Limited Memory Influence Diagrams for Attribute Statistical Process Control with Variable Sample Sizes

Barry R. Cobb cobbbr@vmi.edu
Department of Economics and Business, Virginia Military Institute, Lexington, VA, U.S.A.

Abstract

Limited Memory Influence Diagrams (LIMIDs) are implemented for statistical process control (SPC) to monitor the quality of the output from a production process where the number of defective units in a sample is measured at each time period. The observed defectives provide the input to a decision on whether to stop the process and repair a problematic cause of variation. The model also allows the decision maker to increase the size of the next sample in order to better discern whether or not the process actually requires investigation. The model only requires the user to know the size and result of the current sample to make a decision, in contrast to Bayesian methods that require calculations based on all prior samples and a history of actions. Despite the limited information, the model provides competitive quality costs to existing methods for a wide range of production time horizons.

Keywords: Attribute data; limited memory influence diagram; control chart; defectives; quality control; scatter search; statistical process control.

1. Introduction

Quality can be defined for a business as meeting customer expectations for its products and services. Statistical process control (SPC) tools are utilized to monitor processes to determine whether or not customer expectations are being met. This paper describes a limited memory influence diagram (LIMID) model that can be applied when data on the number of defective items in a sample is available to monitor a production process. The model yields a strategy that informs the manager whether or not the process should be stopped to investigate potential causes of variation that affect the quality of the output. The method also allows the sample size in the interval to be adjusted based on the current result, and this feature reduces the quality costs as compared to static sample size methods.

The production process operates in either an "in-control" and fully functional state or is "out-of-control." A randomly occurring assignable cause of variation shifts the working process from the in-control state. Assignable causes of variation can be traced to a machine, person, or material not performing as intended, and controlling the process requires the operator to distinguish these from normal, common causes of variation. The process state is not directly observable, so a random sample of output is collected at equal intervals and the number of defective units is recorded. The decision problem facing the process manager is whether to stop and intervene to return the process to the in-control state.

Control charts were originally created by Shewhart (1931) to distinguish between common and assignable causes of variation in production processes. The traditional approach to SPC using control charts is to examine a plot of sample data and evaluate whether a sample mean, range, or proportion falls outside of control limits established using either the sample data or a desired process average. The process is allowed to operate for a certain number of periods while the sample data is collected, then a judgment is made using the sample data regarding whether an assignable cause may have occurred.

Duncan (1956) introduced *economic design* of control charts where the user-defined parameters of the control chart – the sample size, sampling interval, and control limits – were chosen to minimize relevant costs. This paper employs an economic design with LIMIDs to determine optimal decision rules and minimize relevant costs. LIMIDs were introduced by Lauritzen and Nilsson (2001) as an alternative to traditional influence diagrams (Howard and Matheson, 2005) and partially-observed Markov decision processes (POMDPs) (Smallwood and Sondik, 1973) where prior observations of chance variables and previous decisions are not required in the solution procedure – a *limited memory* assumption. LIMIDs were recently applied in SPC to monitor process output using qualitative (or attribute) data for the number of defectives in a sample of the output (Cobb, 2021). The goal of this paper is to extend the application of LIMIDs to SPC with data on the number of defectives when the sample size for the next period can be adapted based on the current result. This is referred to as a variable sample size (VSS) feature and is effective at reducing quality-related costs.

Methods for SPC exist that allow sample size and sampling intervals to vary throughout the production horizon, and Bayesian SPC methods have been created that incorporate all sample data collected since the last process repair to be considered in the current decision on whether to investigate the process. The most closely related SPC methods to the one presented in this paper are the POMDP model of Calabrese (1995) that incorporates the number of defectives from a fixed sample size, a Bayesian control chart for sample number of defectives with adaptive sample sizes developed by Kooli and Limam (2009), the adaptive chart for number of defectives introduced by Kooli and Limam (2015), and the LIMID designed to utilize data on sample defectives from fixed sample sizes Cobb (2021). Section 4 will compare the VSS LIMID approach to these methods.

The paper proceeds as follows. The next section describes the process assumptions and the LIMID model. Section 3 gives parameters for example problems and results for solutions to these problems. Section 4 compares the results for the VSS LIMID model to closely related techniques. Section 5 describes potential future research in this area.

2. Model Description

This section describes the production process and LIMID model considered in this paper.

