Designing $\overline{X}$ Control Chart Using DEA Approach

Shervin Asadzadeh , Farid Khoshalhan

Abstract— **Control charts are widely implemented in firms to establish and maintain statistical control of a process which leads to the improved quality and productivity. Design of control charts requires that the engineer selects a sample size, a sampling frequency and the control limits for the chart. In this paper, a possible combination of design parameters is considered as a decision making unit which is identified by three attributes: hourly expected cost, detection power of the chart and in-control average run length. Optimal design of control charts can be formulated as multiple objective decision making (MODM). We have extended cost function from single to multiple assignable causes to near the model to the real situations. An algorithm using DEA is applied to solve the MODM model. A numerical example is used to illustrate the algorithm procedure.**

***Index Terms*— Control chart design, Data envelopment analysis, $\overline{X}$ control chart, Multiple-objective decision making (MODM)**

I. Introduction

If a product is to meet or exceed customer expectations, it should be produced by a process that is stable or repeatable. Statistical process control is a powerful collection of problem solving tools useful in achieving process stability and improving capabilities through the reduction of variability. The main tool of statistical process control is the statistical control chart. The engineering and technical implementation of control charts entails selecting sample sizes, sampling frequencies and the control limits for the chart. Selection of these three parameters is called the design of control chart. Traditionally, control charts have been designed with respect to statistical criteria only, but the design of a control chart has economic aspects too.

The first model in this case was proposed by Duncan [3]. Since that time, the economic approach has received considerable attention and various models suggested in this area. But as declared by Woodall [5], control charts based on optimal economic design, have poor statistical properties. To solve this problem, Saniga [6] noted that some of the criticism of economic design can be overcome by introducing statistical constraints in the problem and solving the model using nonlinear optimization techniques. Del Castillo, Montgomery and Mackin [7] proposed an interactive multi objective algorithm based on this procedure. Also, Chen and Liao [8] formulated optimal design of control charts as a multiple criteria decision making with respect to the constraints proposed by Saniga.

In all these articles, a single assignable cause cost function was used. However in 1971, Duncan[4] developed his previous model and presented a new model in the presence of multiple assignable causes. Since then, many tried to optimize this cost function. Chung [9] carried out subsequent work on Duncan’s model[4]. Chen and Yang [10]considered weibull in-control times with multiple assignable causes. Yu and Hou [11] optimized the control chart parameters with multiple assignable causes and variable sampling intervals. Also, Yu, Tsou and Huang [12] used Duncan's model and the proposed constraints by saniga [6] to investigate economic-statistical design of $\overline{X}$ control chart. Table 1 shows the comparison of different models, mentioned above. However, multiple objective design of $\overline{X}$ control charts with multiple assignable causes has not been addressed up to now. So, the purpose of this paper is to model design parameters in presence of multiple assignable causes. DEA method is used to find the optimum design parameters which satisfy all economic and statistical objectives. A numerical example is given to illustrate the model's working.

II. Economic cost function with multiple assignable causes

In 1971 Duncan[4] generalized his single assignable cause model to multiple one. In this model, there is an in-control state μ, an assignable cause of magnitude δ_j (j=1, 2...,s) which occurs at random, results in a shift in the mean to either $\mu + \delta_j\sigma$ or $\mu - \delta_j\sigma$ and so changes the state until the cause is detected. Meanwhile, during the search for the assignable cause, the process is allowed to continue in operation. The cycle consists of four periods:

1) In-control period

It is assumed that assignable causes occur according to Poisson process with λ_j occurrences per hour. So assuming that process begins in the in-control state, the time interval that the process remains in control is an exponential random variable with mean $1/\lambda$ hour:

$$\frac{1}{\lambda} = \frac{1}{\sum_{j=1}^{s} \lambda j} \qquad (1)$$

2) Out of control period

When the process goes to out of control state, the probability

Manuscript submitted December 1,2007

Corresponding Author :Sh.Asadzadeh, MS student, Department of Industrial Engineering, K.N.Toosi University of Technology,Tehran,Iran. phone: +982188674999; fax: +982188674858
e_mail: sh_asadzadeh@sina.kntu.ac.ir

