A Bayes-Adjusted Cumulative Sum (Cusum)

by

Spencer Graves, Søren Bisgaard, and Murat Kulahci

Affiliations

Spencer Graves
PDF Solutions
333 West San Carlos, Suite 700
San José, CA 95126
spencer.graves@pdf.com

Søren Bisgaard
Eugene M. Isenberg Professor of Technology Management
Eugene M. Isenberg School of Management
University of Massachusetts-Amherst
121 Presidents Drive
Amherst, MA 01003
bisgaard@som.umass.edu

Murat Kulahci
Arizona State University
Dept. of Industrial Engineering
P. O. Box 875906
Tempe, AZ 85287
Murat.Kulahci@asu.edu

A Bayes-Adjusted Cumulative Sum (Cusum)

ABSTRACT

This article shows how the Girshick-Rubin (1952) monitor can be rewritten as a traditional one-sided Cusum plus a correction term that generally provides fairly consistent bias and is often quite small. For an abrupt jump from null to alternative hypotheses, the log odds for the transition is essentially a cumulative sum of log(likelihood ratio) with a floor given by the log(hazard odds). This explains the previously reported virtual equivalence between Girshick-Rubin and a Cusum. We also generalize Girshick-Rubin to non-i.i.d. observations and non-constant hazard. With zero hazard, it becomes the log(odds) formulation of Bayes' theorem, which is also Wald's sequential-probability ratio test (SPRT). Potential applications include On-Board Diagnostics (OBDs) of emissions controls, required on all new automobiles sold today in many countries. It also provides an important new tool for improving preventive maintenance schemes by combining reliability information with current data on the condition of the plant (i.e., system monitored); in applications with increasing hazard, these new preventive maintenance programs could reduce the total cost of operating and maintaining equipment.

KEY WORDS: Monitoring; Bayesian Sequential Updating; Preventive maintenance; Discrete hazard rate; On-board diagnostics (OBDs); Total Productive Maintenance (TPM); Discrete reliability; Fast Initial Response (FIR)

1. INTRODUCTION

The increasing computerization of products from simple to highly complex provides many new opportunities and demands for sophisticated diagnostic monitoring systems. Prime examples are provided by On-board diagnostics (OBDs) to detect malfunctions in emission controls required by law in new automobiles sold today in the US, Canada and Europe and soon in many other countries (e.g., Mondt 2000, p. 144; CARB 1997). Other examples include modern heart pacemakers and implantable defibrillators that monitor both the patient and the device itself (Gunderson 2000): They monitor the patient's condition and intervene only when necessary, and they sound an audible alarm if the battery is low or the electrical connections seem corroded. The techniques could also be used in medical applications to alert appropriate people of a change in a patient's condition, good or bad, suggesting a potential need to change therapy or shortening the time required to evaluate new medical procedures (Steiner et al. 2000).

Box and Luceño (1997, p. 233) describe a country whose national defense includes two kinds of radar: One omnidirectional, the other focused narrowly in the directions from which an enemy would most likely attack. They compare the omnidirectional radar to a Shewhart chart and the focused radar to a Cusum of Fisher's efficient score or log(likelihood ratio) tuned to a specific anticipated change. The massive quantities of data now being collected provide substantial opportunities for increased sophistication in automatic monitoring for many different kinds of changes. Suppose for example that a change in mean indicates an actual process change while an increase in variability suggests metrology problems. Monitors could be programmed to notify different people depending on the nature of a change detected.

A general principle for designing such monitoring systems is provided, we believe, by the two-step Bayesian sequential updating procedure outlined by Graves et al. (2001); see Figure 1: Step 1 uses Bayes' theorem to incorporate new data into current knowledge of the condition of the system being monitored (called the "plant" for consistency with the control theory literature). Step 2 models a possible transition in the condition of the plant before the next observation. For a plant that deteriorates following a normal random walk observed with normal error, this leads to Kalman filtering, a special case of which is an exponentially weighted moving average (EWMA), as explained by Graves et al. (2001) and Graves, Bisgaard and Kulahci (2002a, b). Section 5 below illustrates how models of increasing hazard can be introduced naturally into Step 2, providing an important new method for medical applications and for Total Productive Maintenance (TPM) programs.

(Figure 1 about here)

In this article, we consider an abrupt jump from an "in control" condition $\underline{H}_0$ to an alternate state $\underline{H}_1$; $\underline{H}_1$ may describe one of many possible "out of control" conditions. From this perspective, Shewhart-type charts are tuned to detect single, isolated events such as outliers, while other charts such as Cusums, EWMAs and Kalman filters look for sustained signals.

Girshick and Rubin (1952) considered Bayesian sequential updating with independent, identically distributed (i.i.d.) observations and a constant hazard rate for a transition from a simple (completely specified) hypothesis $\underline{H}_0$ to another simple hypothesis $\underline{H}_1$. They modeled this as a two-state, recurrent Markov chain with a cost model whose optimization involved a non-intuitive iteration; this was generalized to a continuous time stochastic process by Shiryaev (1963). Kenett and Zacks (1998) discuss examples and provide software to compute average run lengths for monitor design.