2.1 Production Process

The production process includes a set of assumptions common to many SPC models, including the early research of Duncan (1956), as well as the Bayesian techniques for attribute data outlined by Calabrese (1995) and Kooli and Limam (2009). The process begins in-control. The rate of occurrence of the assignable cause is exponential with parameter λ such that the mean time until occurrence is $1/\lambda$. The probability of an assignable cause occurring in a time interval of length h is $\gamma = 1 - e^{-\lambda h}$. Inspections occur at time periods indexed by $t = 1, \ldots, T-1$ and the random variable for the system state at time t is S_t with states s_{0t} (in-control) and s_{1t} (out-of-control). For all variables, if the time period is clear from the context, the t in the subscript may be dropped.

The process is evaluated by collecting a random sample of either n_1 or n_2 units, where $n_1 \leq n_2$, from the process in each period t (every h hours). The number of defective units in the sample is a random variable R_t with values $r_0, \ldots, r_{n_1}$ for the smaller sample size and $r_0, \ldots, r_{n_2}$ for the larger sample size. The number of defectives observed at time period t is denoted by r_{kt}. The proportion of defective units produced is p_0 when the process is operating in control. Eventually, an assignable cause of variation will shift the proportion of defective output to $p_1 > p_0$.

Two possibilities are considered regarding the length of the production horizon:

Finite horizon: The production horizon is a finite number of hours H such that the T time periods are each of h hours, $H \approx T \cdot h$. The horizon H can be the time until the next scheduled maintenance.

Infinite horizon: The process operates perpetually and only undergoes maintenance when the model indicates. Though the solution process for the LIMID model requires a finite number of time periods, to approximate the system operating in an infinite horizon we can establish a horizon H long enough that the probability the process shifts to the out-of-control state is high. The resulting strategy is adequate to monitor an infinite horizon process. Two such horizons H will be considered. With $\gamma = 1 - e^{-\lambda h}$, a time horizon that equates to the 95th percentile of the time between occurrences of the assignable cause will be denoted by $H_m = -\ln(0.05)/\lambda$ (subscript m for *maximum*). An alternate solution will be determined for the expected time between assignable causes $H_e = 1/\lambda$ (subscript e for *expected*).

Based on the observation $R_t = r_{kt}$ of the number of defectives, a decision A_t is made with three choices a_{jt} possible: a_{0t} signifying no investigation with a sample of n_1 units in the next period, a_{1t} with no investigation but a larger sample of n_2 units in the next period, and a_{2t} representing an investigation and repair (if required). When the latter action is taken, the sample size returns to n_1 in the next period. These decisions are represented by the rectangles in the graphical representation of the LIMID utilized for monitoring the production process as shown in Figure 1 for the scenario where $T = 5$.

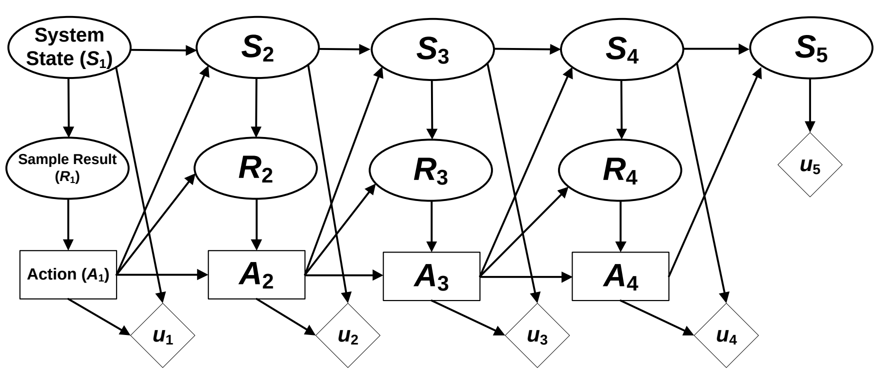

Figure 1: LIMID for monitoring the production process with a variable sample size.

2.2 Numerical Representation

In the LIMID representation of Figure 1, conditional probability potentials are assigned to each random variable represented by an oval in the diagram.

The probability distributions for S_t given the prior state of the system S_{t-1} and prior action A_{t-1} are composed as follows:

$$
\begin{aligned}
&P(S_t = s_0|\{S_{t-1} = s_0, A_{t-1} = a_0\}) = 1-\gamma\ , && P(S_t = s_1|\{S_{t-1} = s_0, A_{t-1} = a_0\}) = \gamma \\
&P(S_t = s_0|\{S_{t-1} = s_0, A_{t-1} = a_1\}) = 1-\gamma\ , && P(S_t = s_1|\{S_{t-1} = s_0, A_{t-1} = a_1\}) = \gamma \\
&P(S_t = s_0|\{S_{t-1} = s_0, A_{t-1} = a_2\}) = 1-\gamma\ , && P(S_t = s_1|\{S_{t-1} = s_0, A_{t-1} = a_2\}) = \gamma \\
&P(S_t = s_0|\{S_{t-1} = s_1, A_{t-1} = a_0\}) = 0\ , && P(S_t = s_1|\{S_{t-1} = s_1, A_{t-1} = a_0\}) = 1 \\
&P(S_t = s_0|\{S_{t-1} = s_1, A_{t-1} = a_1\}) = 0\ , && P(S_t = s_1|\{S_{t-1} = s_1, A_{t-1} = a_1\}) = 1 \\
&P(S_t = s_0|\{S_{t-1} = s_1, A_{t-1} = a_2\}) = 1-\gamma\ , && P(S_t = s_1|\{S_{t-1} = s_1, A_{t-1} = a_2\}) = \gamma
\end{aligned}
$$