Farid Khoshalhan, Assistant Professor, Department of Industrial Engineering, K.N.Toosi University of Technology,Tehran,Iran. phone: 0982188674999; fax: +982188674858
e_mail: khoshalhan@kntu.ac.ir

that it will be detected on any subsequent sample is related to the assignable cause occurred. If the jth assignable cause

Table1.comparison of the models

Model/year	Assumptions	Output
Del Castillo, Montgomery and Mackin(1996)	Single assignable cause Exponential in-control times	Multi-objective design (economic-statistical)
Chen and Liao (2004)	Single assignable cause Exponential in-control times	Multi-objective design (economic-statistical)
Chung (1994)	Multiple assignable causes Exponential in-control times	Economic design
Chen and Yang (2002)	Multiple assignable causes Weibull in-control times	Economic design
Yu and Hou (2006)	Multiple assignable causes variable sampling intervals	Economic design
Yu, Tsou and Huang (2007)	Multiple assignable causes Exponential in-control times	Economic-statistical design (economic objective with statistical constraints)

happens, then the detection power will be:

$$P_j = \int_{-\infty}^{-k-\delta j\sqrt{n}} \Phi(z)dz + \int_{k-\delta j\sqrt{n}}^{+\infty} \Phi(z)dz \quad (2)$$

Where $\phi(z)$, is the probability density function of standardized normal distribution. So the average samples taken after the jth assignable cause happens, is $1/p_j$.

Also, given the occurrence of the jth assignable cause between the uth and u+1st sample, the expected time of occurrence within this interval is:

$$\tau_j = \frac{\int_u^{(u+1)h} \lambda_j e^{-\lambda_j t}(t-uh)dt}{\int_u^{(u+1)h} \lambda_j e^{-\lambda_j t} dt} = \frac{1-(1+\lambda_j h)e^{-\lambda_j t}}{\lambda_j(-e^{-\lambda_j t})} \quad (3)$$

Where h is the sampling frequency.
Therefore, the time required to observe an out of control alarm when the jth assignable cause occurs, will be:

$$\frac{h}{P_j} - \tau_j \quad (4)$$

3) The time to take sample and interpret the results is a constant g proportional to the sample size n, so that gn is the length of this part of the cycle.

4) The time required to find the assignable cause. If this time is Dj for the jth assignable cause, then the expected time in a cycle for detecting assignable cause is

$$\frac{\sum_{j=1}^{s} \lambda_j D_j}{\lambda} \quad (5)$$

Therefore, the expected length of a cycle is:

$$E_{CT} = \frac{1}{\lambda} + \frac{\sum_{j=1}^{S} \lambda_j (\frac{h}{P_j} - \tau_j)}{\lambda} + gn + \frac{\sum_{j=1}^{s} \lambda_j D_j}{\lambda} = \frac{1 + \sum_{j=1}^{s} \lambda_j (\frac{h}{P_j} - \tau_j + gn + D_j)}{\lambda} \quad (6)$$

If the fixed component of sampling cost is a_1 and the variable one is a_2, then the cost of taking a sample of size n will be $a_1 + a_2 n$. The cost of finding an assignable cause j, is a_{3j} and the cost of investigating a false alarm is a_4. The expected number of false alarm generated during a cycle is α, times the expected number of samples taken before the shift or:

$$\frac{\alpha e^{-\lambda h}}{1 - e^{-\lambda h}} \quad (7)$$

In which α is calculated through the below equation:

$$\alpha = \frac{1}{ARL} = 2\int_k^{\infty} \Phi(z)dz \quad (8)$$

Now if one defines a_{5j} , as the hourly penalty cost associated with production in out of control state, then the expected cost per cycle will be

$$E_{CC} = \frac{\sum_{j=1}^{S} a_{5j}\lambda_j (\frac{h}{P_j} - \tau_j + gn + D_j) + \sum_{j=1}^{S} a_{3j}\lambda_j}{\lambda} + \frac{a_4 \alpha e^{-\lambda h}}{1 - e^{-\lambda h}} + (a_1 + a_2 n)\frac{E_{CT}}{h} \quad (9)$$

