
Optimising Shewhart charts in parallel production lines

Ronald D. Fricker, Jr.

Operations Research Department,
Naval Postgraduate School,
Monterey, California 93943, USA
E-mail: rdfricke@nps.edu

Abstract: I describe a methodology for optimising n Shewhart \bar{x} -charts operating on parallel production lines in a factory. The goal is to maximise the factory-wide probability of detecting an out-of-control condition subject to a constraint on the expected number of false signals. I use non-linear programming to appropriately set the \bar{x} -charts' control limits incorporating information about the probability of each production line going out-of-control. Using this approach, factories can set their quality control systems to optimally detect out-of-control conditions. Given some distributional assumptions, I also present a one-dimensional optimisation methodology that allows for the efficient optimisation of very large factories.

Keywords: industrial quality control; statistical process control; x -bar chart; quality engineering; quality technology.

Reference to this paper should be made as follows: Fricker, R.D., Jr. (2009) 'Optimising Shewhart charts in parallel production lines', *Int. J. Quality Engineering and Technology*, Vol. 1, No. 2, pp.125–135.

Biographical notes: Ronald D. Fricker, Jr. is an Associate Professor in the Operations Research Department of the Naval Postgraduate School (NPS). Prior to joining NPS, he was a Senior Statistician at the RAND Corporation. He holds a PhD and an MS in Statistics from Yale University, an MS in Operations Research from The George Washington University and a Bachelor's degree from the United States Naval Academy. His current research is focused on studying the performance of various statistical methods for use in biosurveillance and statistical process control methodologies more generally. He is a Fellow of the American Statistical Association and a former Chair of the Section on Statistics in Defense and National Security and the Committee on Statisticians in Defense and National Security, both of the American Statistical Association.

1 Introduction

Consider a factory with n production lines, each being monitored for quality by a single Shewhart \bar{x} -charts. In such installations chart control limits are usually set equally for all the production lines, often using 3σ limits. Choosing control limits entails making a trade-off between the frequency of adjudicating false positive signals and the speed of detecting an out-of-control condition. The former is usually quantified in terms of the

in-control average run length (ARL_0) and the latter in terms of the out-of-control average run length (ARL_1).

The inherent assumption behind setting the control limits the same on the production lines is that they all have an equal probability of an out-of-control condition occurring. However, such an assumption may not be true, perhaps because of variations in equipment or personnel that are uncorrectable by factory management. In this situation, setting equal control limits could be sub-optimal in the sense that ideally one would want to set the control limits to be more sensitive to catching the line with a higher probability of going out of control.

The methodology presented in this paper provides such a means for optimising control limits under these conditions. It requires a change in the way one thinks about the design of control charts. First, the optimisation is done at the factory level where I assume that there is some fixed level of effort that management desires to devote to adjudicating control chart signals. Second, the problem is not set up in terms of in-control and out-of-control average run lengths, though it is defined in terms that are closely and directly related to the average run lengths.

The paper is organised as follows. In Section 2, I formulate the problem of optimising n Shewhart \bar{x} -charts operating on parallel production lines in a factory and illustrate it on a hypothetical ten production line factory. In Section 3, I derive an equivalent one-dimensional optimisation problem that allows for the efficient application to very large factories. In Section 4, I examine what happens when the optimised set of control charts is applied in situations deviating from the optimisation assumptions, and finally in Section 5 I summarise the results, including providing pointers to other fields to which these results could be applied.

2 Problem formulation

Consider a factory consisting of n independent production lines, each of which is monitored by a Shewhart chart. Let \bar{X}_{ij} denote the statistic to be plotted on the Shewhart chart for production line i at time j , $i = 1, \dots, n, j = 1, 2, \dots$

When the process is in-control, assume:

- the \bar{X}_{ij} are independent and identically distributed
- the rational subgroup size is sufficiently large so it is reasonable to assume the statistics are normally distributed
- the process variance for each of the production lines is known, $\sigma_i, i = 1, \dots, n$.

