

THE USE OF DOWNTON'S ESTIMATOR IN DUAL RESPONSE SURFACE OPTIMIZATION

CH'NG CHUAN KIM¹ & MICHAEL KHOO BOON CHONG²

Abstract. A dual response surface approach consists of two responses of a quality characteristic. These two responses are the mean response and the standard deviation (variance) response, which are estimated from an experimental design after performing a model fitting. The standard deviation response is usually estimated using the sample standard deviation. The main drawback of this estimator by means of sample standard deviation is that it is easily influenced by extreme points. For this case, the fitted model based on the sample standard deviation may not be accurate. Thus, the use of this approach may not produce the correct compromised setting. In this paper, an estimation of the standard deviation based on Downton's estimator in a dual response surface optimization is proposed. A Downton estimator is a robust estimator of standard deviation. A robust estimator is less affected by extreme points compared to the sample standard deviation. Here, a model based on a robust estimator will give better results. An example is used to illustrate the effectiveness of our proposal in optimization. In this example, mean squared error (MSE) will be used as the optimization criterion.

Keywords: Downton's estimator, dual response surface optimization, mean squared error, optimization, compromise setting

Abstrak. Kaedah sambutan dual terdiri daripada dua sambutan bagi suatu ciri an kualiti. Dua sambutan tersebut ialah sambutan min dan sambutan sisihan piawai (varians) yang dianggarkan daripada reka bentuk eksperimen selepas penyuaian model dijalankan. Sambutan sisihan piawai biasanya dianggar daripada sisihan piawai sampel. Kelemahan utama penganggar yang berdasarkan sisihan piawai sampel adalah ia mudah dipengaruhi oleh titik ekstrim. Bagi kes sedemikian, model yang tersuai berdasarkan sisihan piawai sampel adalah mungkin tidak jitu. Oleh itu, penggunaan pendekatan ini mungkin tidak dapat memberi titik kompromi yang betul. Dalam kertas kerja ini, suatu anggaran sisihan piawai berdasarkan penganggar Downton dicadangkan dalam pengoptimuman kaedah sambutan dual. Penganggar teguh kurang dipengaruhi oleh titik ekstrim berbanding dengan sisihan piawai sampel. Dalam hal ini, suatu model tersuai yang berdasarkan penganggar teguh akan memberikan keputusan yang lebih baik. Suatu contoh digunakan untuk mengilustrasikan kecekapan cadangan kami dalam pengoptimuman. Dalam contoh ini ralat kuasadua min (MSE) akan digunakan sebagai ciri pengoptimuman.

Kata kunci: Penganggar Downton, pengoptimuman sambutan dual, ralat min kuasa dua, pengoptimuman, titik kompromi

^{1&2} School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia.
Tel: 604-2644179 E-mail: chngchuankim@yahoo.com Tel: 604-6533941 Fax: 604-6570910 E-mail: mkbc@usm.my/mkbc@tm.net.my

1.0 INTRODUCTION

Practitioners realize that statistical tools like Response Surface Methodology (RSM) should be applied to improve their products so that they can continue to gain customer satisfaction and be the champion in a competitive market place. Dual response surface methodology is one of the favourite choices of practitioners in optimizing the primary response subjected to an appropriate constraint on the value of the secondary response. Due to the demand, researchers like Castillo and Montgomery [1], Copeland and Nelson [2], Fan [3], Fathi [4], Kim and Lin [5], Lin and Tu [6], Myers and Carter [7], Vining and Bohn [8], Vining and Myers [9], and others have contributed significantly to the development of dual response surface optimization.

In dual response surface optimization, models for the mean and standard deviation (or variance) of a quality characteristic are needed in the optimization. These models are based on data collected in an experimental design. Table 1 is a general form of an experimental design for a dual response surface optimization with m design points, k coded process settings, and n replications at each design point.

