing censored data are computationally complex and frequently depend on assumptions of the normality of data. This study presents a procedure for analyzing the censored data obtained in repetitious experiments using the grey system theory. The proposed procedure does not make any statistical assumption and is less conceptual and computationally complex than current methods. Two experiments – one conventional experiment with type II censoring and one Taguchi experiment with type I censoring – are performed to demonstrate the effectiveness of the proposed procedure.

Keywords Censored data · Grey system theory · Repetitious experiments

1 Introduction

The design of experiments (DOE) has been widely adopted in industry to improve the efficiency of the development of new products and to improve their quality. However, only some experiments can be completed in some circumstances. For example, the product life-testing experiment must be completed before the life-time of the product is ended because waiting for long-term test results will delay the launching of the product into the market. Restrictions on equipment and techniques or uncontrollable power failures that interrupt the experiment or censure the experimental data are other reasons for the incompleteness of DOEs. The censored data consist of a single type or many types of censored samples. A censored sample is defined as a sample specimen from a restricted area in a sample space. These censored samples must be combined with uncensored samples to yield pseudo-complete experimental data to be analyzed to improve the quality of the product or process.

The types of censored samples include left-censored samples, right-censored samples, interval-censored samples and multiple censored samples. Nelson [10] described in detail these types of censored samples. Censored samples can also be classified into the following two types – type I censored samples and type II censored samples. A terminal response value assigned before experiments are conducted on N specimens yields a type I censored samples. For example, a life-testing experiment is conducted on ten light bulbs. The experiment maybe terminated once a life test reaches more than 1000 hours. The number of non-failing light bulbs is then a censored observation, and a random variable. A terminal point for censored observations assigned in advance of conducting experiments on N specimens yields a type II censored sample. That is, the response value is a random variable. For example, a life-testing experiment is conducted on ten light bulbs. The experiment is terminated once eight light bulbs fail, if the experiment termination value for the number of non-failed light bulbs is set to two. In such a case, the life times of non-failed light bulbs are censored observations. Accordingly, the failure time of the eighth light bulb is a random variable.

Some studies [6, 8, 9] of censored data analysis have sought to improve the quality of a product or process. However, the developed methods are either computationally complex or depend on a strictly statistical assumption about the data. This study de-

L.-I. Tong (✉)
Department of Industrial Engineering and Management,
National Chiao Tung University,
Hsin-Chu, Taiwan, R.O.C.
E-mail: litong@ccit.edu.tw
Tel.: +886-3-5738196
Fax: +886-3-5733873

C.-H. Wang
Department of Computer Science, Chung Cheng Institute of Technology,
National Defense University,
Taoyuan, Taiwan, R.O.C. 335

L.-C. Hsiao
Department of Industrial Engineering and Management,
National Chiao Tung University,
Hsin-Chu, Taiwan, R.O.C.

velops a novel procedure for solving the problem of censored data obtained in repetitious experiments using grey prediction analysis, taken from the grey system theory. Grey prediction analysis is computationally simple and does not require any statistical assumptions, so engineers without strong statistical backgrounds can use this censored data analysis procedure. Two real experiments, one for type II censored data from a conventional experiment and the other for type I censored data from a Taguchi experiment, were performed using the proposed procedure. The results were compared to those of other developed methods to demonstrate the effectiveness of the proposed procedure.

2 Literature review

Hahn et al. [7] developed the iterative least squares (ILS) method to solve a problem that involved left censored data from a 2^{4-1} factorial experimental design. The characteristics of truncated normal distributions were used to calculate the expected values of the censored observations. An initial regression model fitted using the standard least square method was obtained by determining the censored values at their censoring points or under specific conditions. Iterations were terminated as soon as the sum of the absolute values of the pairwise differences between parameters in two successive iterations fell below a specified low value. Accordingly, the fitted regression model could be used to predict the censored data. Although truncated samples were used to estimate the parameters in the regression model, the censoring points were used to replace the censored data during the first iterative step. In such a case, the determination of significant factors and the follow-up fitted regression model may be biased. Moreover, the iterative procedure is not practical for use by engineers because of its computational complexity.

Taguchi [12] developed the minute accumulating analysis (MAA) method to predict interval-censored data. The MAA divides time into many intervals and assumes that the time intervals are independent of each other. Binary data bits (0 or 1) are initially assigned according to the situation when a specimen test fails or survives in the specified interval. The MAA applies analysis of variance (ANOVA) to the generated binary data, treating them as having been in a split-plot experiment. The factors studied herein are main-plot factors and the generated time is considered to be a sub-plot factor. A shortcoming of the MAA is that the assumption of independence of the time intervals does not hold. Moreover, a large number of degrees of freedom in the ANOVA table make the ANOVA result unconvincing. Furthermore, MAA neglects censoring information, and treats the censoring time as the actual failure time, leading to the making of incorrect decisions to improve the product or process quality.