In this article, we generalize Girshick-Rubin to non-i.i.d. observations with non-constant hazard. This generalization opens up many new areas of potential application, including the on-board diagnostics (OBDs) of plants using data whose behavior changes dramatically with the dynamics of the plant and whose susceptibility to failure may change both with mode of use and with age, as illustrated by the example in Section 5 below.

The basic derivation and an i.i.d. example appear in the next section. With zero hazard, Bayesian sequential updating can be written as a cumulative sum (Cusum) of log(likelihood ratio) that is the log odds formulation of Bayes' theorem. With non-zero hazard, the log odds for the transition can never go below the log(hazard odds), and the resulting iteration closely approximates a one-sided Cusum.

This derivation also suggests a relationship between the threshold of Page's one-sided Cusum and the increase in posterior log(odds) for $\underline{H}_1$ against $\underline{H}_0$; this relationship is discussed in Section 3. This is followed by a brief discussion of costs and run lengths in monitor design. An extensive example involving non-identically distributed increments and non-constant hazard appears in Section 5. The difficulties of generalizing these results to composite hypotheses and to situations where the distribution of observations from $\underline{H}_1$ depends on the time of fault onset are briefly mentioned in Section 6. This is followed by concluding remarks.

2. NON-I.I.D. OBSERVATIONS, NON-CONSTANT HAZARD

We wish to generalize the Girshick-Rubin (1952) procedure to non-i.i.d. observations and non-constant hazard. In particular, suppose we observe random variables y_t that have a density $f_{i,t} = f_{i,t}(y_t \mid y_{t-1}, y_{t-2},...)$, $i = 0, 1$ for $\underline{H}_0$ and $\underline{H}_1$ (where $f_{0,t}$ and $f_{1,t}$ are densities with respect to a common dominating measure and may vary with t). Typically, $\underline{H}_0$ represents "good" and $\underline{H}_1$ "bad", and we will use those terms in this article, though that is not required.

Consider a random variable t_0 = the change point from $\underline{H}_0$ to $\underline{H}_1$. We further assume that the $f_{i,t}$'s are completely specified, which also implies that $f_{1,t}$ does not depend on t_0.

Let h_t = hazard rate = Pr{ $\underline{H}_1$ at t +1 | $\underline{H}_0$ at t }. Since $f_{1,t}$ does not depend on t_0, the Bayesian posterior can be summarized in a single number, Pr{ $\underline{H}_0$ at $t \mid y_1, ..., y_t$ }. In typical applications, we want to know when this change has occurred so appropriate action can be taken.

Step 1 of the two-step iteration of Figure 1 requires us to compute the posterior probability of $\underline{H}_0$ at time t given the prior, g_{t-1}. In this context, Step 1 (Bayes' theorem) gives us the following:

$$\Pr\{\ \underline{H}_0 \text{ at } t \mid y_1, ..., y_t\ \} = \frac{f_{0,t}\, g_{t-1}}{f_{0,t}\, g_{t-1} + f_{1,t}\left(1 - g_{t-1}\right)}.$$

Step 2 allows for a possible transition, giving us the prior for the next observation given the past as follows:

$$g_t = \Pr\{ \underline{H}_0 \text{ at } (t+1) \mid y_1, ..., y_t \}$$

$$= (1 - h_t) \Pr\{ \underline{H}_0 \text{ at } t \mid y_1, ..., y_t \} = \frac{(1-h_t) f_{0,t} g_{t-1}}{f_{0,t} g_{t-1} + f_{1,t}(1-g_{t-1})}, \qquad (1)$$

where g_0 = the initial prior probability that observation y_1 is collected under $\underline{H}_0$. We assume that this is strictly positive. Otherwise, with $g_0 = 0$ expression (1) says $g_t = 0$ for all t. But $g_0 = 0$ means that we are certain that the transition of interest has already occurred, and we will therefore take the action before the first observation is collected.

More generally, if an observation is highly informative, either $f_{0,t} >> f_{1,t}$ or $f_{0,t} << f_{1,t}$. In either of these cases, expression (1) essentially discards the past immediately, giving us $g_t \cong (1-h_t)$ in the first case and $g_t \cong 0$ in the second, almost regardless of g_{t-1}. At the other extreme, if an observation is completely noninformative, $f_{0,t} = f_{1,t}$, so $g_t = (1-h_t)g_{t-1}$. When all observations are noninformative, $g_t = g_0 \prod (1-h_t)$ = a deterministic march to 0 following the reliability distribution. Except in cases where the difference between $f_{0,t}$ and $f_{1,t}$ is always negligible {so $g_t \cong g_0 \prod (1-h_t)$}, the influence of g_0 on g_t will over time become dominated by the information contained in the data and in the reliability distribution.

We typically think of g_t as Pr("good"|...), though it might be something else. We focus on the prior of $\underline{H}_0$ given the past, because we are usually more concerned with the future than the past; if for example we planned to discontinue use before the next observation, we might not even collect y_t. In most cases, the hazard rate will be so small that the difference obtained by introducing $(1 - h_t)$ into (1) will be small; however, in certain applications with rapidly increasing hazard, this difference could be important; see the case study discussed with Figure 6 below.

In the following, we will convert g_t into log(odds). In so doing, we will obtain a cumulative sum with a floor given by the log hazard odds and with the initial prior log odds being a fast initial response (FIR) parameter (Lucas and Crossier 1982). This helps us understand FIR: it is essentially equivalent to specifying g_0 = the initial prior on the condition of the plant.