The probabilities when the process was previously in state s_1 and action a_2 was not pursued are defined assuming the process does not spontaneously repair itself.

The sample results R_t are binomially distributed as:

$$
\begin{aligned}
&R_t|\left\{A_{t-1} = a_0, S_t = s_0\right\} \sim B\left(n_1, p_0\right)\ , && R_t|\left\{A_{t-1} = a_0, S_t = s_1\right\} \sim B\left(n_1, p_1\right) \\
&R_t|\left\{A_{t-1} = a_1, S_t = s_0\right\} \sim B\left(n_2, p_0\right)\ , && R_t|\left\{A_{t-1} = a_1, S_t = s_1\right\} \sim B\left(n_2, p_1\right) \\
&R_t|\left\{A_{t-1} = a_2, S_t = s_0\right\} \sim B\left(n_1, p_0\right)\ , && R_t|\left\{A_{t-1} = a_2, S_t = s_1\right\} \sim B\left(n_1, p_1\right)
\end{aligned}
$$

2.3 Utility (Cost) Functions

The following cost assumptions are made in the problem. An investigation cost $I > 0$ is incurred if the process is stopped in order to determine whether an assignable cause is present. If the process requires repair, a repair (fix) cost $F \geq 0$ is added to I to bring the process back into control. The cost per hour with higher defective output is $M > 0$.

If the process shifts from s_0 to s_1 in time period t, the expected time of occurrence in the interval is

$$
\tau = \left(\int_{th}^{(t+1)h} e^{-\lambda u} \lambda (u - th)\, du \right) \Big/ \left(\int_{th}^{(t+1)h} e^{-\lambda u} \lambda\, du \right) = \frac{1 - (1 + \lambda h) e^{-\lambda h}}{\lambda (1 - e^{-\lambda h})}\ .
$$

This means that $(h - \tau)$ is the expected length of the process shift in the h hour interval, so $\gamma M(h - \tau) = M(h - \gamma/\lambda)$ is the expected cost of producing at the higher defect rate, with the inspection cost I and repair cost F added when action a_2 is taken. The fixed cost of drawing a sample is c_1, and the cost per unit for sampling is c_2, so sampling costs of either $c_1 + c_2 \cdot n_1$ or $c_1 + c_2 \cdot n_2$ are incurred depending on the action taken. An additional $c_1 + c_2 \cdot n_1$ is added in the first period assuming the first sample is size n_1, and no sampling cost is incurred in period $T-1$ because no sample is drawn in period T. These utility values are consistent with the closely related prior research (Calabrese, 1995; Kooli and Limam, 2009, 2015; Cobb, 2021).

The utility function u_t for period t is summarized in Table 1.

Table 1: Utility (cost) values for the LIMID.

S_t	A_t	u_1	u_{T-1}
s_0	a_0	$M(h-\gamma/\lambda)+2\cdot(c_1+c_2\cdot n_1)$	$M(h-\gamma/\lambda)$
s_0	a_1	$M(h-\gamma/\lambda)+2c_1+c_2\cdot(n_1+n_2)$	$M(h-\gamma/\lambda)$
s_0	a_2	$I+M(h-\gamma/\lambda)+2\cdot(c_1+c_2\cdot n_1)$	$I+M(h-\gamma/\lambda)$
s_1	a_0	$Mh+2\cdot(c_1+c_2\cdot n_1)$	Mh
s_1	a_1	$Mh+2c_1+c_2\cdot(n_1+n_2)$	Mh
s_1	a_2	$I+F+M(h-\gamma/\lambda)+2\cdot(c_1+c_2\cdot n_1)$	$I+F+M(h-\gamma/\lambda)$
S_t	A_t	u_t, $t=2,\ldots,T-2$	u_T
s_0	a_0	$M(h-\gamma/\lambda)+c_1+c_2\cdot n_1$	$M(h-\gamma/\lambda)$
s_0	a_1	$I+M(h-\gamma/\lambda)+c_1+c_2\cdot n_2$	$M(h-\gamma/\lambda)$
s_0	a_2	$I+M(h-\gamma/\lambda)+c_1+c_2\cdot n_1$	$M(h-\gamma/\lambda)$
s_1	a_0	$Mh+c_1+c_2\cdot n_1$	Mh
s_1	a_1	$Mh+c_1+c_2\cdot n_2$	Mh
s_1	a_2	$I+F+M(h-\gamma/\lambda)+c_1+c_2\cdot n_1$	Mh