Then, without loss of generality, we can assume that when the factory is in-control $\bar{X}_{ij} \sim F_0 = N(0,1)$ for all i and all j while, if production line i goes out of control at time τ , $\bar{X}_{ij} \sim F_1 = N(\delta,1), \delta \neq 0, j = \tau, \tau + 1, \dots$

Average run length is the standard measure of control chart performance. For production line i , the goal is to set the control limit h_i such that when the line is in-control ARL_0 is suitably large and when it goes out-of-control ARL_1 is suitably small.

For production line i at time j , the probability a two-sided Shewhart control chart gives a false signal is

$$1 - \int_{x=-h_i}^{h_i} f_0(x) dx = 2 \times \Phi(-h_i) = \alpha_i, \quad (1)$$

and the probability it fails to signal during an out-of-control condition is

$$\int_{x=-h_i}^{h_i} f_1(x) dx = 2 - \Phi(-h_i + \delta) - \Phi(h_i + \delta) = \beta_i. \quad (2)$$

Thus, for production line i , $ARL_0 = 1/\alpha_i$ and $ARL_1 = 1/(1 - \beta_i)$.

While average run lengths are useful metrics for setting control limits for an individual production line, from a factory perspective one might prefer metrics that quantify the combined performance of all the charts (particularly if each of the production lines can set a different control limit). One such metric is the average time between false signals for all the control charts in the factory, or the *combined in-control ARL* (C- ARL_0), calculated as

$$C-ARL_0 = \left(\sum_{i=1}^n \alpha_i \right)^{-1}.$$

Defining p_i as the proportion of times that production line i goes out of control out of the total number of times any production line in the factory goes out of control, we have

$\sum_{i=1}^n p_i = 1$. And, we can think of the p_i s as probabilities in the sense that, at some random point in time, p_i is the probability that line i will next go out of control. In a Bayesian framework, we can also think about $\mathbf{p} = \{p_1, p_2, \dots, p_n\}$ as a sort of prior distribution. Then, given that an out-of-control condition occurs in some future time period according to \mathbf{p} , a second metric is the probability the out-of-control condition is detected in that time period: $P_d(\mathbf{h}) = \sum_{i=1}^n p_i (1 - \beta_i)$.

Given these factory-level metrics, we formulate the problem of choosing control limits as maximising the probability of detecting an out-of-control condition occurring on one of the production lines according to \mathbf{p} , subject to a minimum constraint on C- ARL_0 . That is, defining $\mathbf{h} = \{h_1, \dots, h_n\}$:

$$\begin{aligned} \max_{\mathbf{h}} \quad & P_d(\mathbf{h}) \\ \text{s.t.} \quad & C-ARL_0 \geq \kappa'. \end{aligned} \quad (3)$$

This factory level optimisation is akin to choosing a control limit that minimises ARL_1 subject to a lower bound on ARL_0 at the production line (i.e., control chart) level.

Restating the constraint in terms of the expected number of false signals in a particular time period, $\sum_{i=1}^n \alpha_i$, which is a measure of the cost of operating the factory for one time period when everything is in-control, an equivalent form is:

$$\begin{aligned} \max_{\mathbf{h}} \quad & \sum_{i=1}^n p_i (1 - \beta_i) \\ \text{s.t.} \quad & \sum_{i=1}^n \alpha_i \leq \kappa. \end{aligned} \quad (4)$$

Finally, to be explicit about the assumption of normality, we can also express the problem as:

$$\begin{aligned} \max_{\mathbf{h}} \quad & \sum_{i=1}^n p_i [2 - \Phi(h_i + \delta) - \Phi(h_i - \delta)] \\ \text{s.t.} \quad & 2 \sum_{i=1}^n \Phi(-h_i) \leq \kappa. \end{aligned} \quad (5)$$

Note that in this formulation of the problem we are maximising the probability of detecting a single out-of-control condition that occurs somewhere in the factory. This is a conservative detection probability, in the sense that if multiple out-of-control conditions occur simultaneously then the actual probability of detection will be greater than $P_d(\mathbf{h})$.