With one important restriction, nuisance parameters could be handled by tracking them using the same 2-step Bayesian sequential update cycle of Figure 1, then applying this formulation to the "predictive distribution" after integrating out the nuisance parameters. The important restriction is the assumption that the $f_{i,t}$'s do not depend on t_0, which will not be strictly true when the posterior distribution of the nuisance parameters depends on the change point, t_0. However, Menke and Maybeck (1995) and Eide and Maybeck (1996) seem to get good results with essentially this procedure ignoring the impact of the change point on the posterior (see also Graves, et al. 2001, 2002a, b).

To turn (1) into a Cusum, we rewrite it in term of $B_t = (1-g_t)/g_t$ = the odds for $\underline{H}_1$ ("bad" in certain applications) at time (t + 1), as follows:

$$B_t = [\ h_t + (f_{1,t}/f_{0,t})B_{t-1}\ ] / (1 - h_t)$$

or

$$B_t = H_t + z_t B_{t-1}, \qquad (2)$$

where $H_t = h_t/(1 - h_t)$ = hazard odds, and $z_t = (f_{1,t}/f_{0,t})/(1 - h_t)$ = adjusted likelihood ratio. Without the transition (i.e., if $h_t = 0$), this is merely the odds formulation of Bayes' theorem: The posterior odds is the likelihood ratio times the prior odds.

Girshick and Rubin converted (1) into a recursion for $Z_t = \left[1/g_t - 1/(1-h_t)\right]/H_t$ = $(z_t B_{t-1}/H_t)$. By definition, B_{t-1} = $\{1/g_{t-1} - 1\}$ = $\{1/g_{t-1} - 1/(1-h_{t-1}) + H_{t-1}\}$ = $\{Z_{t-1} + 1\}H_{t-1}$. Whence,

$$Z_t = z_t\left(1+Z_{t-1}\right)\left(H_{t-1}/H_t\right). \tag{3}$$

Girshick and Rubin obtained the constant-hazard simplification of this: $Z_t = z_t\left(1+Z_{t-1}\right)$. A simple recursion such as (3) can be quite valuable for computations. However, we have been unable to find a simple interpretation for Z_t. Moreover, the presence in (3) of the factor $\left(H_{t-1}/H_t\right)$ suggests that any meaning that might be assigned to a fixed point like $Z_t = 1$ could change with non-constant hazard. We shall therefore ignore Z_t.

Computationally, (2) [and (3)] often lead to numeric difficulties, which can be avoided by using logarithms. Let β_t = log(B_t) = log(odds for $\underline{H}_1$), η_t = log(H_t) = log(hazard odds), and ζ_t = log(z_t) = log[likelihood ratio(t)] – log(1 – h_t) = "adjusted log(likelihood ratio)". Then (2) can be rewritten as

$$\beta_t = \eta_t + \log[1 + (z_t B_{t-1} / H_t)\ ],$$

$$= \eta_t + \log[1 + \exp(\Delta_t)], \tag{4}$$

where $\Delta_t = \zeta_t + \beta_{t-1} - \eta_t$. But ζ_t exceeds log[likelihood ratio(t)] by [– log(1 – h_t)] > 0 as long as $0 < h_t < 1$. In most practical applications, h_t is so small that $\log(1 - h_t) \cong (-h_t) \cong 0$, which makes ζ_t essentially the log(likelihood ratio).

To relate (4) to something more concrete, note that if the observations $(y_t|\underline{H}_i) \sim N(\mu_i, \sigma^2)$, $i = 0, 1$, then

$$\zeta_t = (-1)^{i+1} d\,\frac{\left(y_t - \bar{\mu}\right)}{\sigma} - \log\left(1-h_t\right) \sim N\left[(-1)^{i+1} 0.5\,d^2 - \log\left(1-h_t\right),\ d^2\right], \tag{5}$$

where $d = \left(\mu_1 - \mu_0\right)/\sigma$, and $\bar{\mu} = \left(\mu_0 + \mu_1\right)/2$. We return to this case later.

An alternative to (4) can be obtained by factoring ($z_t B_{t-1}$) out of (2) to produce the following:

$$\beta_t = \zeta_t + \beta_{t-1} + \log[1 + \exp(-\Delta_t)]. \tag{6}$$

If we use (4) when $0 > \Delta_t = (-|\Delta_t|)$ and (6) when $0 \leq \Delta_t = |\Delta_t|$, we get

$$\beta_t = \max\{\ \eta_t,\ \zeta_t + \beta_{t-1}\ \} + \log[1 + \exp(-|\Delta_t|)]. \tag{7}$$

For any Δ_t,

$$0 < \log[1 + \exp(-|\Delta_t|)] \leq \log(2).$$

Except when $|\Delta_t|$ is small, this term will be negligible. In that case, (7) is a cumulative sum with a floor at η_t and with β_0 = the initial prior log odds for $\underline{H}_1$ being a fast initial response (FIR) parameter, as mentioned following (1). If the data are only marginally informative, $|\Delta_t|$ will tend to be small under $\underline{H}_0$, but getting larger after the transition. If the data are nearly always highly informative, $|\Delta_t|$ will tend to be large even under $\underline{H}_0$. In this latter case, the "Bayes-adjustment term" $\log[1 + \exp(-|\Delta_t|)]$ will nearly always be negligible.