Table 1 General form of an experimental design for a dual response surface optimization

Design point	Coded process settings				Replication			
	\mathbf{x}_1	\mathbf{x}_2	...	\mathbf{x}_k	1	2	...	\mathbf{n}
1	-1	-1	...	-1	Y_{11}	Y_{12}	...	Y_{1n}
2	0	-1	...	-1	Y_{21}	Y_{22}	...	Y_{2n}
:	:	:	:	:	:	:	:	:
m	1	1	1	0	Y_{m1}	Y_{m2}	...	Y_{mn}

Copeland and Nelson [2], Kim and Lin [5], Lin and Tu [6], and Vining and Myers [9] make use of the arithmetic mean and the sample standard deviation (a.k.a. S.S.D.) in their model fitting. The arithmetic mean and the S.S.D. at a design point are defined as:

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad (1)$$

and

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (2)$$

respectively.

In this paper, we would like to propose Downton's estimator (a.k.a. D.E.) as a replacement for the S.S.D. in estimating the standard deviation of a design point. The Downton estimator and its properties will be discussed in the next section. Next, a data set described in Box and Draper [10] will be used to illustrate its use. Finally, conclusions will be drawn with regards to our proposal.

2.0 THE DOWNTON ESTIMATOR

Let Y_1, Y_2, \dots, Y_n be a random sample from a Normal distribution with the mean μ and variance σ^2 , that is, $Y_1, Y_2, \dots, Y_n \sim N(\mu, \sigma^2)$. If we rearrange the observations of the sample in ascending order so that $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(n)}$, Downton [11] proposed the following estimator:

$$\begin{aligned}\sigma^* &= \sqrt{\pi} \sum_{i=1}^n \frac{(2i-n-1)Y_{(i)}}{n(n-1)} \\ &= \frac{2\sqrt{\pi}}{n(n-1)} \sum_{i=1}^n \left[i - \frac{1}{2}(n+1) \right] Y_{(i)}\end{aligned}\quad (3)$$

where $\sqrt{\pi} = 1.772453851$ and σ^* is an estimator of σ at a design point.

Downton [11] mentioned that this estimator is extremely simple, unbiased (that is, the expected value of Downton's estimator is equal to standard deviation) and uncorrelated with the mean (that is, Downton's estimator is independent of the mean). In addition, this estimator has a high efficiency relative to the root-mean-squared estimator even for small values of n . Details concerning this estimator and other properties can be found in Downton [11].

3.0 EXAMPLE

This data set is taken from Box and Draper [10] and it has been used frequently by other researchers. This data set is a study made to investigate the effect of speed x_1 , pressure x_2 , and distance x_3 upon a printing machine's ability to apply coloured inks onto package labels. This experiment is a factorial design with 3^3 replicates ($n = 3$) at each design point. Since $n = 3$ the D.E. can be simplified to the form shown below:

$$\begin{aligned}
\sigma^* &= \frac{2\sqrt{\pi}}{3(3-1)} \sum_{i=1}^n \left[i - \frac{1}{2}(3+1) \right] Y_{(i)} \\
&= \frac{\sqrt{\pi}}{3} \sum_{i=1}^3 (i-2) Y_{(i)} \\
&= \frac{\sqrt{\pi}}{3} (Y_{(3)} - Y_{(1)})
\end{aligned} \tag{4}$$

Table 2 is the data set plus an additional 3 columns, which contain the computations of the arithmetic mean \bar{y} , the S.S.D., and the D.E. for each design point. We will assume that a cuboidal region is of interest, that is, $-1 \leq x_i \leq 1$ for $i = 1, 2$ and 3 . This is because Vining and Bohn [8] restricted the optimization to the cube since the