And the expected cost per hour can be indicated as:

$$E_{HC} = \frac{E_{CC}}{E_{CT}} = \frac{(a_1 + a_2 n)}{h} + \left(\frac{\sum_{j=1}^{S} a_{5j}\lambda_j (\frac{h}{P_j} - \tau_j + gn + D_j) + \sum_{j=1}^{S} a_{3j}\lambda_j}{\lambda} + \frac{a_4 \alpha e^{-\lambda h}}{1 - e^{-\lambda h}} \right) / \left(\frac{1 + \sum_{j=1}^{S} \lambda_j (\frac{h}{P_j} - \tau_j + gn + D_j)}{\lambda} \right) \quad (10)$$

So economic design of $\overline{X}$ control chart involves determination of optimal parameters n, h and k which minimize E_{HC}.

III. Multiple objective optimal design of $\overline{X}$ control chart

To establish the multiple objective decision making model, we should first determine a set of conflicting objectives that define the problem for the quality control manager. Due to the nature of the DEA method used in algorithm, various combinations of design parameters n, h, k should also be set in advance. Taking into account the Saniga's constraints, the multi-objective model is:

Max ARL(D)
Max P(D)
Max E_{HC}(D)
s.t (11)
$P_j \geq P_{lj}$ (j=1, 2...,s)
$\alpha \leq \alpha_u$
$ATS_j \leq ATS_{uj}$ (j=1, 2...,s)

D is a possible combination of design parameters that has been shown in bracket for the entire three objectives for emphasizing on the fact that it does have an impact on the values of objectives.
The aims of MODM models are to find solutions that can satisfy and set a balance among all objectives.. To solve MODM problems, the DEA method is one of the most powerful and popular method to optimize the feasible combinations of design parameters specifically when measuring the efficiencies of similar units is under consideration.

IV. Data Envelopment Analysis (DEA)

DEA is the optimization method of linear programming to generalize the Farrell [1] single input, single output technical efficiency measure to the multiple-input, multiple-output case by constructing a relative efficiency score of a group of competing decision making units (DMU). Applications and implementations of DEA in modeling performance measurement [13] has gained a lot of attention in recent years. In this paper we have used the CCR model(Charnes, Cooper and Rhodes[2]). The objective in CRR model is to maximize the relative efficiency value of each of DMUs from among a reference set of design D, by selecting the optimal weights associated with the inputs and outputs. The algebraic model is as follows:

$$Max \quad E_i(D) = \frac{\sum_{r=1}^{z} U_r Y_{ri}(D)}{\sum_{j=1}^{m} V_j X_{ji}(D)}$$

s.t. (12)

$$\frac{\sum_{r=1}^{z} U_r Y_{ri}(D)}{\sum_{j=1}^{m} V_j X_{ji}(D)} \leq 1 \qquad \textit{for other designs } D$$

Where
U_r : the weights given to output r
Y_{ri} : amount of output r from unit i
V_j : weight given to input j
X_{ji} : amount of input j from unit i
To solve the model, it is necessary to convert it into linear form so that methods of linear programming can be applied. This nonlinear programming is equivalent to two linear programming: 1) setting its denominator to one and maximizing its numerator (output maximization) 2) setting its numerator to one and minimizing denominator (input minimization). Because CCR model considers constant retunes to scale, there exists no difference which one to choose and CCR yields the same efficiency score. So,the linear programming will be:

$$Max \quad E_i(D) = \sum_{r=1}^{z} U_r Y_{ri}(D)$$

s.t. (13)

$$\sum_{j=1}^{m} V_j X_{ji}(D) = 1$$

$$\sum_{r=1}^{z} U_r Y_{ri}(D) - \sum_{j=1}^{m} V_j X_{ji}(D) \leq 1$$

$$\textit{for other designs } D$$

$$U_r, V_j > 0$$

If E_i^* =1, that means no other design is more efficient than design i under its own weights. If E_i^*<1, then there is at least one other design that is more efficient under optimal set of weights determined. Calculation should be done for each DMU to find the relative efficiency of each one.