2.1 An illustrative example

Consider a hypothetical factory that consists of ten production lines, each of which has a probability of going out-of-control (p_i) as depicted in Table 1. In this factory, production line #1 is more likely to go out of control than the others lines. In fact p_1 is an order of magnitude greater than the other production lines, perhaps due to older equipment or inexperienced operators.

Table 1 An illustrative factory with ten production lines

Production line (i)	p_i	Common control limit #1	Optimal control limits	Common control limit #2
1	0.55	3.00	2.28	2.69
2	0.05	3.00	3.48	2.69
3	0.05	3.00	3.48	2.69
4	0.05	3.00	3.48	2.69
5	0.05	3.00	3.48	2.69
6	0.05	3.00	3.48	2.69
7	0.05	3.00	3.48	2.69
8	0.05	3.00	3.48	2.69
9	0.05	3.00	3.48	2.69
10	0.05	3.00	3.48	2.69
	P_d	0.159	0.245	0.245
	C-ARL ₀	37.0	37.0	14.0

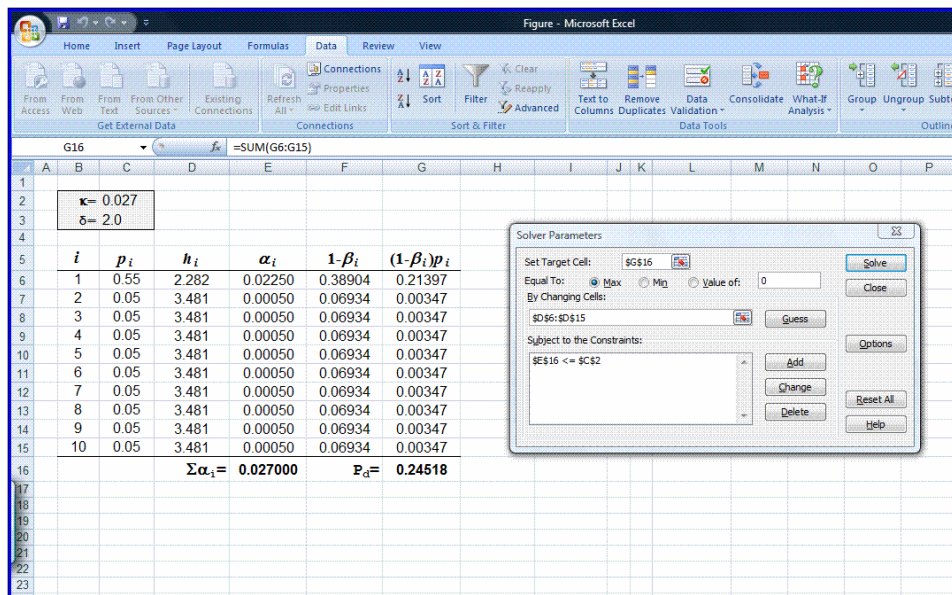
Assume that $F_0 = N(0, 1)$ and $F_1 = N(2, 1)$. The column labelled ‘common control limit #1’ shows that the factory would achieve a probability of detection of $P_d = 0.159$ with a combined in-control ARL of 37.0 using 3σ control limits for all production lines. However, by optimising the control limits, the ‘optimal control limit’ column shows that a probability of detection of $P_d = 0.245$ can be achieved for the same combined in-control ARL – a more than 50% improvement. This is achieved by lowering the control limit (i.e., increasing the probability of detecting an out-of-control condition) in the production line most likely to go out-of-control while raising the control limits in those locations less

likely to go out-of-control. Finally, the column labelled ‘common control limit #2’ shows that to achieve the same optimal $P_d = 0.245$ with a common control limit (2.69) the factory would have a combined in-control ARL of 14.0 – which more than doubles the false signal rate for the factory. This means that using a common control limit costs the factory 62% more effort in terms of personnel time spent investigating false positive control chart signals to achieve the same probability of detecting an out-of-control condition.