Table 2 The printing study data

Design point	x_1	x_2	x_3	y_1	y_2	y_3	\bar{y}	S.S.D.	D.E.
1	-1	-1	-1	34	10	28	24.0	12.49	14.18
2	0	-1	-1	115	116	130	120.3	8.39	8.86
3	1	-1	-1	192	186	263	213.7	42.83	45.49
4	-1	0	-1	82	88	88	86.0	3.46	3.54
5	0	0	-1	44	178	188	136.7	80.41	85.08
6	1	0	-1	322	350	350	340.7	16.17	16.54
7	-1	1	-1	141	110	86	112.3	27.57	32.49
8	0	1	-1	259	251	259	256.3	4.62	4.73
9	1	1	-1	290	280	245	271.7	23.63	26.59
10	-1	-1	0	81	81	81	81.0	0.00	0.00
11	0	-1	0	90	122	93	101.7	17.67	18.91
12	1	-1	0	319	376	376	357.0	32.91	33.68
13	-1	0	0	180	180	154	171.3	15.01	15.36
14	0	0	0	372	372	372	372.0	0.00	0.00
15	1	0	0	541	568	396	501.7	92.5	101.62
16	-1	1	0	288	192	312	264.0	63.50	70.90
17	0	1	0	432	336	513	427.0	88.61	104.57
18	1	1	0	713	725	754	730.7	21.08	24.22
19	-1	-1	1	364	99	199	220.7	133.82	156.57
20	0	-1	1	232	221	266	239.7	23.46	26.59
21	1	-1	1	408	415	443	422.0	18.52	20.68
22	-1	0	1	182	233	182	199.0	29.44	30.13
23	0	0	1	507	515	434	485.3	44.64	47.86
24	1	0	1	846	535	640	673.7	158.21	183.74
25	-1	1	1	236	126	168	176.7	55.51	64.99
26	0	1	1	660	440	403	501.0	138.94	151.84
27	1	1	1	878	991	1161	1010.0	142.45	167.20

experimental design is cuboidal. In addition, a target for the mean of 500 is used here because Kim and Lin [5], Vining and Bohn [8], and Vining and Myers [9] fixed $T = 500$.

In this paper, we started with a full cubic model to fit the models for the mean, the S.S.D. and the D.E.. The full cubic model includes the terms $x_1, x_2, x_3, x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3, x_2x_3, x_1^3, x_2^3, x_3^3, x_1^2x_2, x_1^2x_3, x_2^2x_1, x_2^2x_3, x_3^2x_1, x_3^2x_2$ and $x_1x_2x_3$. Model selection procedures in SPSS, version 7.5.2, that is, Forward (a procedure begin with no regressors in the model and attempts to insert regressors until an optimal model is obtained), Backward (a procedure begin with all regressors in the model and attempts to eliminate regressors until an optimal model is obtained), and Stepwise (combination of Forward and Backward) will be employed here. We used the default settings in SPSS, namely, the probability of F for entry is 0.05 and for removal is 0.10. For each model selection procedure, models for the mean, the S.S.D. and the D.E. will be obtained. In order to make this study meaningful, the adjusted R squared value will be used to choose the best models. For example, SPSS give 8 models for the mean when the Stepwise procedure is used. Thus, we will choose the model with the highest adjusted R squared value among the 8 models before proceeding to the optimization.

3.1 Model Selection Procedure: Stepwise

Let $\hat{\omega}_\mu$ denotes the fitted response surface for the mean, $\hat{\omega}_{S.S.D.}$ denotes the fitted response surface for the S.S.D. and $\hat{\omega}_{D.E.}$ denotes the fitted response surface for the D.E.. Then, the best models for the mean, the S.S.D. and the D.E. are, respectively:

$$\begin{aligned} \hat{\omega}_\mu = & 314.674 + 177.011x_1 + 147x_2 + 131.467x_3 + 66.033x_1x_2 + 75.458x_1x_3 \\ & + 43.583x_2x_3 - 56.367x_3^2x_2 + 82.788x_1x_2x_3 \end{aligned} \quad (5)$$

$$\hat{\omega}_{S.S.D.} = 47.944 + 29.190x_3 + 29.565x_1x_2x_3 \quad (6)$$

$$\hat{\omega}_{D.E.} = 53.939 + 34.006x_3 + 34.414x_1x_2x_3 \quad (7)$$

Table 3 shows the R squared and adjusted R squared values for (5), (6) and (7) when the Stepwise procedure is used. Table 3 clearly indicates that the adjusted R squared values for (6) and (7) are close to each other.

Table 3 R Squared and adjusted R squared values for (5), (6) and (7)

Model	R squared	Adjusted R squared
(5)	0.966	0.952
(6)	0.373	0.321
(7)	0.381	0.329

Table 4 Comparison of the optimal settings for S.S.D. and D.E.