V. Solution algorithm

Unlike many multiple-objective models that the DM has an implicit unknown value function, here the values of E_{HC}(D), P(D) and ARL (D) must be calculated for each potential combination D according to formula 1 to 10 in advance. Due to the complicated multi-assignable cause cost function, all calculations have been facilitated by Excel software. Also, to evaluate and compare the efficiencies of DMUs, Microsoft Excel with XlDEA has been implemented. Chen and Liao [8] proposed a solution procedure for their multi-criteria decision making model. In this paper, we have employed their 4-step algorithm to solve our multi-objective model. They applied this procedure for their model with one assignable cause cost function. The procedure is approximately the same except steps 1 and 2 which have been converted a bit to suit our proposed model.
The four-step procedure will be:
1) Determining all possible solutions by putting bounds on each parameter. In this paper the scope of sample size n is set from 1 to 35, increased by 1. Scope of sampling frequency is confined from 0.1 to 4 increased by 0.1h and finally the scope of control limit width k is considered from 0.1 to 3 in terms of standard deviation increased by 0.1. Contemplating all possible combinations the number of potential solutions will be 35*40*30=42000

2) In this step we have added another constraint too, in order to take into account the value of parameter h, because in the two previous constraints Chen and Liao used, there was no sign of parameter h. subsequently, we eliminate infeasible solutions by the following constraints:
$\alpha \leq \alpha_u$ $\quad P_j \geq P_{lj}$ $\quad ATS_j \leq ATS_{uj}$

3) Partial optimization. Remain the elements with Pareto optimality for each subset Q_n . A solution "s" with Pareto optimization in a set Q_n means that there is no other solution

in the same set such that "s" is dominated in terms of statistical properties and cost.

4) Global optimization. Merge all the remainders into a set W and select the elements with highest relative efficiency among W. The selected elements will afford to DM to make final decision.

VI. Numerical example

In this section, Duncan's [4] data were employed to illustrate the use of the proposed model and algorithm.
The numbers of assignable causes are assumed to be 12. When an assignable cause j with the average occurrence of λ_j occurs, it produces a shift of size δ_j in the mean. The cost of taking a sample that is independent of sampling is 1$ and the variable cost per item of sampling, testing and plotting is 0.1$. An average time of 0.05h is needed to test and analyze a sample item and the cost of looking for trouble when none exists, is estimated 25$. Values of other parameters have been tabulated in table 2.

Table2.input values of parameters

δ_j	λj	Dj	a_{3j}	a_{5j}
0.75	0.001098	4.17	19.68	7.22
1.25	0.000855	3.08	14.57	27.6
1.75	0.000666	2.50	11.81	76.14
2.25	0.000519	2.08	9.84	165.69
2.75	0.000404	1.92	9.06	302.36
3.25	0.000314	1.84	8.66	433.64
3.75	0.000245	1.77	8.37	570.32
4.25	0.000191	1.72	8.17	659.86
4.75	0.000148	1.70	8.05	708.4
5.25	0.000115	1.68	7.93	728.97
5.75	0.000090	1.66	7.83	735.78
6.25	0.000070	1.64	7.73	737.56

Also our statistical constraints in false alarm rate α , detection power P_j and average time to signal ATS_j are

$$\alpha \leq 0.1 \qquad P_j \geq 0.9 \qquad ATS_j \leq 4$$

The optimization procedure can be carried out as described.Table3 illustrates the results.

As indicated by *, two design parameters combinations have received score 1 and therefore offered to the DM for final selection. Then the DM may choose the first combination if low cost is of paramount importance for him/her. Similarly if he/she is much more interested in the outgoing quality, then the second combination with large average run length and detection power may be the final choice.

VII. Conclusion

A multi-objective model for designing $\overline{X}$ control chart in presence of multiple assignable causes, is proposed. For this model, various combinations of n, h, k are contemplated as DMUs. DEA method is employed to assess the efficiency of DMUs and to select the optimum designs with large average run length, high detection power and low expected cost. Numerical example is given based on the Duncan's [4] data to illustrate the solution procedures .Other interesting research areas for future research involve multi-objective design of $\overline{X}$ control chart under weibull shock and multi-objective design of adaptive $\overline{X}$ control chart.