Expression (7) can be written in a more familiar form by letting $Q_t^* = \beta_t - \eta_t$ = the excess over the log hazard odds of the log odds for $\underline{H}_1$. However, with non-constant hazard, Q_t^* seems to lose its utility. In such cases, we will focus on β_t, as in the example discussed with Figure 6 below. Before considering that case, however, we will explore the simplifications obtainable from using Q_t^* with constant hazard, $h_t = h$, so $Q_t^* = \beta_t - \eta$.

Subtracting η from both sides of (7), we get the following:

$$Q_t^* = \max\{\ 0,\ Q_{t-1}^* + \zeta_t\ \} + \log[1 + \exp(-|Q_{t-1}^* + \zeta_t|)]. \tag{8}$$

As mentioned above, the term log[1 + exp(–| Q^*_{t-1} + ζ_t |)] will be negligible except when |Δ_t| is small. If we drop it from (8), and replace Q^*_t with Q^+_t to signify the change, we get

$$Q^+_t = \max\{\ 0,\ Q^+_{t-1} + \zeta_t\ \}. \tag{9}$$

This is essentially the standard one-sided Cusum of log(likelihood ratio) due to Page (1954) except for the "log(1 – h_t)" adjustment included in the definition of ζ_t just before (4).

Figure 2 presents a typical simulation comparing (8) and (9) with 20 observations changing from $\underline{H}_0$ to $\underline{H}_1$ at t = 11. In this, we assume h = 0.001, $\underline{H}_i$: $y_t \sim N(\mu_i,\ \sigma^2)$, σ = 1, i = 0, 1, μ_0 = 0, and μ_1 = 1. The log(likelihood ratio) in this case is given by (5). This makes Q^+_t a standard one-sided Cusum except for the "adjustment" $[-\log(1-h)]$ to the log(likelihood ratio) in ζ_t. Since h_t = 0.001, this adjustment is $[-\log(1-h)] \cong 0.001$, which is not visually detectable in Figure 2.

(Figure 2 about here)

Four vertical scales are provided in Figure 2. The first is the "natural" Cusum scale for Q^*_t and Q^+_t, starting at 0. But Q^*_t is the excess in the log odds for $\underline{H}_1$ (β_t) over the hazard odds (η), recalling the definition of Q^*_t preceding (8). We use this to obtain the second scale, per β_t = Q^*_t + η, where $\eta = \log[h/(1-h)] = \log(0.001/0.999) = (-6.91)$; the log(hazard odds) is a floor for the log odds for the transition. The third and fourth scales in Figure 2 simply transform the log odds for $\underline{H}_1$ into odds, B_t, and probability, (1 – g_t).

The behavior here is typical of what we have seen in other simulations with different values of μ_1 and differing numbers of observations before and after the change: Q_t^* and Q_t^+ tend to go up and down together, possibly after an initial drift to an fairly predictable bias generated by the Bayes-adjustment term $\log\left[1+\exp\left(-\left|Q_{t-1}^*+\zeta_t\right|\right)\right]$ in (8). When the plant is good, the movements in Q_t^* and Q_t^+ are nearly identical (except possibly for an initial drift); the similarities only increase after the plant becomes bad. The effect of the Bayes-adjustment term is displayed graphically in Figure 3.

(Figure 3 about here)

For the normal example introduced with (5), $E\zeta_t$ = $(-1)^{i+1}0.5d^2-\log(1-h)$ under $\underline{H}_i$, i = 0, 1, with d = $(\mu_1-\mu_0)/\sigma$. Thus, after the transition (under $\underline{H}_1$), the traditional Cusum Q_t^+ increases on average $0.5d^2/\sigma^2-\log(1-h)$ for each observation; meanwhile, the Bayes-adjusted Cusum Q_t^* increases slightly faster initially but converges quickly to this growth rate. In the case considered in Figure 3, this asymptote was for practical purposes reached in one observation.

To understand better the bias apparent in Figure 2, we computed the means of Monte Carlo trajectories for Q_t^* and Q_t^+ and their difference for a variety of combinations of d and h. Under $\underline{H}_0$, all began with a drift towards an asymptote, $E(Q_\infty^*|\underline{H}_0)$ or $E(Q_\infty^+|\underline{H}_0)$. We found that the logarithms of these two asymptotes were approximately linear in θ_0 = log(–δ_0), where δ_0 = $E_0(\zeta_t/d)$ = $-0.5d-\ln(1-h)/d$ = expected increment rescaled to unit variance under $\underline{H}_0$ (assuming d > 0). Analysis of the means of 20,000 simulated trajectories at each of 17 different levels for *d* between 0.15 to 3 and with *h* positive but computationally 0 produced the following:

$$\log(Q_\infty^* / d | \underline{H}_0) = -0.367 - 1.290\,\theta_0 - 0.351\theta_0^2 - 0.125\theta_0^3 - 0.017\,\theta_0^4 + e\,,$$
$$\log(Q_\infty^+ / d | \underline{H}_0) = -2.083 - 2.634\,\theta_0 - 0.993\theta_0^2 - 0.341\theta_0^3 - 0.049\,\theta_0^4 + e\,, \qquad (10)$$

where $\theta_0 = \log(-\delta_0)$.