2.4 Policies and Strategies

The strategy δ composed of policies $\delta^1,\ldots,\delta^{T-1}$ for $A_1,\ldots,A_{T-1}$ has values $\delta^t_{i,j,k}$ for $i=0,1,2$ and $j=0,1,2$. The values for δ^t are defined for $k=0,\ldots,n_2$ when $i=1$ and for $k=0,\ldots n_1$ otherwise. The index i references the prior action a_0, a_1, or a_2 taken at node A_{t-1}, and the index j represents the selected value of the current action node A_t. The index k references an observed number of defects in the current sample result R_t.

Each policy δ^t is composed of $(6\cdot(n_1+1))+(3\cdot(n_2+1))$ elements. Suppose that for $n_1=1$ and $n_2=3$, the policy δ^t contains the value 1 for

$$\left\{\delta^t_{0,0,0},\delta^t_{0,1,1},\delta^t_{1,0,0},\delta^t_{1,1,1},\delta^t_{1,1,2},\delta^t_{1,2,3},\delta^t_{2,0,0},\delta^t_{2,1,1}\right\}$$

and values of zero for all other elements. For example, the strategy dictates with the value $\delta^t_{0,1,1}=1$ that when $A_{t-1}=a_0$ and the current sample is $R_t=r_1$ (one defective unit), the action $A_t=a_1$ should be selected and the next sample should be of size n_2. For a prior value of $A_{t-1}=a_1$, a sample size of $n_2=3$ was employed and the policy specifies three scenarios. If $R_t=0$ (no defective units), the value $\delta^t_{1,0,0}=1$ recommends that the sample size revert to $n_1=1$. Values $\delta^t_{1,1,1}=\delta^t_{1,1,2}=1$ specify that for an observation of one or two defective units, the sample size of $n_2=3$ should be maintained in the next period. If three defectives are observed, action a_2 should be pursued and the process should be investigated and (if necessary) repaired (indicated by $\delta^t_{1,2,3}=1$).

2.5 Solution

The LIMID is solved by single policy updating (SPU) via message passing in a junction tree (Cowell et al., 1999). The SPU algorithm proceeds by passing messages toward a root node containing a decision variable and its parents from all other nodes in the junction tree. As each root node for each decision variable, A_t, is encountered, the policy δ^t is updated. A forward and backward pass through all the root nodes containing decisions constitutes one iteration of SPU. In a subsequent iteration, some policies δ^t can change to improve the solution based on the best policy for the other decision variables. Iterations of SPU are continued until the minimum total cost TC^* is unchanged. This is the optimal solution to the LIMID and the current strategy set δ is the best set of policies. Further details of the solution algorithm are omitted here, but the reader can refer to Lauritzen and Nilsson (2001) for complete details.

The LIMID solution procedure avoids the *no forgetting* assumption of traditional IDs that would require the current action A_t to be a function of all prior actions $A_1, \ldots, A_{t-1}$ and all prior sample and current sample results $R_1, \ldots, R_t$. Despite utilizing limited information, the iterative nature of the SPU procedure allows the LIMID to develop decision policies that provide expected utility values that are similar to a traditional ID model in many cases. While there is some expected utility loss due to a lack of memory as compared to traditional influence diagrams or the method employed by Kooli and Limam (2009), the LIMID is more tractable when implemented for longer time horizons, as will be shown in the next section.

For some production processes, the parameters T, n_1, and n_2 may be established operationally based on the judgment of a manager. In this situation, the LIMID is simply solved to determine the best strategy and expected cost for those parameters. In other cases, an economic design can be used to select T, n_1, and n_2 to minimize quality control costs. This involves solving the LIMID for a wide range of values for each parameter to establish a set of parameters that performs well. In this paper a scatter search approach is utilized (Laguna and Martí, 2003) to iteratively narrow the search for a desired parameter set, $\{T^*, n_1^*, n_2^*\}$. Scatter search begins by selecting a set of random values for the decision variables within reasonable operational bounds. For examples in the next section, the initial solutions include n_1 and n_2 in the range $[1, 150]$ and T in the range $[2, 2H]$. The model is solved for the initial parameter sets to create a *reference set* of solutions ranked by total cost. The parameter set is then refined to include convex combinations of those parameters in the previous reference set that produced the lowest cost solutions, i.e. these are in the *neighborhood* of the best solutions in the previous reference set. Some new randomly generated parameter sets are added at each iteration to reduce the likelihood that the procedure becomes settled on a local minimum. This continues until there is no change in the minimum total cost for the best parameter set on two iterations.