Table3.Non-dominated solutions with largest efficiencies

(n,h,k)	Cost	P	ARL	
(27,2.9,2.6)	5.9662	0.9773	107.5269	*
(28,2.9,2.6)	6.0481	0.9801	107.5269	
(29,2.9,2.6)	6.1301	0.9825	107.5269	
(29,2.9,2.7)	6.1103	0.9789	144.9275	
(30,3,2.6)	6.2109	0.9846	107.5269	
(30,3,2.7)	6.1918	0.9814	144.9275	
(30,2.9,2.8)	6.1771	0.9777	196.0784	
(31,3,2.6)	6.2917	0.9866	107.5269	
(31,3,2.7)	6.2725	0.9837	144.9275	
(31,3,2.8)	6.2601	0.9803	196.0784	
(32,3.1,2.6)	6.3723	0.9883	107.5269	
(32,3,2.7)	6.3533	0.9856	144.9275	
(32,3,2.8)	6.3387	0.9826	196.0784	
(32,3,2.9)	6.3291	0.9791	270.2703	
(33,3.1,2.6)	6.4519	0.9898	107.5269	
(33,3.1,2.7)	6.4334	0.9874	144.9275	
(33,3.1,2.8)	6.4193	0.9846	196.0784	
(33,3,2.9)	6.4084	0.9814	270.2703	
(33,3,3)	6.4003	0.9777	370.3704	*
(34,3.1,2.6)	6.5316	0.9911	107.5269	
(34,3.1,2.7)	6.5131	0.989	144.9275	
(34,3.1,2.8)	6.4989	0.9865	196.0784	
(34,3.1,2.9)	6.4883	0.9836	270.2703	
(34,3.1,3)	6.4804	0.9802	370.3704	
(35,3.2,2.6)	6.6109	0.9922	107.5269	
(35,3.1,2.7)	6.5926	0.9904	144.9275	
(35,3.1,2.8)	6.5785	0.9881	196.0784	
(35,3.1,2.9)	6.5679	0.9855	270.2703	
(35,3.1,3)	6.5599	0.9824	370.3704	

References

[1] Farrell M.J (1957) The measurement of productive efficiency. J.R. Statis. Soc. Series A 120(3): 253-290
[2] Chanes A, Cooper W.W, Rhodes E (1978) Measuring the efficiency of decision making units. European Journal of Operational Research 2(6): 429-444
[3] Duncan A.J (1956) The economic design of x charts used to maintain current control of a process.Journal of the American Statistical Association 51(274): 228-242
[4] Duncan A.J (1971) The economic design of x charts when there is a multiplicity of assignable causes Journal of the American Statistical Association 66(333): 107-121
[5] Woodall W.H (1986) Weakness of the economic design of control charts. Technometrics 28(4): 408–410
[6] Saniga E. M (1989) Economic statistical control chart designs with an application to $\overline{X}$ and R charts. Technometrics 31(3): 313-320

[7] Del Castillo E, Mackin P , Montgomery D. C (1996) Multiple-criteria optimal design of $\overline{X}$ control charts. IIE Transactions 28(6): 467-474

[8] Chen Y.K, Liao H.C (2004) Multi-criteria design of an $\overline{X}$ control chart. Computers and Industrial Engineering 46(4): 877-891

[9] Chung k.j (1994) An algorithm for computing the economically optimal $\overline{X}$ control charts for a process with multiple assignable causes. European Journal of Operation Research 72(2): 350-363

[10] ChenY.S, Yang Y.M (2002) Economic design of $\overline{X}$ control charts with weibull in-control times when there are multiple assignable causes. International Journal of Production Economics 77(1): 17-23

[11] Yu F.J, Hou J.L (2006) Optimization of design parameters for control charts with multiple assignable cause. Journal of Applied Statistics 33(3): 279 – 290

[12] Yu F.J, Tsou C.S , Huang K.I (2007) An Economic-Statistical Design of $\overline{X}$ Control Charts with Multiple Assignable Causes. The 22 nd European Conference on Operational Research.

[13] Cook W.D , Zhu J (2005) Modeling Performance Measurement, Springer