Method	Optimal settings	$\hat{\omega}_\mu$	Variance	MSE
$(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{S.S.D.}^2$	(1.0000, 1.0000, -0.5665)	497.8238	216.3959	221.1314
$(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{D.E.}^2$	(1.0000, 1.0000, -0.5676)	497.3982	228.2214	234.9919

Here, we used the mean squared error (MSE) criterion proposed by Lin and Tu [6] as the optimization criterion. The MSE consists of two terms: the bias and the variance. In addition, the Generalized Reduced Gradient algorithm in the “Solver” option in Microsoft Excel is used to minimize the MSE and to find the optimal settings of this example. The starting point of this example is $x' = (1, 1, 1)$. Two types of minimization had been carried out separately, that is, minimize $(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{S.S.D.}^2$ subject to $-1 \leq x_i \leq 1$ for $i = 1, 2, 3$ and $(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{D.E.}^2$ subject to $-1 \leq x_i \leq 1$ for $i = 1, 2, 3$.

Table 4 summarizes the optimization of S.S.D. and D.E. in the dual response problem. The use of S.S.D. led to the optimal setting $(x_1, x_2, x_3) = (1.0000, 1.0000, -0.5665)$ which resulted in an expected mean of 497.8238, and a variance of 216.3959 (MSE = 221.1314). With D.E., the optimal setting $(x_1, x_2, x_3) = (1.0000, 1.0000, -0.5676)$ resulted in an expected mean of 497.3982 and a variance of 228.2214 (MSE = 234.9919). These two estimators typically produced almost identical results.

The Appendix describes the Microsoft Excel spreadsheet implementation used for this example when Stepwise is the model selection procedure.

3.2 Model Selection Procedure: Backward

We repeated the whole analysis again but now with the use of the Backward procedure. The best models for the mean, the S.S.D. and the D.E. are, respectively:

$$\begin{aligned} \hat{\omega}_\mu = & 327.630 + 32.011x_1^2 - 22.389x_2^2 - 29.056x_3^2 + 66.033x_1x_2 + 75.458x_1x_3 \\ & + 43.583x_2x_3 + 177.011x_1^3 + 147x_2^3 + 131.467x_3^3 - 56.367x_3^2x_2 + 82.788x_1x_2x_3 \end{aligned} \quad (8)$$

$$\begin{aligned} \hat{\omega}_{S.S.D.} = & 36.809 + 16.778x_2^2 + 14.082x_2x_3 + 36.495x_1^3 + 30.442x_2^3 - 29.190x_3^3 \\ & - 22.677x_1^2x_2 - 37.451x_2^2x_1 + 29.565x_1x_2x_3 \end{aligned} \quad (9)$$

$$\begin{aligned} \hat{\omega}_{D.E.} = & 41.029 + 19.366x_3^2 + 15.409x_2x_3 + 42.154x_1^3 + 34.463x_2^3 + 21.270x_3^3 \\ & - 24.814x_1^2x_2 + 19.103x_1^2x_3 - 43.917x_2^2x_1 + 34.414x_1x_2x_3 \end{aligned} \quad (10)$$

Table 5 shows the R squared and adjusted R squared values for (8), (9), and (10) when the Backward procedure is used. Both (9) and (10) produce almost identical

Table 5 R squared and adjusted R squared values for (8), (9) and (10)

Model	R squared	Adjusted R squared
(8)	0.977	0.960
(9)	0.680	0.538
(10)	0.701	0.543

adjusted R squared values. Tables 3 and 5 also indicate that the Backward procedure can give a better model than the Stepwise procedure, based on the adjusted R squared values.

The "Solver" option will be used again here with the same information (including $T = 500$, $-1 \leq x_i \leq 1$ for $i = 1, 2, 3$, minimize $(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{S.S.D.}^2$ and $(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{D.E.}^2$, separately) except that equations (8), (9), and (10) would replace equations (5), (6), and (7). Table 6 summarizes the optimization of S.S.D. and D.E. in the dual response problem. The use of the S.S.D. led to the optimal setting $(x_1, x_2, x_3) = (1.0000, 1.0000, -0.7009)$ which resulted in an expected mean of 498.6252 and a variance of 125.8929 (MSE = 127.7836). With the D.E., the optimal setting $(x_1, x_2, x_3) = (1.0000, 1.0000, -0.6990)$ which resulted in an expected mean of 499.5843 and a variance of 8.5574 (MSE = 8.7299). The use of optimal settings obtained by the D.E. can obviously reduce the MSE by up to 93%, which is a great achievement. Generally, the great reduction of MSE was due to the variance. Thus, the D.E. can help practitioners in finding better optimal settings with smaller MSEs.