The simulated differences $E[(Q_t^* - Q_t^+) | \underline{H}_i]$, $i = 0, 1$, clearly approached asymptotes with increasing *t*. Figure 4 presents this bias for $t = 500$ under $\underline{H}_0$ and for $t = 250$ after an abrupt jump to $\underline{H}_1$. The following was fit to these data under $\underline{H}_0$:

$$E\{\log\left[\left(Q_t^* - Q_t^+\right)/d\right] | \underline{H}_0\} =$$
$$0.608 + 1.584\,\delta_0 + 0.766\,\delta_0^2 + 0.470\,\delta_0^3 + 0.114\,\delta_0^4\,. \qquad (11a)$$

(Figure 4 about here)

Under $\underline{H}_1$, the three sets of simulations with $d < 0.2$ had not quite converged during the period simulated after the transition (250 independent observations), so they were excluded from the fitting process. The following model was fit to the remaining 14 points:

$$E\{\log\left[\left(Q_t^* - Q_t^+\right)/d\right] | \underline{H}_1\} = 0.664 - 0.717\,\delta_1 - 0.037\,\delta_1^2\,. \qquad (11b)$$

Expressions (11) can help us in various ways, e.g., to modify threshold selection procedures developed for Page's one-sided Cusum so they can be used with the Bayes' Cusum. Expressions (10) and (11) should be considered preliminary, as some of the simulations upon which they are based exhibited a subtle non-Markovian behavior that we could not explain; see the Appendix.

Before discussing threshold selection in more detail, we consider briefly how this analysis changes if the hazard rate is zero: First, if the hazard is constant, it disappears from (3). If it is zero, it disappears from (2), and then the logarithm of (2) becomes the log-odds formulation of Bayes' theorem:

$$\beta_t = \zeta_t + \beta_{t-1}. \tag{12}$$

In words, the posterior log odds (β_t) is the prior log odds (β_{t-1}) plus the log likelihood ratio (ζ_t). This is a traditional two-sided Cusum. It is clearly the correct answer if we are testing to evaluate an unchanging property of nature, as noted by Wald (1947).

However, this is different from the Girshick-Rubin criterion (3) without the hazard factor. To understand this, note that zero hazard turns $\left(H_{t-1}/H_t\right)$ in (3) into the indeterminate form (0/0) in (3) while also turning η_t into $(-\infty)$ and introducing either $(\infty - \infty)$ or similar nonsense into (4) and (6)-(8).

But monitoring applications look for changes, which imply a non-zero hazard. In such situations, a traditional two-sided Cusum, (12), is inappropriate except when used with a time-varying threshold such as a traditional V-mask. But a two-sided Cusum is mathematically and in usage equivalent to two one-sided Cusums, where the slope of the "V" in the V-mask is related to the difference in expectations between $\underline{H}_0$ and $\underline{H}_1$ of the log(likelihood ratio). As previously noted, the zero level for Page's one-sided Cusum is provided by the log(hazard odds), and with zero hazard, this floor is at negative infinity.

In sum, if the hazard rate is truly zero, a two-sided Cusum performing a Bayesian formulation of a Wald sequential test (12) is appropriate. Meanwhile a non-zero but constant and small hazard rate calls for a monitor that is virtually equivalent to a one-sided Cusum (9).

In sum, the comparison in this section of the Bayes-adjusted Cusum and Page's one-sided Cusum analysis helps us understand previous reports, e.g., by Srivastava and Wu (1993, p. 665), who said that with low thresholds and high false alarm rates, the Bayes' procedure was better than the traditional Cusum. With higher thresholds, Srivastava and Wu agreed with the earlier simulation comparison by Roberts (1966), who found the Cusum and the Bayesian approaches essentially equivalent.

We next consider selection and interpretation of a detection threshold.

3. CUSUM THRESHOLD AND INCREASE IN POSTERIOR LOG(ODDS)

An obvious decision criterion for Bayesian monitoring is to set an alarm when the posterior probability of $\underline{H}_1$ exceeds a threshold, which may be tied to the economics of the problem (e.g., Girshick and Rubin 1952). This translates into a threshold for the posterior log odds for $\underline{H}_1$, β_t. With constant hazard, this is equivalent to setting a threshold for our Bayes-adjusted Cusum, $Q_t^* = \beta_t - \eta$; this threshold on Q_t^* becomes the "increase" in log odds for a bad plant (β_t) over the log hazard odds (η) required to set an alarm. This equivalence is illustrated in Table 1 for a constant hazard rate of 0.01. To deepen our understanding of this connection, suppose we are collecting one sample per day from a sewage treatment plant and preparing a Cusum chart of the result. And suppose that $\zeta_t = \log(f_{1,t}/f_{0,t}) - \log(1 - h)$ in (8) has standard deviation of 1 and mean (-0.5) if the system is good and $(+0.5)$ if bad. Suppose also that the plant has an upset (goes bad, from $\underline{H}_0$ to $\underline{H}_1$) roughly once every 100 days, which means it has roughly a constant hazard rate of 0.01. Then from Table 1, we see that a threshold for Q_t^* of 4 is roughly equivalent to deciding to declare an upset when the posterior probability of $\underline{H}_1$ is 0.36 or greater [ignoring the bias from the Bayes-adjustment term]. If the model is correct, then this 0.36 reflects the proportion of plants with comparable histories that are bad; it is not (merely) a subjective probability.