The Hamada-Wu (HW) procedure [8] established a linear model of response value on the factor/level combination, using the maximum likelihood estimation (MLE) method with an iterative procedure to predict the censored data. The HW procedure replaced the ILS procedure by the MLE method to solve the problems of ILS and MAA, neglecting censoring information. The HW procedure considers only responses of quantities, such as failure times. The idea behind the HW procedure is that modeling positive response using power transformation is convenient. The convenient characteristics of the normally distributed responses can be exploited. Accordingly, an initial normality transformation must be applied to the data if the data are not normally distributed. The censored data are then predicted using the MLE method with an iterative procedure. The HW simulation procedure is more accurate than ILS. Additionally, the HW procedure can check the adequacy of the model and determine the effects of significant factors on the response. However, in many cases, the initial model does not incorporate the MLE [9] and the MLE calculations are hard to perform.

Tong and Su [13] used regression analysis and a non-parametric method to solve the problem of censored data in repetitious experimental design. Using this method, a regression model is initially established based on non-censored data to predict the censored data. Ranks are then assigned to all responses, including the predicted censored data. Two regression models of the mean and variance are established using the ranked data. The optimal combination of factor levels is determined accordingly. Tong and Su [13] overcame the shortcomings of the MAA and ILS methods, neglecting censoring information. However, the means of determining significant factors remains uncertain because significant factors are determined subjectively, from a normality probability plot.

Su and Miao [11] developed a procedure to solve the problem of right and left censored data in repetitious experiments, using a back-propagation (BP) neural network. Although the neural network does not depend on any statistical assumptions, the construction of the network, including the determination of the number of hidden layers and the corresponding number of neurons, must be trained using a complex trial and error process. Moreover, a neural network cannot determine which factor significantly affects the response. Additionally, the constructed network depends on a large number of data and often yields inconsistent results when different people seek to solve the same problem.

3 Grey prediction

Deng [3] developed the grey system theory in 1982. This method can effectively solve problems that are uncertain or incomplete, or which involve systems with incomplete information, using system relational analysis, model construction, and forecasting and decision analysis. The grey prediction method is the most important component of the grey system theory and is often used to describe and analyze future data from past and current data. The calculation of grey prediction is rather simple because it requires only a few samples. The grey system theory has been successfully used in various fields, including industry, agriculture, economics and civil engineering. The grey prediction method involves accumulated generating operation (AGO), the inverse generating operation (IAGO) and grey modeling. Deng [4] described these operations in detail. The method is described briefly below.

3.1 Grey generating

Grey generating is a data processing method. The idea behind grey generating is to determine a regular pattern and obtain effective information while reducing the disorder and randomness of the data. The grey prediction model is constructed from grey generating data.

Let $x^{(0)}$ be the original data sequence, expressed as follows.

$$x^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)\right) = \left(x^{(0)}(k);\ \ k = 1, 2, 3, \dots, n\right) . \quad (1)$$

The operation of the first order AGO (1-AGO) of $x^{(0)}$ is expressed as follows.

$$x^{(1)} = \left(\sum_{k=1}^{1} x^{(0)}(k), \sum_{k=1}^{2} x^{(0)}(k), \dots, \sum_{k=1}^{n} x^{(0)}(k)\right) \quad (2)$$

where $x^{(1)}$ is the 1-AGO of $x^{(0)}$. The sequence $x^{(0)}$ can be obtained after the IAGO operation is performed on sequence $x^{(1)}$, represented as IAGO $x^{(1)} = x^{(0)}$. That is, $x^{(0)}$ can be expressed as follows.

$$x^{(0)}(k) = x^{(1)}(k) - x^{(1)}(k+1) . \quad (3)$$

3.2 Grey modeling

Grey modeling (GM) generally involves GM (1, 1), GM (1, N) and GM (0, N) models, of which GM (1, N) represents the first order derivative model, and N is the number of variables. Xue [14] developed the second order differential equation model, GM (2, 1). Grey modeling involves establishing a model with the properties of sequential linear differential equations, called a grey differential equation model. The conventional first-order differential equation model is expressed as follows.

$$\frac{\mathrm{d}x}{\mathrm{d}t} + ax = b . \quad (4)$$

The differential equation is used mainly to differentiate continuous data. However, the sequence data from the grey system is discrete and non-differentiable. Accordingly, the notion of constructing a grey differential equation model is self-contradictory. Deng [3] used the conventional linear differential equation model to construct a first-order linear differential equation based on $x^{(1)}$, as follows.

$$\frac{\mathrm{d}x^{(1)}}{\mathrm{d}t} + ax^{(1)} = b . \quad (5)$$

Accordingly, the most commonly used GM (1, 1) model for grey prediction is constructed as follows.

$$x^{(0)}(k) + aZ^{(1)}(k) = b \quad (6)$$

$$Z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1) \quad (7)$$

where "a" and "b" are parameters of the GM (1, 1) model: "a" represents a developing coefficient and "b" represents a grey input to the corresponding model. Parameters "a" and "b" are obtained using the least square method. The following equation for the GM (1, 1) model can be obtained using Eq. 6 for $k = 2, \dots, n$.