2.2 Optimising control limits

For a small factory, with F_0 and F_1 normal distribution functions, it is a simple matter to optimise (5) in an excel spreadsheet using the NORMDIST function and the Solver. See Figure 1. For this example, I used the Solver in Excel 2007 to find the optimal control limits, which ran quickly (a fraction of a second) and reliably found the optimal solution. (Within the Solver, I used the Newton search method with Precision= 1×10^{-6} , Tolerance= 5×10^{-2} , and Convergence= 1×10^{-4} – the default settings.) However, note that the Solver is limited to 200 adjustable cells (<http://support.microsoft.com/kb/75714>), which puts an upper bound on the number of control charts that can be optimised using this approach.

Figure 1 Screen shot of Excel using the Solver to get the optimal control limits for the Section 2.1 example (see online version for colours)



The fundamental problem is that every additional production line adds a variable to (5). As the dimensionality of the problem grows, more specialised optimisation software such as the MINOS solver in GAMS may suffice, though very large factories may exceed the capacity of even these programs to solve via brute force. This suggests a need for an alternative solution methodology that reduces the dimensionality of the problem.

3 An equivalent one-dimensional optimisation problem

Even though it is easy to show that under some relatively mild conditions the objective function in (5) is strongly quasiconvex over the constraint region, because this is a maximisation problem a globally-optimal solution is not guaranteed. However, assuming F_0 and F_1 are normally distributed and the out-of-control condition manifests itself as a shift in the mean, with the following theorem we can simplify this from an n -variable optimisation problem to a one-variable optimisation problem with a guaranteed optimal solution.

Theorem 1: If $F_0 = N(0, 1)$ and $F_1 = N(\delta, 1)$, $\delta \neq 0$, then the optimisation problem reduces to finding μ to satisfy

$$\sum_{i=1}^n \Phi \left(\mu - \frac{1}{|\delta|} \ln(p_i) \right) = n - \kappa/2, \quad (6)$$

and the optimal solution is $h_i = \mu - \frac{1}{|\delta|} \ln(p_i)$.

Proof: It is easy to show that the optimal solution lies on the boundary of the constraint, so from (5) we can express the upper control limit for production line #1 as

$$h_1 = \Phi^{-1} \left(n - \kappa/2 - \sum_{i=2}^n \Phi(h_i) \right).$$

The result then follows from reformulating the constrained maximisation problem as an unconstrained problem:

$$\begin{aligned} \max_{\mathbf{h}} f = & p_1 \left[2 - \Phi \left(\Phi^{-1} \left(n - \kappa/2 - \sum_{i=2}^n \Phi(h_i) \right) + \delta \right) \right. \\ & \left. - \Phi \left(\Phi^{-1} \left(n - \kappa/2 - \sum_{i=2}^n \Phi(h_i) \right) - \delta \right) \right] \\ & + \sum_{i=2}^n p_i \left[2 - \Phi(h_i + \delta) - \Phi(h_i - \delta) \right]. \end{aligned}$$

The partial differential equations with respect to each of the h_i , for $i = 2, 3, \dots, n$, are

$$\begin{aligned} \frac{\partial f}{\partial h_i} = & -\frac{1}{\sqrt{2\pi}} \exp \left(\frac{-h_i^2 - \delta^2}{2} \right) \left[p_i \exp[h_i \delta] + p_i \exp[-h_i \delta] \right] \\ & + p_i \exp \left(\sqrt{2} \delta \operatorname{Erf}^{-1} \left\{ n - \kappa - \sum_{i=2}^n \operatorname{Erf} \left[\frac{h_i}{\sqrt{2}} \right] \right\} \right) \\ & + p_i \exp \left(-\sqrt{2} \delta \operatorname{Erf}^{-1} \left\{ n - \kappa - \sum_{i=2}^n \operatorname{Erf} \left[\frac{h_i}{\sqrt{2}} \right] \right\} \right), \end{aligned} \quad (7)$$

where $\text{Erf}(z/\sqrt{2}) = \frac{2}{\sqrt{\pi}} \int_0^z \exp(-t^2) dt$ and $\text{Erf}^{-1}(\text{Erf}(z)) = z$.