3. Results

To examine the VSS LIMID method, the example case with parameters in Table 2 is solved as a baseline scenario and changes in individual inputs are considered in sensitivity analysis.

Table 2: Parameters for baseline example case.

M	I	F	c_1	c_2	λ	p_0	p_1	H_e
100	500	250	1	1	0.01	0.05	0.20	100

Table 3 displays the average hourly cost (AHC) for the LIMID model with a static sample size along with the AHC for the VSS LIMID model, along with the optimal parameters for each model. Note that the hourly interval between samples $h^* = H/T^*$. Results for time horizons up to $H_m = 300$ are displayed. The term *optimal* in this case simply means the best solution as determined by the SPU technique, not the best solution that could be provided by a traditional influence diagram model, if it were in fact tractable. This is discussed extensively in (Cobb, 2021). The rows corresponding to time horizons H_e and H_m for baseline assumptions are shown in *italics* and **bold** font, respectively. The left panel of Figure 2 displays the AHC values for production horizons between $H = 40$ and $H_m = 300$.

Table 3: Results for AHC in example case.

	Static				VSS						Static	VSS
H	*AHC*	n^*	T^*	h^*	*AHC*	n_1^*	n_2^*	T^*	h^*	c_2	*AHC*	*AHC*
40	16.04	21	6	6.67	14.34	1	36	45	0.89	0.10	11.59	11.52
60	16.77	22	9	6.67	15.18	2	35	40	1.50	0.25	13.34	12.80
80	17.13	30	11	7.27	15.53	1	28	83	0.96	0.50	15.11	14.31
100	*17.32*	*31*	*13*	*7.69*	*15.79*	*2*	*20*	*73*	*1.37*	0.75	16.34	15.06
120	17.46	30	16	7.50	16.04	2	22	73	1.64	1.00	*17.32*	*15.79*
150	17.59	31	20	7.50	16.27	3	29	80	1.88	1.25	18.20	16.51
180	17.67	31	23	7.83	16.41	3	29	84	2.02	1.50	18.79	17.23
240	17.77	31	31	7.74	16.88	4	34	90	2.67	1.75	19.41	17.94
300	**17.84**	**31**	**38**	**7.89**	**17.06**	**4**	**34**	**100**	**3.00**	2.00	20.01	18.64

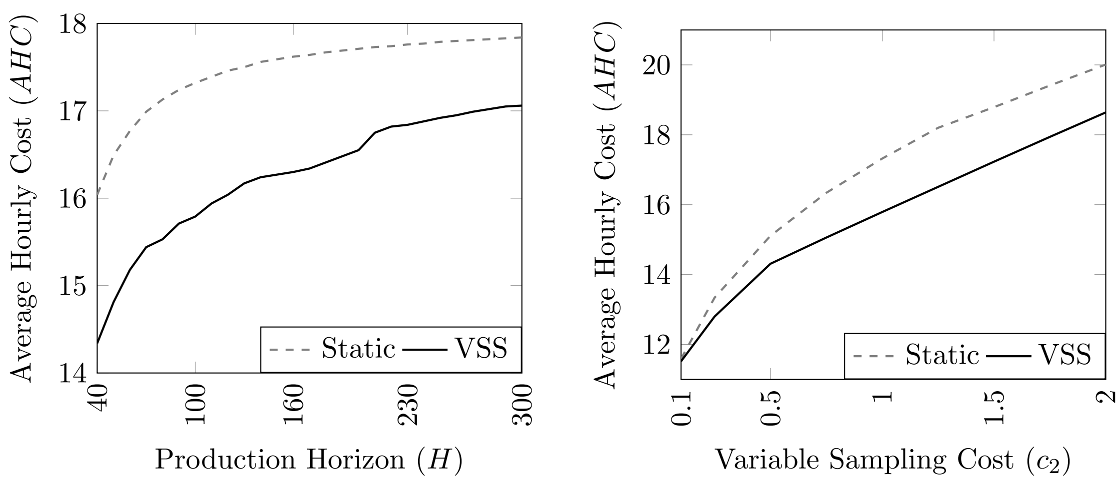

Figure 2: Average hourly cost for LIMID models in example case for various production horizons H and variable sampling costs c_2.