$$\begin{aligned} &x^{(0)}(2) + az^{(1)}(2) = b \\ &x^{(0)}(3) + az^{(1)}(3) = b \\ &x^{(0)}(4) + az^{(1)}(4) = b \\ &\dots \\ &x^{(0)}(n) + az^{(1)}(n) = b . \end{aligned} \quad (8)$$

Let

$$\boldsymbol{Y}_N = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ -z^{(1)}(4) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \quad \hat{\boldsymbol{a}} = \begin{bmatrix} a \\ b \end{bmatrix} .$$

Eq. 8 can also be expressed in matrix form, as follows.

$$\boldsymbol{Y}_N = \boldsymbol{B}\hat{\boldsymbol{a}} . \quad (9)$$

The following equation can be obtained after performing the matrix operation using Eq. 9.

$$\boldsymbol{B}^T \boldsymbol{Y}_N - \boldsymbol{B}^T B\hat{\boldsymbol{a}} = 0 . \quad (10)$$

The model parameters "a" and "b" can be obtained using Eq. 11, which is derived by the least squares method.

$$\hat{\boldsymbol{a}} = (\boldsymbol{B}^T \boldsymbol{B})^{-1} \boldsymbol{B}^T \boldsymbol{Y}_N . \quad (11)$$

The whitened response equation can then be obtained after substituting the values of the model parameters "a" and "b" determined from Eq. 11 into the GM (1, 1) model, as follows.

$$x^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right] \mathrm{e}^{-ak} + \frac{b}{a} . \quad (12)$$

The sequence $x^{(0)}$ is obtained by applying the IAGO operation to the sequence $x^{(1)}$. Therefore, $x^{(0)}(k)$ is used as a predictive model, as follows.

$$\begin{aligned} x^{(0)}(k+1) &= x^{(1)}(k+1) - x^{(1)}(k) \\ &= (1 - \mathrm{e}^{a}) \left[x^{(0)}(1) - \frac{b}{a}\right] \mathrm{e}^{-ak} . \end{aligned} \quad (13)$$

A prerequisite for the use of the grey prediction model is that the class ratio should be between 0.1353 and 7.389 in constructing the GM (1, 1) model, based on $x^{(0)}$ [5]. Let $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n))$ be an original sequence; the class ratio of $x^{(0)}$ is defined as follows.

$$\sigma^{(0)}(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, \quad k = 2, 3, \dots, n . \quad (14)$$

The procedure for constructing a grey prediction model is summarized below.

Step 1 Given original data $x^{(0)}(k)$, $k = 1, \ldots, n$, form an original sequence $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \ldots, x^{(0)}(n))$ and determine the class ratio.
Step 2 Apply AGO to $x^{(0)}(k)$ to generate $x^{(1)}(k)$.
Step 3 Compute $Z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)$.
Step 4 Calculate the model parameters "a" and "b" using the least square method based on Eq. 11.
Step 5 Substitute model parameters "a" and "b" obtained from Step 4 in the whitened response equation, Eq. 12, to yield the GM (1, 1) model, proceeding with the prediction.

4 Proposed procedure

The proposed procedure for analyzing censored data includes the following two stages.

Stage I. Predict the censored data in each experimental run

Step 1 Distinguish the censored data Y_C from the uncensored data Y_U obtained in the repetitious experiments.
Step 2 Construct an original sequence $x^{(0)}$ based on the uncensored data Y_U of each experimental run and check the rationality of the corresponding class ratio.
Step 3 Perform AGO to construct GM (1, 1) models based on $x^{(0)}$ obtained in Step 2 in each experimental run.
Step 4 Predict the uncensored data in each experimental run using the GM (1, 1) model obtained in Step 3.
Step 5 Combine the predicted censored data generated in Step 4 with the uncensored data, to yield pseudo-complete experimental data.

Stage II. Perform ANOVA on the pseudo-complete experimental data obtained in Step 5 of Stage I
to determine the optimal combination of factor levels

Figure 1 presents the proposed procedure of two-stage censored data analysis.

5 Illustrative examples

Two experiments, one involving type II censored data obtained in a conventional experiment and one involving type I censored data obtained in a Taguchi experiment, were used to demonstrate the effectiveness of the proposed procedure. These two experiments originally contained no censored data. They were also analyzed using artificially manipulated censored data to demonstrate the effectiveness of the proposed procedure. Comparisons between the censored data and the uncensored data obtained in the original experiment were drawn. These two experiments were described as follows.