(Table 1 about here)

By comparison, a standard Cusum chart might put the threshold for Q_t^+ at 4, which would have an Average Run Length (ARL) of roughly 350 to a false alarm and 8.5 to a valid alarm (e.g., Bissell 1969). Ignoring the bias modeled by (11), this is roughly equivalent to declaring an upset when the posterior probability of $\underline{H}_1$ is 0.36. In a real application, we would want to adjust these numbers to consider the bias per (11). However, the point here is the conceptual equivalence, so we will not bother now with this refinement.

4. COSTS AND RUN LENGTHS

In many applications, the expected cost of a delay to detection will be proportional to the average run length (ARL) to an alarm after the plant (i.e., the system being monitored) transitions from $\underline{H}_0$ to $\underline{H}_1$. Meanwhile, the expected cost of a false alarm might be proportional to the probability of an alarm under $\underline{H}_0$. Obviously, increasing the threshold increases the ARL (under $\underline{H}_1$) while reducing the false alarm rate. Therefore, selecting a threshold implies a certain assessment of the cost of a false alarm relative to the cost of a delay of one more observation. This gives us three equivalent ways to select a threshold for a Cusum / Bayes' monitor:

(*a*) Select a posterior probability [or log(odds)] for $\underline{H}_1$ above which an malfunction is declared.

(*b*) Select a threshold to balance some characteristics of the run length distributions for good and bad systems.

(*c*) Specify the cost of a false alarm relative to the cost of waiting one more observation before declaring a transition to $\underline{H}_1$.

In practical monitor design, it may be wise to evaluate all three perspectives before making the final choice of threshold.

This three-part equivalence assumes the model is correct, which often is not the case. For example, a monitor may be designed ignoring serial dependencies in the data, because the application may not justify the additional expense of modeling them. The final evaluation and selection of a threshold might be made by extrapolating from run length data collected using artificially low thresholds in prototype systems, as suggested by Bisgaard et al. (2002). These thresholds would automatically adjust for model inadequacies such as serial dependence. Meanwhile, the threshold implied by this equivalence for the posterior probability of $\underline{H}_1$, ignoring serial dependence and model inadequacies, might be ridiculously close to 1; in such cases, the formally computed posterior is not a realistic assessment of the relative frequency of plants with comparable histories that are bad. This does not negate the value of the monitor, only the posterior probability interpretation of it.

5. AN EXAMPLE WITH NON-IDENTICALLY DISTRIBUTED INCREMENTS AND NON-CONSTANT HAZARD

Box et al. (2000) outline an "8-step process" for developing an on-board diagnostic (OBD) to detect malfunctions in emissions controls required on all new automobiles sold today in the US, Canada and Europe. This process begins with defining "good" and "bad", collecting data on each, and building models so $\underline{H}_0$ and $\underline{H}_1$ can be defined. This "8-step process" is now routinely taught in workshops for engineers organized by the Society of Automotive Engineers. A standard example in these workshops is a "defective paper helicopter", i.e., one with excessive wing loss. Paper helicopters have a long history as a tool for teaching experimental design (Box 1992), including as an illustration of sequential experimentation in a response surface investigation (Box and Liu 1999).

The problem of designing an OBD to detect excessive wing loss becomes difficult (and similar to more substantive applications) by assuming that the wing area cannot easily be measured directly but must instead be inferred from the fall time from a known height with a known "passenger load" (number of staples in the body of the paper helicopter). The automotive engineers in our SAE workshop assure us that this is quite similar to their applications, except that paper helicopters can be tested in minutes with a miniscule budget rather than months and many thousands of dollars to build and test prototypes.

This "plant" produces noisy data whose distribution depends on operating conditions (number of staples and fall height) as well as the condition of the plant (percent wing loss). The resulting diagnostic incorporates non-identically distributed increments, which are similar to what engineers might want to use to diagnose, for example, misfires from deficits in angular acceleration, where the deficit depends on the fuel load, and the noise level seems to be related to the engine speed (Graves, Bisgaard and Kulahci 2002b).

Our "good" paper helicopter is 2.54 by 10.16 cm. (1 by 4 inches) with (*a*) the wings occupying the top half and (*b*) the bottom half of the body folded vertically in thirds where staples are affixed. "Deterioration" is simulated by removing between 0 and 100% of the wings. Data on which to build a diagnostic were collected from a 3 x 3 experiment with (0, 2, and 4) staples by (0, 50, and 100%) wing loss, augmented by a 2 x 2 with (1 and 3) stapes by (25 and 75%) wing loss. This allowed us to estimate cubics and quartics along with a full parabolic and some higher order interactions. This design allowed us to look for nonlinearity, suspected from earlier tests conducted with workshops held from Turin, Italy, to Pasadena, California.

Two paper helicopters for each of these 13 designs were prepared and dropped twice from three different heights, 2.53, 4.06, and 5.96 m. with the fall time being recorded. In addition to these 156 "calibration" drops, we also completed prototypes so we had a full factorial in (0, 20, 25, 40, 50, 60, 75, 80, 100%) wing loss by (0, 1, 2, 3, 4) staples, which were all dropped once after the "calibration" drops at each height.