The AHC increases with the length of the production horizon H because the process starts in-control and there is a greater likelihood the process enters the out-of-control state more often for longer H. The VSS LIMID provides a lower AHC than the LIMID with a fixed sample size for all time horizons, with larger savings occurring for shorter time horizons.

The last two columns of Table 3 also show the sensitivity of the AHC in the two models to the variable sampling cost parameter c_2, all for time horizons of $H_e = 100$. When variable sampling cost is very small, the static LIMID model has a very similar AHC to the VSS model; however, as sampling becomes more expensive on a per unit basis, the VSS model is able to provide a decision rule and sampling policy that significantly reduces the average costs of quality. The right panel of Figure 2 shows the AHC values for each model in these scenarios.

The policies for the baseline scenario with $H_e = 100$ are shown in Figure 3. Separate elements of the policies are required for the scenarios where the system was operating in $A_{t-1} = a_0$ and $A_{t-1} = a_1$. The state a_0 involves drawing a small sample size of $n_1 = 2$. In the left panel of Figure 3, note that for $t = 1$ the policy dictates that the system will always continue in state a_0 to period $t = 2$. Similarly, the system will never deviate from a_0 after period $t = 64$. In all other periods, the system stays in a_0 if r_{0t} is observed, and moves to state a_1 and a larger sample size of $n_2 = 20$ if 1 or 2 defectives are observed. When $A_{t-1} = a_1$ was selected (see the right panel of Figure 3), the system moves back to a_0 if $r_{kt} < 2$ for most of the production horizon, while $A_t = a_1$ will be maintained when 2 or 3 defectives are observed. When 4 or more defectives are sampled, the system moves to a_2 and an investigation occurs. This policy holds until $t = 63$ when only a move from a_0 to a_2 is recommended at $r_{3,63} = 3$. After $t = 63$, the system always returns to a_0.

Figure 4 displays the policies that are determined for the case where the variable sampling cost is $c_2 = 0.25$ (right panel of Figure 2). The optimal sample sizes are $n_1 = 5$ and $n_2 = 52$, and the decision rules suggest periodically relaxing and tightening the boundaries for sample results that require a larger sample size or an inspection. This pattern provides a balance between sampling cost and the other costs of quality.

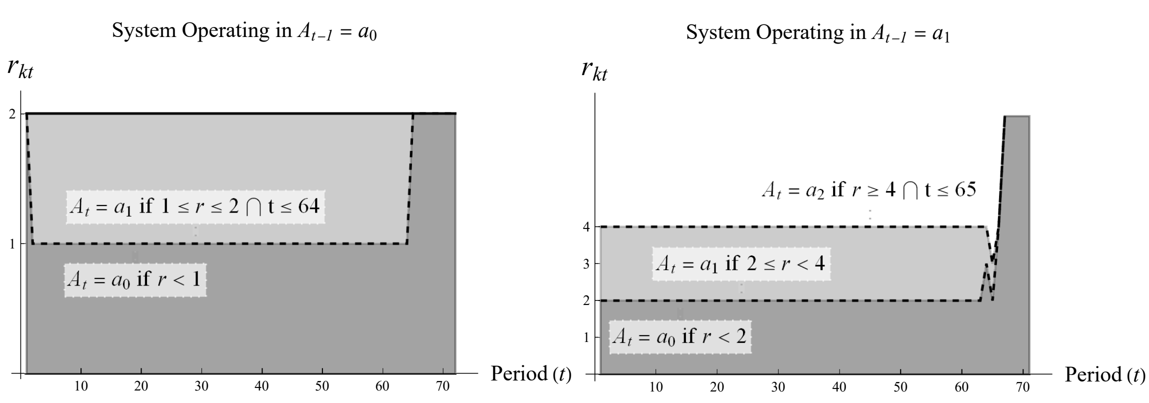

Figure 3: Policies for baseline case when system is operating in $A_{t-1} = a_0$ (left) and $A_{t-1} = a_1$ (right).