Experiment 1: type II censored data obtained
in a conventional repetitious experiment

The aim of this experiment, taken from Condra [2], was to determine the significant factors and the optimal combination of factor levels to maximize the life of a surface-mounted capacitor. Table 1 lists the control factors and their corresponding levels. The experimental response was time-to-failure. Each run involved eight repeated experiments. Restated, eight capacitors were tested in each experimental run. Table 2 summarizes the results of the original experiment which contained no censored data. To imitate the problems associated with type II right-censored data, this experiment is artificially terminated after once six surface-mounted capacitors fail. That is, the experiment was terminated when the number of non-failing capacitors equaled two in the lift-testing experiments performed on the eight capacitors. Table 3 summarizes the ranked results of the lifetime-tests. According to Table 3, the lifetime data in the last

Table 1. Control factors and their corresponding levels

	Control factor	Level	
A	Dielectric composition	Low	High
B	Processing temperature	Low	High

Fig. 1. Proposed procedure

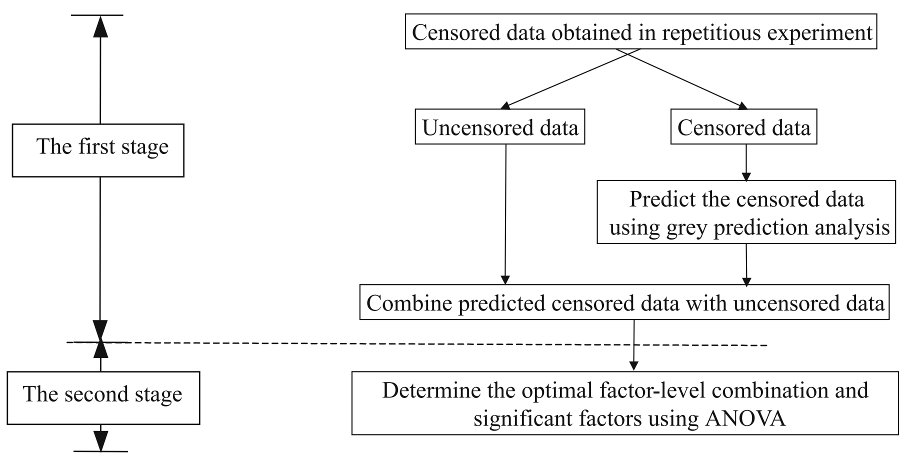

Table 2. Original experimental data

No	A	B	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
1	1	1	430	950	560	210	310	230	250	230
2	1	2	1080	1060	890	450	430	320	340	430
3	2	1	890	1060	680	310	310	310	250	230
4	2	2	1100	1080	1080	460	620	370	580	430

Table 3. Ranked results from the life-tests

No	A	B	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
1	1	1	210	230	230	250	310	430	560	950
2	1	2	320	340	430	430	450	890	1060	1080
3	2	1	230	250	310	310	310	680	890	1060
4	2	2	370	430	460	580	620	1080	1080	1100

two columns (variables y_7 and y_8) are imitated as censored observations with unknown value.

According to the proposed procedure, the experimental data were analyzed as follows.

Stage I. Predict the censored data in each experimental run

Step 1 Distinguish the censored data Y_C and the uncensored data Y_U obtained in repetitious experiments in each experimental run.

Table 4. Uncensored data Y_U and censored data Y_C

No	A	B	Y_U						Y_C	
			y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
1	1	1	210	230	230	250	310	430		
2	1	2	320	340	430	430	450	890		
3	2	1	230	250	310	310	310	680		
4	2	2	370	430	460	580	620	1080		

Table 5. Original sequences and corresponding class ratio

No	$x^{(0)}$	Class ratio (0.1353–7.389)
1	(210, 230, 230, 250, 310, 430)	(–, 0.91, 1.00, 0.92, 0.81, 0.72)
2	(320, 340, 430, 430, 450, 890)	(–, 0.94, 0.79, 1.00, 0.96, 0.51)
3	(230, 250, 310, 310, 310, 680)	(–, 0.92, 0.81, 1.00, 1.00, 0.46)
4	(370, 430, 460, 580, 620, 1080)	(–, 0.86, 0.93, 0.79, 0.94, 0.57)

Table 7. Predicted censored data in each experimental run

No	The predicted censored data	
1	$x^{(0)}(7) = 477.185$,	$x^{(0)}(8) = 571.777$
2	$x^{(0)}(7) = 980.143$,	$x^{(0)}(8) = 1261.268$
3	$x^{(0)}(7) = 742.012$,	$x^{(0)}(8) = 973.303$
4	$x^{(0)}(7) = 1243.65$,	$x^{(0)}(8) = 1608.261$

Table 8. Pseudo-complete experimental data (censored observations are given in the last two columns)

No	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
1	210	230	230	250	310	430	477.185*	571.777*
2	320	340	430	430	450	890	980.143*	1261.268*
3	230	250	310	310	310	680	742.012*	973.303*
4	370	430	460	580	620	1080	1243.65*	1608.261*

Table 4 summarizes the uncensored data Y_U and censored data Y_C.

Step 2 Construct the original sequences $x^{(0)}$ based on the uncensored data Y_U of each experimental run and check the rationality of the corresponding class ratio.
Table 5 summarizes the original sequences and the corresponding class ratio. All the values in the class ratio fall between the rational interval, 0.1353 and 7.389.