The "calibration" data set was used to estimate a regression model, whose results were incorporated into a diagnostic, which was evaluated using the "confirmation" data set. Selected predictions from the model with standard 95% confidence limits appear in Figure 5. The local minimum appearing in Figure 5 between 85 and 100% wing loss may to us possible lack of fit. However, we think the local maximum appearing between 0 and 30% wing loss is probably real. Most importantly, the model predictions at 40 and 60% wing loss are likely sufficiently accurate for present purposes. The 40 and 60% numbers were selected as "worst acceptable" ($\underline{H}_0$) and "best unacceptable" ($\underline{H}_1$; see Box et al. 2000). A competent engineer might use a model like this and ignore the apparent lack of fit unless there was strong justification for the additional investment required to develop a better model.

(Figure 5 about here)

Figure 6 presents two Bayes-adjusted Cusums based on this model: one with constant hazard and the other with a discrete Weibull. A *.zip file downloadable from "www.prodsyse.com" contains a drawing of the paper helicopters used for this study plus an Excel file with the experimental data used to estimate the model fits summarized in Figure 5 and an additional 131 observations collected with gradually increasing levels of wing loss; this Excel file also includes the Cusum computations for Figure 6. In these additional observations, wing loss jumps from 25 to 40% at observation 45, then to 50% at observation 66, to 60% at observation 78, and 75% at observation 93. The first 100 of these 131 additional observations are plotted in Figure 6. The last 31 observations, not plotted, continue this steep upward trend.

(Figure 6 about here)

For Figure 6, we assumed h_t = $\left[S(t)-S(t+1)\right]/S(t)$ = $1-\exp\left\{\left[(t+1)/\tau\right]^c-\left(t/\tau\right)^c\right\}$ with characteristic life τ = 100 and with shape c = 1 and 3 for the constant hazard and the increasing failure rate lines. This is the discrete hazard rate for the Weibull distribution (Lawless 1982, p. 10; Kalbfleisch and Prentice 1980, pp. 35-36; Salvia and Bollinger 1982; it is different from the "discrete analogue to the

Weibull" described by Padgett and Spurrier 1985). In discrete applications such as the present, the hazard rate is actually a probability; continuous hazards must be converted to probabilities as in this example.

Figure 6 is labeled with 3 vertical axes. The scale is linear in the axis on the left, the "naive log odds for bad", which is translated into nonlinear "naive odds for bad" and "naive Pr(Bad)" on the right [ignoring the bias (11)]. We say "naive", because the model assumes an abrupt jump from 40 to 60% wing loss, but we are applying it to data collected with a gradual drift in wing loss from 0 to 75%. The lines drawn would represent probabilities in the relative frequency sense if the points plotted actually represented a plant as modeled, beginning with 40% wing loss changing abruptly to 60% at a random point according to a probability distribution whose discrete hazard is the floor under Bayes-adjusted Cusum. Since the malfunction mechanism here was not the single abrupt jump modeled, the computed "posterior log odds, odds and probabilities" are labeled "naive" as a reminder of the violation of the assumptions. However, for many purposes, we would expect the Bayes-adjusted Cusum to be as robust as the traditional one-sided Cusum (see Box and Luceño 1997), though the posterior probability interpretation may not be as robust.

Both the constant and Weibull ($c = 3$) lines in Figure 6 are assumed to start with a prior log odds for bad of 0, simulating a cheap repair that is only effective half the time. This shows clearly how the "head start" or "fast initial response" of Lucas and Crossier (1982) is related to the prior log odds for the plant being "bad".

The Bayes-adjusted Cusum seems to linger around 0 until observations 14 and 15 pull it down to its hazard rate floor, consistent with expression (7). At that point, the hazard rate line for Weibull (c = 3) is substantially below that for constant hazard with the same characteristic life, so it makes it easy to see the difference. From there, the Bayes-adjusted Cusum tends to mostly follow the hazard rate until a change in the condition of the plant starts pulling it upwards fairly consistently, as we see in Figure 6.

6. IF THE $\underline{H}_1$ DISTRIBUTION DEPENDS ON THE CHANGEPOINT

While the present discussion has focused on simple hypotheses, the derivation here works with incompletely specified hypotheses, cusumming the log of the ratio of the predictive (marginal) densities with parameters estimated sequentially as in Kalman filtering, provided that (*a*) $f_{1,t}$ does not depend on t_0, and (*b*) we can ignore the question of when we can start accumulating information about unknown parameters of $f_{1,t}$.

For such situations, West and Harrison (1999) recommend "Bayes' factors", whose logarithms are the one-sided Cusums just described [without the Bayes' adjustment, like (9)]. Eide and Maybeck (1996) and Menke and Maybeck (1995) call similar techniques "Multiple Model Adaptive Estimation"; see also Graves et al. (2001, sec. 7).

For a situation where the first condition is not met, consider tool wear with t_0 = the time at which the hardened surface of a tool is worn through, after which the tool wears much faster. In such case, the posterior cannot be summarized in one number such as g_t in (1). However, Bayesian sequential updating might still provides a theoretical best procedure, evaluated using, e.g., Markov Chain Monte Carlo (MCMC; see, e.g., Roberts and Casella 1999 or Carlin and Louis 1998). Such evaluations could help us evaluate the adequacy of simpler methods.