Table 6 Comparison of the optimal settings for S.S.D. and D.E.

Method	Optimal settings	$\hat{\omega}_\mu$	Variance	MSE
$(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{S.S.D.}^2$	(1.0000, 1.0000, -0.7009)	498.6252	125.8929	127.7836
$(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{D.E.}^2$	(1.0000, 1.0000, -0.6990)	499.5843	8.5574	8.7299

3.3 Model Selection Procedure: Forward

The use of the Forward procedure in this example produces the same models for the mean, the S.S.D., the D.E., R squared and adjusted R squared values as the Stepwise procedure. Thus, the Forward procedure as a model selection procedure will give exactly the same results (optimal settings, expected mean, variance, and MSE) as the Stepwise procedure.

4.0 CONCLUSION

Great improvements in dual response surface optimization can be achieved with the use of optimal settings obtained by using Downton's estimator in estimating the standard deviation of every design point. As indicated in the previous section, a reduction of the MSE of up to 93% can be achieved when the Backward procedure is considered with a little bias from the target value. In addition, the Downton's estimator and the sample standard deviation gave comparable results when the Stepwise and Forward procedures were used.

REFERENCES

- [1] Castillo, E. D., and D. C. Montgomery. 1993. A Non-linear Programming Solution to the Dual Response Problem. *Journal of Quality Technology*. 25(3): 199-204.
- [2] Copeland, K. A. F., and P. R. Nelson. 1996. Dual Response Optimization via Direct Function Minimization. *Journal of Quality Technology*. 28(3): 331-336.
- [3] Fan, S. K. S. 2000. A Generalized Global Optimization Algorithm for Dual Response Systems. *Journal of Quality Technology*. 32(4): 444-456.
- [4] Fathi, Y. 1991. A Nonlinear Programming Approach to the Parameter Design Problem. *European Journal of Operational Research*. 53: 371-381.
- [5] Kim, K. J., and D. K. J. Lin. 1998. Dual Response Surface Optimization: A Fuzzy Modeling Approach. *Journal of Quality Technology*. 30(1): 1-10.
- [6] Lin, D. K. J., and W. Tu. 1995. Dual Response Surface Optimization. *Journal of Quality Technology*. 27(1): 34-39.
- [7] Myers, R. H., and W. H. Carter. 1973. Response Surface Techniques for Dual Response Systems. *Technometrics*. 15(2): 301-317.
- [8] Vining, G. G., and L. L. Bohn. 1998. Response Surface for the Mean and Variance Using a Nonparametric Approach. *Journal of Quality Technology*. 30(3): 282-291.
- [9] Vining, G. G., and R. H. Myers. 1990. Combining Taguchi and Response Surface Philosophies: A Dual Response Approach. *Journal of Quality Technology*. 22(1): 38-45.
- [10] Box, G. E. P., and N. R. Draper. 1987. *Empirical Model-Building and Response Surfaces*. New York: John Wiley & Sons.
- [11] Downton, F. 1966. Linear Estimates with Polynomial Coefficients. *Biometrika*. 53:129-141.

APPENDIX

MICROSOFT EXCEL SPREADSHEET

In this paper, a full cubic model is used to obtain the best models for the mean, the S.S.D. and the D.E.. Hence, these three models have the following general form:
Expected Value

$$= b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_1^2 + b_5x_2^2 + b_6x_3^2 + b_7x_1x_2 + b_8x_1x_3 + b_9x_2x_3 + b_{10}x_1^3 + b_{11}x_2^3 + b_{12}x_3^3 + b_{13}x_1^2x_2 + b_{14}x_1^2x_3 + b_{15}x_2^2x_1 + b_{16}x_2^2x_3 + b_{17}x_3^2x_1 + b_{18}x_3^2x_2 + b_{19}x_1x_2x_3$$

Figure 1 shows the Microsoft Excel spreadsheet used to optimize the dual response surface. Using Excel's notation, the estimated parameters of models are as follows:

- (i) the mean is shown in cells B3:K3 and B8:K8,
- (ii) the S.S.D. is shown in cells B4:K4 and B9:K9, and
- (iii) the D.E. is shown in cells B5:K5 and B10:K10

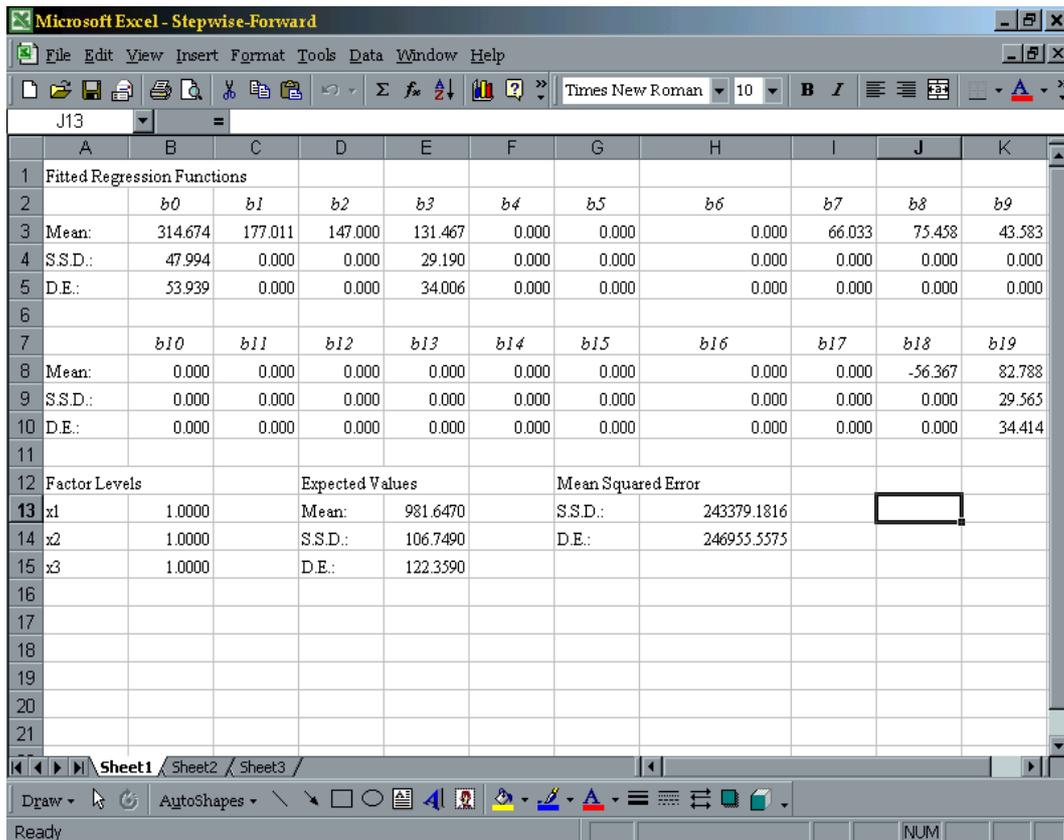


Figure 1 Microsoft Excel implementation

In addition, the values of the factor levels (x) are shown in cells B13:B15. Initially, cells B13:B15 had all ones. The expected values for the mean, the S.S.D. and the D.E. at x are found in E13, E14 and E15, respectively. Finally, our objective is to minimize the functions $(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{S.S.D.}^2$ and $(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{D.E.}^2$, separately. They are shown in cells H13 and H14, respectively.

The Generalized Reduced Gradient algorithm is invoked by selecting the “Solver” option of the Microsoft Excel. In the solver menu, we choose to minimize cell H13 by changing cells B13:B15. Constraints $-1 \leq x_i \leq 1$ for $i = 1, 2$ and 3 are included. Thus, the optimal settings, the expected mean, and the variance will be shown after we press the “solve” button. After getting the results for the objective function $(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{S.S.D.}^2$, we proceed to minimize cell H14 to get the results for the objective function $(\hat{\omega}_\mu - 500)^2 + \hat{\omega}_{D.E.}^2$.