Step3 Perform AGO to construct GM (1, 1) models based on $x^{(0)}$ obtained in Step 2 in each experimental run.
Table 6 shows the GM (1, 1) models in each experimental run.

Step 4 Predict the uncensored data in each experimental run using the GM (1, 1) model obtained in Step 3.
Table 7 presents the predicted censored data in each experimental run.

Step 5 Combine the predicted censored data generated in Step 4 with the uncensored data to yield pseudo-complete experimental data.
Table 8 presents the pseudo-complete experimental data.

Stage II. Perform ANOVA on the pseudo-complete experimental data obtained from Step 5 in Stage I to determine the optimal combination of factor levels

Table 9 presents the ANOVA results. Factor B significantly affects the life of the capacitor, according to Table 9, because its p-value = 0.012091, which is below the level of significance $\alpha = 0.05$. Moreover, the interaction between factors A and B is

Table 6. GM (1, 1) models in each experimental run

No	$x^{(0)}$	GM (1, 1) models
1	(210, 230, 230, 250, 310, 430)	$x^{(0)}(k+1) = 974.5835(1-e^{-0.18085})e^{0.18085k}$
2	(320, 340, 430, 430, 450, 890)	$x^{(0)}(k+1) = 968.482(1-e^{-0.25217})e^{0.25217k}$
3	(230, 250, 310, 310, 310, 680)	$x^{(0)}(k+1) = 613.0233(1-e^{-0.27133})e^{0.27133k}$
4	(370, 430, 460, 580, 620, 1080)	$x^{(0)}(k+1) = 1172.946(1-e^{-0.2571})e^{0.2571k}$

Table 9. ANOVA result of pseudo-experimental data (bold type indicates significance)

Source	SS	df	MS	F	p-value
A	178033.3	1	178033.3	1.654911	0.208827
B	**774715**	**1**	**774715**	**7.201373**	**0.012091**
A × B	1177.909	1	1177.909	0.010949	0.917408
Error	3012206	28	107578.8		
Total	3966132	31			

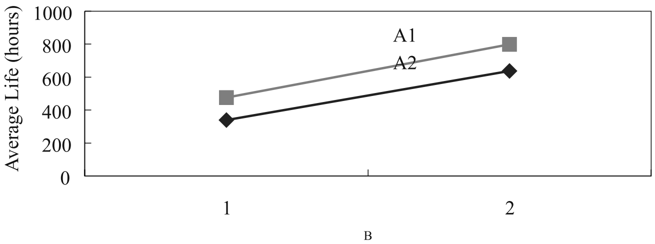

Fig. 2. Main effects of factors A and B

Table 10. ANOVA results of the original experiment contained no censored data (bold type indicates significance)

Source	SS	df	MS	F	p-value
A	79003.13	1	79003.13	0.83488	0.368667
B	**385003.1**	**1**	**385003.1**	**4.068591**	**0.053367**
A × B	703.125	1	703.125	0.00743	0.931921
Error	2649588	28	94628.13		
Total	3114297	31			

insignificant. Accordingly, the optimal combination of factor levels can be determined from individual factor effects. Figure 2 shows the main effects of factors A and B. The optimal factor level is determined as A_1B_2, according to Fig. 2, because it yields the largest desired response value.

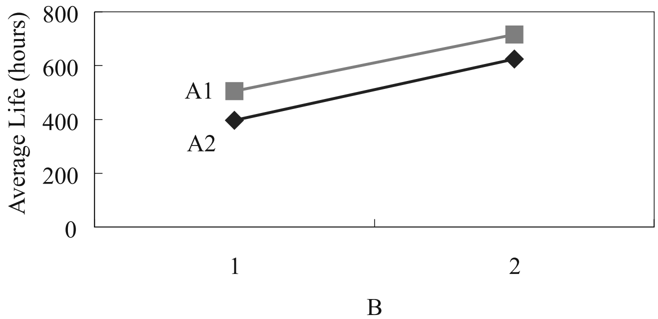

Fig. 3. Main effect of factors A and B in the original experiment contained no censored data

The original experiment, which contained no censored data, was also performed using the ANOVA method. Table 10 presents the ANOVA results. According to Table 10, the p-value of factor B is 0.053367, which is approximately the level of significance $\alpha = 0.05$. Consequently, factor B significantly affects the life of the capacitor. The interaction between factors A and B is insignificant. Accordingly, the individual factor effect determines the optimal combination of factor levels can be determined according to the individual factor effect. Figure 3 shows the main effects of factors A and B. The optimal combination of factor levels was determined as A_1B_2, from Fig. 3. The pseudo-complete experimental data yield ANOVA results that are consistent with those of the original experiment which contained no censored data. Hence, the proposed procedure can effectively solve the problem associated with the censored data obtained in repetitious experiments.