Now, (7) is equal to zero if

$$p_i \exp[h_i \delta] = p_1 \exp\left(\sqrt{2} \delta \text{Erf}^{-1}\left\{n - \kappa - \sum_{i=2}^n \text{Erf}\left[\frac{h_i}{\sqrt{2}}\right]\right\}\right)$$

and

$$p_i \exp[-h_i \delta] = p_1 \exp\left(-\sqrt{2} \delta \text{Erf}^{-1}\left\{n - \kappa - \sum_{i=2}^n \text{Erf}\left[\frac{h_i}{\sqrt{2}}\right]\right\}\right).$$

Simplifying gives

$$\text{Erf}\left[\frac{h_i + (\ln(p_i) - \ln(p_1))/\delta}{\sqrt{2}}\right] = n - \kappa - \sum_{i=2}^n \text{Erf}\left[\frac{h_i}{\sqrt{2}}\right]$$

and

$$\text{Erf}\left[\frac{h_i - (\ln(p_i) - \ln(p_1))/\delta}{\sqrt{2}}\right] = n - \kappa - \sum_{i=2}^n \text{Erf}\left[\frac{h_i}{\sqrt{2}}\right].$$

Since $\text{Erf}(z/\sqrt{2}) = 2\Phi - 1$, after some algebra we have that

$$\Phi\left(h_i + \frac{1}{\delta} \ln(p_i) - \frac{1}{\delta} \ln(p_1)\right) + \sum_{i=2}^n \Phi(h_i) = n - \kappa/2$$

and

$$\Phi\left(h_i - \frac{1}{\delta} \ln(p_i) + \frac{1}{\delta} \ln(p_1)\right) + \sum_{i=2}^n \Phi(h_i) = n - \kappa/2.$$

The result in Theorem 1 follows by setting $h_i = \mu - \frac{1}{|\delta|} \ln(p_i)$.

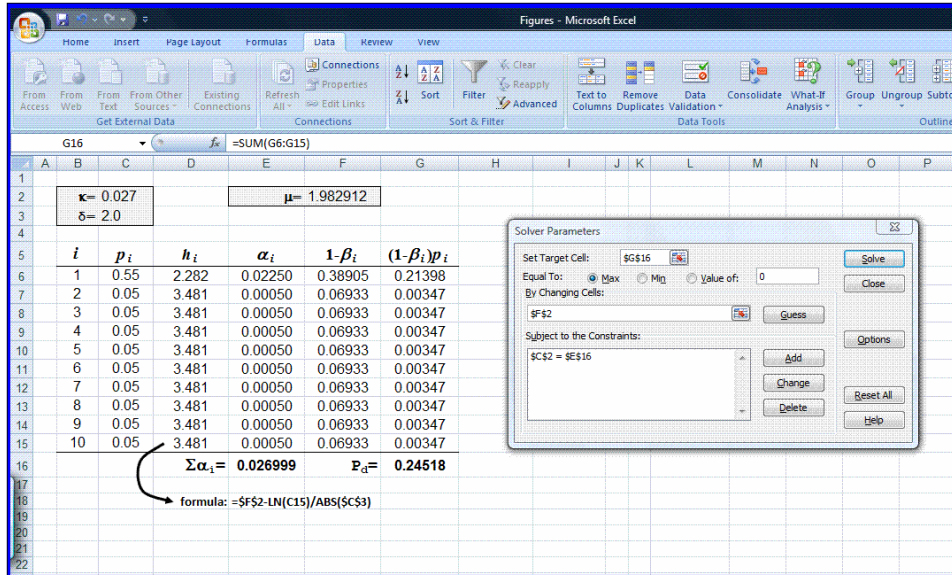
Figure 2 demonstrates that applying Theorem 1 to the hypothetical ten production line example gives the same result as was shown in Figure 1.

One way to think about the one-dimensional optimisation in Theorem 1 is in terms of finding μ such that the sum of the probabilities that each of n normally distributed random variables (all with the same mean but possibly different variances) is greater than some constant equals $n - \kappa/2$. Specifically, find μ such that

$$\sum_{i=1}^n \mathbb{P}\left(X_i > \frac{1}{|\delta|}\right) = n - \kappa/2, \quad (8)$$

where $X_i \sim N\left(\mu, [\ln(p_i)]^2\right)$.