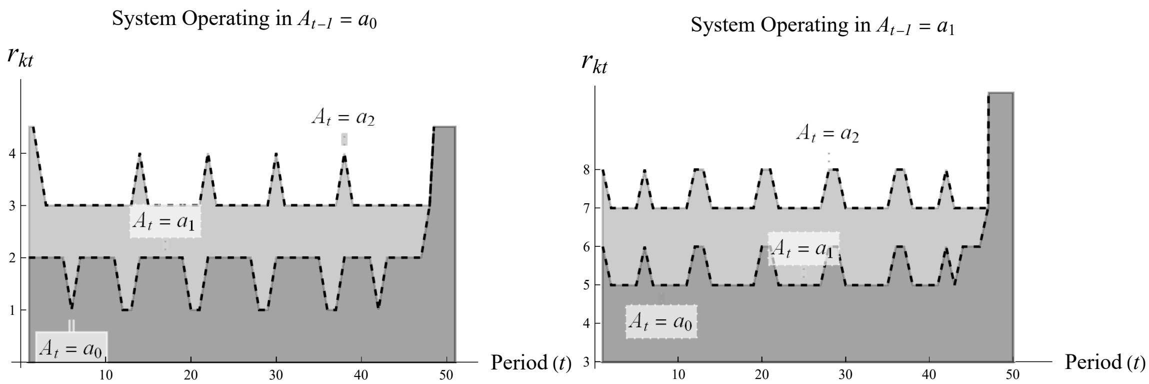

Figure 4: Policies for scenario where $c_2 = 0.25$ when system is operating in $A_{t-1} = a_0$ (left) and $A_{t-1} = a_1$ (right).

4. Comparison

This section provides a comparison of the VSS LIMID model to the four most closely related models in the literature.

4.1 Comparison to Kooli and Limam (2009)

The Bayesian VSS model presented by Kooli and Limam (2009) is solved via dynamic programming and enumerates the expected continuation cost associated with each potential sample size at each sampling interval. The next sample size is selected based on these calculations. Implementing the model in practice requires all of these continuation values to be recalculated during the production horizon based on the up-to-date history of actions and sample results. Kooli and Limam (2009) only implement the model for relatively short production horizons, likely due to the computational complexity of the method.

The VSS LIMID was implemented for one of the cases in Kooli and Limam (2009) with parameters and results shown below:

M	I	F	c_1	c_2	λ	p_0	p_1	H
100	100	600	10	0.30	0.01	0.05	0.20	120

VSS LIMID					Bayesian VSS		
n_1^*	n_2^*	T^*	h^*	AHC	T^*	h^*	AHC
38	120	17	7.06	13.69	16	7.50	13.58

The Bayesian VSS has an average hourly cost that is about 0.8% lower than the VSS LIMID. The examples in Kooli and Limam (2009) are primarily ones that have a relatively high fixed sampling cost c_1 and a low variable sampling cost c_2. This scenario is likely to be one (based for instance on evidence in Figure 2) where the VSS feature has the least value and the hourly time between samples (and thus the number of sampling periods) is likely to be low. The latter would be a more tractable scenario for the Bayesian VSS.

Interestingly, when the VSS LIMID model was implemented, the decision rules typically only recommended a larger sample size early in the production horizon, as the optimal n_1^* was relatively large and proved to be adequate for most time periods.

The Bayesian VSS may be well-suited to short time horizons where sampling costs are low, but carries the burden of a significant computational requirement during production and may not scale well to longer production horizons.

4.2 Comparison to Kooli and Limam (2015)

The economically designed *np* control charts with variable sample sizes introduced by Kooli and Limam (2015) are primarily structured to operate in processes with infinite production horizons. This is due to the fact that selection of optimal parameters relies on analytical formulas for total costs that are based on long-run expected values. These analytical solutions are used to enumerate a large number of potential solutions to determine the optimal sample sizes, time between samples, and control limits.

The VSS LIMID was implemented for one of the cases in Kooli and Limam (2015) with parameters and results shown below:

M	I	F	c_1	c_2	λ	p_0	p_1	H_m
160	250	250	1	0.10	0.008	0.03	0.1153	375

VSS LIMID					VSS *np* chart			
n_1^*	n_2^*	T^*	h^*	AHC	n_1^*	n_2^*	h^*	AHC
71	91	120	3.13	10.13	72	103	3.27	10.20

Because the VSS *np* chart is designed for an infinite production horizon, the VSS LIMID has been tested over the long production horizon of $H_m = 375$. There is a 95% probability that the assignable cause occurs during that period of time, so this period is used to make a reasonable comparison with the infinite horizon. The VSS LIMID average cost is lower than the VSS *np* chart for this case. If the VSS *np* chart is implemented for a shorter time horizon, the VSS LIMID strategy would likely adapt so that the average cost would be lower, as illustrated in the left panel of Figure 2.

4.3 Comparison to Calabrese (1995)

The POMDP method of Calabrese (1995) utilizes a static sample size and is implemented for relatively short time horizons. The comparison of the static LIMID with the POMDP method is made extensively in Cobb (2021). In general, the static LIMID has average hourly costs that are very similar (sometimes lower) to the POMDP and the computational time is often significantly lower. Additionally, the LIMID (either the static or VSS version) does not require additional calculations during the production horizon, whereas the POMDP requires the updated probability that the system is in-control to be calculated after each sample is collected.