Experiment 2: type I censored data obtained in a Taguchi's experimental design

This experiment was taken from Byrne and Taguchi [1], who performed an experiment on the fabrication of elastomeric connectors using a car engine's nylon pipe. The aim of their experiment was to maximize the pull-off performance. Accordingly, a larger

Table 11. Experimental pull-off performance (bold type indicates a pull-off weight above 22 lbs)

					8	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**No.**
					2	2	2	2	1	1	1	1	**E**
					2	2	1	1	2	2	1	1	**F**
					1	1	2	2	2	2	1	1	**E × F**
					2	1	2	1	2	1	2	1	**G**
					1	2	1	2	2	1	2	1	**E × G**
					1	2	2	1	1	2	2	1	**F × G**
					2	1	1	2	1	2	2	1	**e**
No	**A**	**B**	**C**	**D**	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	**S/N**
1	1	1	1	1	19.1	20	19.6	19.6	19.9	16.9	9.5	15.6	24.025
2	1	2	2	2	21.9	**24.2**	19.8	19.7	19.6	19.4	16.2	15	25.522
3	1	3	3	3	20.4	**23.3**	18.2	**22.6**	15.6	19.1	16.7	16.3	25.335
4	2	1	2	3	**24.7**	**23.2**	18.9	21	18.6	18.9	17.4	18.3	25.904
5	2	2	3	1	**25.3**	**27.5**	21.4	**25.6**	**25.1**	19.4	18.6	19.7	26.908
6	2	3	1	2	**24.7**	**22.5**	19.6	14.7	19.8	20	16.3	16.2	25.326
7	3	1	3	2	21.6	**24.3**	18.6	16.8	**23.6**	18.4	19.1	16.4	25.711
8	3	2	1	3	**24.4**	**23.2**	19.6	17.8	16.8	15.1	15.6	14.2	24.832
9	3	3	2	1	**28.6**	**22.6**	**22.7**	**23.1**	17.3	19.3	19.9	16.1	26.152

Table 12. Pseudo-complete experimental data (censored observations are on the right)

No	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
1	9.5	15.6	16.9	19.1	19.6	19.6	19.9	20
2	15	16.2	19.2	19.4	19.7	19.8	21.9	22.5769*
3	15.6	16.3	16.7	18.2	19.1	20.4	21.5536*	22.8578*
4	17.4	18.3	18.6	18.9	18.9	21	20.9401*	21.5843*
5	19.7	18.6	19.4	21.4	22.7891*	24.4757*	26.2872*	28.2328*
6	14.7	16.2	16.3	19.6	19.8	20	21.8742*	23.2105*
7	16.4	16.8	18.4	18.6	19.1	21.6	22.2243*	23.4832*
8	14.2	15.1	15.6	16.8	17.8	19.6	20.6440*	22.0701*
9	16.1	17.3	19.3	19.9	21.5465*	23.0672*	24.6953*	26.4383*

response is desired. Table 11 summarizes the experimental data on pull-off performance, wherein A, B and C represent the control factors and E, F and D represent the noise factors. Terminal pull-off performance was determined at 22 lbs to imitate type I right censored data obtained in repetitious experiments. Thus, pull-off weights above 22 lbs in Table 1 are taken as the censored observations.

The experimental data are analyzed as follows.

Step 1 – Step 5 Predict the censored data and combine them with the uncensored data to generate pseudo-complete experimental data.
Table 12 summarizes the pseudo-complete experimental data.

Step 6 Calculate the signal to noise (S/N) ratio using the Taguchi method.
The S/N ratio in each experimental run can be obtained using the S/N ratio equation for the larger-the-better response. Table 13 presents the results.

Step 7 Perform ANOVA on the SN values.
Table 14 presents the results of applying ANOVA to the S/N values obtained in Step 6. The sum of the squares (SS) of factor B is pooled with the SS of factor D to yield an SS for the error in ANOVA. Factors A and C significantly affect the pull-off performance, according to the corresponding contribution percentage and the F values taken from Table 14.
Table 15 summarizes the S/N ratios given each combination of factor levels. Figure 4 plots the corresponding factor effects. The optimal combination of factor levels was determined to be $A_2B_2C_3D_1$, which maximizes the pull-off performance.

Table 13. S/N ratio in each experimental run

No	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	S/N
1	0.1053	0.0641	0.0592	0.0524	0.0510	0.0510	0.0503	0.0500	24.0253
2	0.0667	0.0617	0.0521	0.0515	0.0508	0.0505	0.0457	0.0443	25.4518
3	0.0641	0.0613	0.0599	0.0549	0.0524	0.0490	0.0464	0.0437	25.2856
4	0.0575	0.0546	0.0538	0.0529	0.0529	0.0476	0.0478	0.0463	25.7120
5	0.0508	0.0538	0.0515	0.0467	0.0439	0.0409	0.0380	0.0354	26.8287
6	0.0680	0.0617	0.0613	0.0510	0.0505	0.0500	0.0457	0.0431	25.2648
7	0.0610	0.0595	0.0543	0.0538	0.0524	0.0463	0.0450	0.0426	25.6408
8	0.0704	0.0662	0.0641	0.0595	0.0562	0.0510	0.0484	0.0453	24.6922
9	0.0621	0.0578	0.0518	0.0503	0.0464	0.0434	0.0405	0.0378	26.1277