Figure 2 Applying the results of Theorem 1, a screen shot of Excel using the Solver to get the optimal control limits for the Section 2.1 example (see online version for colours)



Given the continuity of the normal distribution, (8) makes it clear that an optimal solution is guaranteed to exist. Furthermore, it is a relatively simple problem to solve for μ by starting with a large value and gradually decreasing it until the sum of one minus each cdf evaluated at $1/|\delta|$ in (8) equals $n - \kappa/2$.

If the application calls for one-sided Shewhart charts, the following theorem applies.

Theorem 2: If $F_0 = N(0, 1)$ and $F_1 = N(\delta, 1)$, $\delta > 0$ then the optimisation problem reduces to finding μ to satisfy $\sum_{i=1}^n \Phi\left(\mu - \frac{1}{\delta} \ln(p_i)\right) = n - \kappa$, and the optimal solution is

$$h_i = \mu - \frac{1}{\delta} \ln(p_i).$$

The proof follows the same steps as Theorem 1.

4 Discussion

In the hypothetical ten production line example in Section 2.1, the control limits were set assuming $C\text{-ARL}_0 = 37$ and $\delta = 2$. Setting $C\text{-ARL}_0$ is a matter of resources and should be based on an organisational assessment of the resources to be devoted to investigating false positive signals. In the example, we set $C\text{-ARL}_0 = 37$ simply to be consistent with what would occur with ten Shewhart charts each using 3σ limits. Of course, for a fixed number of control charts, one can improve the factory-wide probability of detection by

increasing the expected number of false signals allowed. Table 2 shows the trade-off in probability of detection for the ten production line example for four levels of δ and for five values of C-ARL₀.

Table 2 Optimal probabilities of detection in the hypothetical ten production line example for various values of δ and C-ARL₀

P_d	C-ARL ₀ = 10	C-ARL ₀ = 20	C-ARL ₀ = 30	C-ARL ₀ = 37	C-ARL ₀ = 50
$\delta = 1$	0.145	0.094	0.072	0.062	0.051
$\delta = 2$	0.395	0.310	0.266	0.245	0.217
$\delta = 3$	0.733	0.654	0.606	0.581	0.546
$\delta = 4$	0.941	0.910	0.887	0.874	0.855

Choosing the value of δ over which to optimise is a subjective judgement based on the minimum increase that the factory wishes to detect. As shown in Table 2, once the choice is made and the control limits set, an out-of-control condition manifested as a small value for δ will be harder to detect and will result in a lower probability of detection. Conversely, an out-of-control condition manifested as a larger δ will make it easier to distinguish between F_0 and F_1 and thus will result in a higher the probability of detection.

That said, a relevant question is how sensitive the resulting probability of detection is to the misspecification of δ during the optimisation. For example, what happens if the control limits are chosen using an optimisation based on $\delta = 2$ and then the actual outbreak manifests itself with $\delta = 1$ or $\delta = 3$? Table 3 shows the actual probabilities of detection that would occur in the ten production lines example using the optimal control limits determined for $\delta = 2$. Comparing Table 3 to Table 2 we see that there is some degradation in P_d if the actual out-of-control condition manifests at some δ other than the value used to optimise the factory, but the loss in detection probability is not large.