For the baseline scenario in the previous section, the POMDP model provides an AHC of 17.73 with a static sample size of $n^* = 13$ and $T^* = 27$ for the $H_e = 100$ production horizon. The computational time is 145 seconds. A VSS version of the POMDP method could be developed, but it seems unlikely that it would outperform the VSS LIMID in terms of cost or computational burden, given the performance of the static version.

4.4 Comparison to Cobb (2021)

This comparison has been made directly in this paper and the reduction in average total cost from applying the VSS feature to the LIMID is apparent in the results listed in Table 3. Both the static and VSS methods can be implemented for both short and long time horizons, with the VSS providing lower costs when changing the sample size is operationally feasible. Neither method requires further calculations during the production horizon based on the past history of actions and sample results.

Moving from the static to the VSS model does entail additional computational complexity to determine the optimal sample size(s) and interval between samples. The baseline scenario from the last section requires about 0.50 seconds to run in Wolfram Mathematica® 13.0 on a computer with a 2.80 GHz processor with 16.0 GB of RAM. The VSS model requires 18 seconds of computing time for the same example. Employing the scatter search technique typically required evaluating about 60 parameter sets for the static model and 100 parameter sets for the VSS model.

4.5 Summary

The significance of the VSS LIMID model is that it provides improved quality control costs to most of the previous models utilized for the production process introduced in Section 2.1. This process involves a single assignable cause that shifts the percentage of defective items produced. The only model that slightly outperforms the VSS LIMID model is the Bayesian VSS model, although that model would likely be difficult to extend to longer production horizons. Applying the LIMID in some form to other production processes, such as those that are subject to more than one assignable cause, can be a topic of future research. The next section mentions additional future opportunities to adapt LIMID models for SPC.

5. Future Research

The LIMID model can be extended to additional problems in quality control, including:

1. **Variable sampling intervals** – Varying the sampling interval based on current sample results, either in place of or in additional to varying the sample size, is an option to improve SPC models. A LIMID model could be offered as an alternative to the variable sampling interval *np* control chart of Kooli and Limam (2015).

2. **Attribute data on nonconformities** – Some aspects of quality control are appropriately measured using data on the number of errors or nonconformities, as opposed to data on the proportion of defective units. No economically designed or Bayesian SPC models exist for this scenario. An alternative to the statistically designed c control chart of Epprecht et al. (2003) would be an interesting extension of the LIMID in both the static and VSS scenarios.

3. **Variable data** – This paper has dealt with attribute data on the number of defectives in a sample, but a wide range of SPC model exist for use with variable data measured

on a continuous scale. Variable control chart techniques outlined by Nenes (2013) would potentially be a point of comparison for new LIMID methods.

References

J. M. Calabrese. Bayesian process control for attributes. *Management Science*, 41(4): 637–645, 1995.

B. R. Cobb. Statistical process control for the number of defectives with limited memory. *Decision Analysis*, 18(3):203–217, 2021.

R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter. *Probabilistic Networks and Expert Systems*. Springer, New York, 1999.

A. J. Duncan. The economic design of $\overline{X}$ charts used to maintain current control of a process. *Journal of the American Statistical Association*, 51(274):228–242, 1956.

E. K. Epprecht, A. F. B. Costa, and F. C. T. Mendes. Adaptive control charts for attributes. *IIE Transactions*, 35(6):567–582, 2003.

R. Howard and J. Matheson. Influence diagrams. *Decision Analysis*, 2(3):127–143, 2005.

I. Kooli and M. Limam. Bayesian *np* control charts with adaptive sample size for finite production runs. *Quality and Reliability Engineering International*, 25(4):439–448, 2009.

I. Kooli and M. Limam. Economic design of an attribute *np* control chart using a variable sampling policy. *Applied Stochastic Models in Business and Industry*, 31(4):483–494, 2015.

M. Laguna and R. Martí. *Scatter Search: Methodology and Implementations in C*. Springer, New York, 2003.

S. L. Lauritzen and D. Nilsson. Representing and solving decision problems with limited information. *Management Science*, 47(9):1235–1251, 2001.

G. Nenes. Optimisation of fully adaptive Bayesian charts for infinite-horizon processes. *International Journal of Systems Science*, 44(2):289–305, 2013.

W. A. Shewhart. *Economic Control of a Quality Manufactured Product*. D. Van Nostrand, New York, NY, 1931.

R. Smallwood and E. Sondik. The optimal control of partially observable Markov decision processes over a finite horizon. *Operations Research*, 21(5):1071–1088, 1973.