Table 14. ANOVA results

Source	df	SS	MS	S'	ϱ(%)	F
A	2	1.5500	0.7750	1.1709	22.82%	4.0884
B*	2*	0.4799*				
C	2	2.8220	1.4110	2.4429	47.62%	7.4437
D*	2*	0.2783*				
Pooled e	4	0.7582	0.1896	1.5165	29.56%	
Total	8	5.1303		5.1303	100%	

Table 15. S/N values for each combination of factor level

	A	B	C	D
Level 1	24.9606	25.2135	24.7278	25.6950
Level 2	26.0458	25.7538	25.8593	25.5194
Level 3	25.5650	25.6042	25.9844	25.3571

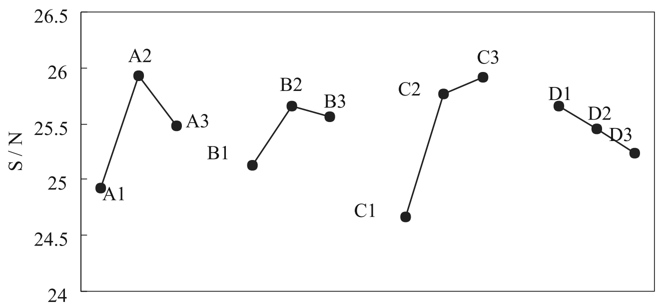

Fig. 4. Effect of each factor

Table 16. ANOVA result obtained using MAA

Source	df	SS	MS	S'	ϱ (%)	F
S_m	1	548.33	548.33		83%	
A	2	0.77	0.38	0.44	0.12%	2.33
B	2	0.13*	0.07			
C	2	1.41	0.71	1.08	0.21%	4.29
D	2	0.07*	0.04			
e_1	63	10.83*	0.17			
Pooled_e_1	67	11.03	0.16			
ω	10	55.73	5.57	55.08	8%	
$A \times \omega$	20	1.37	0.07	0.06	0.21%	1.07
$B \times \omega$	20	0.76**	0.04			
$C \times \omega$	20	1.81	0.09	0.49	0.27%	1.42
$D \times \omega$	20	0.73**	0.04			
e_2	630	37.05*	0.07			
Pooled_e_2	670	38.54	0.06			
Total	792	659				

Table 17. Effects of each factor level

	A	B	C	D
Level 1	0.7917	0.8144	0.7727	**0.8447**
Level 2	**0.8674**	0.8371	0.8561	0.8295
Level 3	0.8434	**0.8447**	**0.8670**	0.8220

Table 18. ANOVA results obtained using the Taguchi method

Source	df	SS	MS	S'	$\varrho(\%)$	F
A	2	1.7743	0.8872	1.4552	27.52%	5.5593
B	2*	0.4670*				
C	2	2.8749	1.4374	2.5557	48.33%	9.0076
D	2*	0.1713*				
Pooled e	4	0.6383	0.1596	1.2766	24.14%	
Total	8	5.2875		5.2875	100.00%	

Table 19. S/N ratio of each factor level

	A	B	C	D
Level1	24.9606	25.2135	24.7278	**25.6950**
Level 2	**26.0458**	**25.7538**	25.8593	25.5194
Level 3	25.5650	25.6042	**25.9844**	25.3571

Pseudo-complete experimental data were also generated by the MAA method; the results were compared with those obtained by the proposed procedure. The value of the response obtained by the MAA method was divided into 11 intervals because the censored point was set to be 22. Each interval was two, such that $\omega_1 = 2, \omega_2 = 4, \ldots, \omega_{11} = 22$. A 0 or 1 binary data set was initially assigned to this case. Where a specimen failed or survived a test in the specified interval. ANOVA was then performed on this generated binary data, treating them as if they were obtained in a split-plot experiment. Table 16 presents the ANOVA results of MAA. The F value of factor C was 4.29, which exceeds $F_{0.05,2,67} = 3.08$, according to Table 16. Factor C therefore significantly affects the pull-off of the part fabrication process at the 5% significance level. Table 17 summarizes the effect of the each combination of factor levels. The optimal combination of factor levels was determined to be $A_2B_3C_3D_1$.

The original experiment, which contained no censored data, was also performed using the Taguchi method. Table 18 presents the results of ANOVA. According to Table 18, factors A and C significantly affected the pull-off part fabrication process. Table 19 summarizes S/N ratio of each combination of factor levels and Fig. 5 plots the corresponding effects. The optimal combination of factor levels is determined to be $A_2B_2C_3D_1$ according to Table 19 and Fig. 5.