Table 3 Actual probabilities of detection for the ten production line example when the factory is optimised for $\delta = 2$ but F_1 occurs with δ as shown in the left column of the table

P_d	C-ARL ₀ = 10	C-ARL ₀ = 20	C-ARL ₀ = 30	C-ARL ₀ = 37	C-ARL ₀ = 50
Observed $\delta = 1$	0.131	0.086	0.067	0.058	0.048
Observed $\delta = 2$	0.395	0.310	0.266	0.245	0.217
Observed $\delta = 3$	0.716	0.635	0.587	0.562	0.527
Observed $\delta = 4$	0.923	0.883	0.856	0.841	0.818

5 Conclusions

In this paper I have described a framework for optimising control limits for a system of n Shewhart \bar{x} -charts where, for whatever reason, some of the production lines are more likely to go out of control than others. Using standard practices, the factory would likely set the control limits equally on all the Shewhart charts. However, that would mean

less-than-optimal factory performance since, ideally, one would want to set the control limits to be more sensitive to catching the line or lines more likely to go out of control. The methodology presented in this paper provides such a means for optimising the control limits. It requires a change in the way one thinks about the design of control charts since the measures used to find the optimal control limits are at the factory level and are not in terms of in-control and out-of-control average run lengths.

Clearly this approach applies when there is a differential probability that parallel production lines will go out of control. The greater the disparity, the more relevant and important it is to take this approach rather than the traditional one of setting the control limits equally among all the production lines. An extreme example: Consider a factory with two production lines, one of which never goes out of control. Obviously a control chart applied to the line that never goes out of control is a waste of resources since it will only result in false positive signals. For a fixed amount of resources for investigating and adjudication false signals at the factory, it makes most sense to apply all of those resources only to the line that can go out-of-control. That is what the optimisation would do as well by setting the control limits so wide on the 'perfect' line that false positives would be impossible and appropriately smaller on the other line which would, as a result, have more false positives (to the level specified), but it would also be able to more quickly signal when the line when it goes out of control.

My motivation for this problem is a factory using Shewhart \bar{x} -charts. This also allowed for an important assumption that greatly simplified the optimisation calculations, namely that control chart signals are independent over time. However, there are other control charting methods that use both current and historical information, such as the CUSUM and EWMA, for which additional research is required to determine how to implement an equivalent approach. Certainly the idea is relevant – those control charts could also be applied to production lines with unequal probabilities of going out of control – but because the distribution at each time period is conditional on the history up to that time period, the calculations for probability of detection and combined in-control ARL are surely more complicated.

I conclude by noting that this methodology does not only apply to industrial quality control systems using Shewhart charts. Systems of threshold-based sensors (i.e., radar and sonar) have historically been used in military applications and, with today's increasing computing power and miniaturisation, systems of sensors are proliferating well beyond the military. Applications are present in many diverse fields such as meteorology, supply chain management, equipment and production monitoring, healthcare, production automation, traffic control, habitat monitoring, and health surveillance. See, for example, Gehrke and Liu (2007), Xu (2007), Intel (2007), Trigoni (2004), and Bonnet (2004). This methodology can potentially be applied to any application that uses threshold detection-based sensors. See Fricker and Banschbach (to appear) for one such example.

Acknowledgements

The author's work was partially supported by Office of Naval Research grant N0001407WR20172.

References

- Bonnet, P. (2004) *Sensor Network Applications*, available at www.diku.dk/undervisning/2004v/336/Slides/SN-applications.pdf (accessed on October 2, 2007).
- Fricker, R.D., Jr. and Banschbach, D. (to appear) 'Optimizing biosurveillance systems that use threshold-based event detection methods', *Information Fusion*.
- Gehrke, J. and Liu, L. (2007) *Sensor Network Applications*, available at <http://dsonline.computer.org/portal/site/dsonline/menuitem.6dd2a408dbe4a94be487e0606bcd45f3/indexjsp?&pName=dsolevel1article&TheCat=1015&path=dsonline/2006/04&file=w2gei.xml&> (accessed on October 2, 2007).
- Intel (2007) *Sensor Nets/RFID Website*, available at www.intel.com/research/exploratory/wireless_sensors.htm (accessed on October 2, 2007).
- Trigoni, N. (2004) *Sensor Networks: Applications and Research Challenges*, available at <http://locationprivacy.objectis.net/talks/5trigoni> (accessed on October 2, 2007).
- Xu, N. (2007) *A Survey of Sensor Network Applications*, available at <http://courses.cs.tamu.edu/rabi/cpsc617/resources/sensor%20nw-survey.pdf> (accessed on October 2, 2007).