The results of the analysis of the data obtained in the original experiment, which contained no censored data, were compared with those of the pseudo-complete data, using the proposed procedure, the MAA method and neural network analysis. Table 20 summarizes these comparisons. The significant factors and the

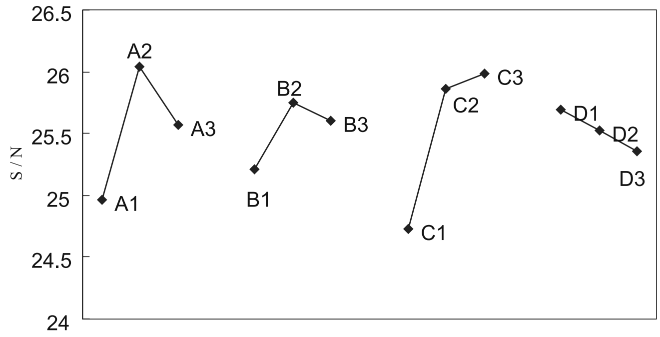

Fig. 5. Effect of each factor

Table 20. Comparison of each developed method with the original experiment contained no censored data

Method	Significant factors	Optimal factor level combination
Original experiment contained no censored data	**A·C**	$\mathbf{A_2B_2C_3D_1}$
The proposed procedure	A·C	$A_2B_2C_3D_1$
The MAA	C	$A_2B_3C_3D_1$
The BLUE	A·C	$A_2B_2C_3D_1$
The neural network	Cannot determine	$A_2B_2C_3D_1$

optimal combination of factor levels obtained using the MAA method differed from that obtained using the original experimental results which contained no censored data. The neural network analysis yielded consistent results with those of the original experiment which contained no censored data. However, the neural network analysis could not determine significant effects of the factor on the response. The proposed censored data analysis procedure yielded consistent results with those obtained in the original experiment which contained no censored data, as shown in Table 20, and the calculations are quite simple.

6 Conclusion

The conventional experimental design and the Taguchi method have been extensively applied in industry to determine the optimal combination of factor levels and significant effects of the factor on the response. These methods are essential in improving the quality of processes or products in industry. However, experimental data may be censored because of time and cost constraints, equipment constraints and other unpredictable factors. Consequently, developing an optimization procedure that solves the problem of censored data is an important issue. Some studies on censored data have used conventional statistical methods, including maximum likelihood estimation and iterative least square methods. However, these methods depend on the collection of sufficient data and assume the normality of the data; they are therefore impractical for use in engineering situations. Although, constructing a neural network may overcome the complexity of the problem of censored data. Consequently, inconsistent results can be produced when different individuals, who program the process, tackle the same problem. Additionally, neural network analysis cannot determine the significant effects of the factors on the response.

This study used the grey prediction method to solve the problem of censored data. This method does not depend on strict assumptions regarding data and can be applied effectively to few discrete data. Two experiments were performed to demonstrate the effectiveness of the proposed procedure. Results obtained using the proposed procedure were compared with those obtained using other methods. The proposed procedure provides the following advantages.

1. The grey system theory does not make strict assumptions regarding the data and it involves rather simple calculations. It is therefore practical for use by engineers without a strong statistical background.
2. The proposed procedure can be applied simultaneously to conventional experimental and Taguchi experimental designs.
3. The proposed procedure can be applied simultaneously to type I and type II censored data obtained in repetitious experimental designs. Therefore, it can be used broadly.
4. The proposed procedure reduces the number of experiments required to solve the problem of censored data when the cost (of so doing) is so high as to be unbearable by the manufacturer.

References

1. Byrne DM, Taguchi S (1987) The Taguchi approach to parameter design. Quality Progress, Milwaukee, WI, pp 19–26
2. Condra LW (1993) Reliability improvement with design of experiments. Dekker, New York
3. Deng J (1982) Control problems of grey system. Syst Control Lett 5:288-294
4. Deng J (1989) Introduction to grey system theory. J Grey Syst 1(1):1–24
5. Julong DengJ (1993) On judging the admissibility of grey modeling via class ratio. J Grey Syst 5(4):249–252
6. Hahn GJ, Nelson W (1974) A comparison of methods for analyzing censored life data to estimate relationships between stress and product life. IEEE Trans Reliab 23(1):2–11
7. Hahn GJ, Morgon CB, Schmee J (1981) The analysis of a fractional factorial experiment with censored data using iterative least square. Technometrics 23:33–36
8. Hamada M, Wu CFJ (1991) Analysis of censored data from highly fractionated experiments. Technometrics 33(1):25–38
9. Lu JC, Cemal U (1994) Process characterization and optimization based on censored data from highly fractionated experiments. IEEE Trans Reliab 43(1):145–155.
10. Nelson W (1982) Applied life data analysis. Wiley, New York
11. Su CL, Miao CL (1998) Neural network procedures for experimental analysis with censored data. Int J Qual Sci 3(3):239–253
12. Taguchi G (1987) System of experimental design. Unipub/Kraus International, White Plains, NY
13. Tong LI, Su CT (1997) A non-parametric method for experimental analysis with censored data. Int J Quality Reliab Manage 14(5):456–463
14. Xue J (1990) Prediction of tea-leaves yield with grey system models. J Grey Syst 2(2):81–88