A superficially attractive alternative is provided by generalized likelihood ratio (GLR) monitors studied by Lai (1995), Basseville and Nikivorov (1993) and others. These procedures compute cumulative sums of ratios of fragments of likelihood maximized over ($\underline{H}_0 \cup \underline{H}_1$) vs. $\underline{H}_0$, similar to generalized likelihood ratio procedures that have been primary tools of frequentist statistical methods since the 1928-1938 pioneering work of Neyman and Pearson (1966; see also Lehmann 1986). Recent simulation work by Chang and Fricker (1999) discovered that Cusums and exponentially weighted moving averages (EWMAs) "perform surprisingly well compared to the GLR test, usually outperforming it." These results invite the speculation that GLR procedures may lose power by optimizing over implausible alternatives. When history is relevant to predicting the future, we might expect that reasonable Bayesian procedures might outperform the excessive conservatism of minimax-type GLR procedures.

7. DISCUSSION

In this article, we looked for an abrupt jump from $\underline{H}_0$ to $\underline{H}_1$. We found that when the hazard rate is low and relatively constant, Bayesian sequential updating is reasonably well approximated by a Cusum of log(likelihood ratio); for a recent overview of the traditional Cusum literature, see Hawkins and Olwell (1998). If the hazard rate is not constant (i.e., the distribution of time to a problem is not exponential), then the theory presented here provides a natural way to improve diagnostic performance through the use of that information. This could be quite valuable for designing preventive maintenance procedures that combine periodic data collection with reliability models or for biostatistical protocols that combine current observations for changes in a patient's condition with hazard rates estimated from other patients with comparable conditions and superficially similar therapies. This holds promise for developing procedures that outperform any procedure that considers only one source of information.

The literature on preventive maintenance includes various techniques for using reliability information (e.g., Kardon and Fredendall 2002). Gertsbakh (1977) describes how to determine optimal inspection intervals for preventive maintenance. However, we know of no previously published techniques for combining noisy data from imperfect inspections with reliability information; a Bayes-adjusted Cusum provides, we believe, a useful way to address this common situation. We believe that it is simple enough, especially with modern computers, that it could become part of routine programs for "Overall Equipment Effectiveness (OEE)" and "Total Productive Maintenance (TPM)" (Productivity Development Team 1999).

If instead of an abrupt jump from $\underline{H}_0$ to $\underline{H}_1$, we assume a normal random walk in a state space observed with normal error, Bayesian sequential updating produces Kalman filtering (Graves et al. 2001), a special case of which is a Bayesian exponentially weighted moving average (EWMA), whose design combines reliability information with measurement noise models based on studies of gage repeatability and reproducibility (Graves, Bisgaard and Kulahci 2002a).

When looking for abrupt jumps between hypotheses $\underline{H}_0$ and $\underline{H}_1$ where at least one is incompletely specified, the obvious solution of running parallel sequential Bayes procedures (e.g., Kalman filters) has been tried with apparent success, as discussed in Section 6 above.

With Cusum increments that are not identically distributed, such as the paper helicopters of Section 5 with varying "passenger loads" and drop heights, an acceptable monitor should perform well over a range of usage patterns. Variations in the usage pattern complicate the design problem: Environmental protection agencies might consider noncompliant an OBD that did not function properly for drivers who never drove over 40 kilometers per hour (25 miles per hour), even if it performed perfectly for the other 99% of the population.

Engineers could deal with this problem by considering different usage patterns, e.g., distributions of drop heights and passenger loads. They would then try to select thresholds that provide a sufficiently quick response with an acceptable false alarm rate under all scenarios. Many applications such as legally mandated OBDs on emissions controls in automobiles mandate a maximum delay from the onset of a "worst acceptable" condition to an alarm. The image of Figure 5 suggests that an acceptable diagnostic would be much easier to achieve with many staples and a high drop height than with no staples and a low drop height: The difference between "good" and "bad" is large in the first case and small in the second. Any threshold low enough to trigger an alarm sufficiently quickly with repeated short drops and zero staples might have an unacceptable false alarm rate for some usage pattern.

In sum, we believe we have described a new tool that can improve solutions for a variety of problems while providing a Bayesian interpretation for traditional Cusum techniques. In conjunction with related work described, e.g, by Graves et al. (2001) and Graves, Bisgaard and Kulahci (2002a, b), we believe this helps to establish the 2-step

Bayesian sequential updating cycle of Figure 1 as a general foundational principle for solving monitoring problems.

APPENDIX

No precision estimates are given with (10) and (11) because we saw a non-Markovian behavior in the supporting simulations that we could not explain. Since modeling these asymptotes and the bias is tangential to the main thrust of this article, we briefly describe our concerns in this appendix and leave their resolution to future research.

Specifically, for certain values of θ_0, the averages of large numbers of simulated trajectories using both the Bayes-Adjusted (8) and Page's one-side (9) Cusum starting from the asymptotes in (10) initially fell quickly then climbed gradually back to the models (10). Several efforts were made to identify possible bugs or non-random behavior of the pseudo-random number generators, to no avail. The simulations used in (10) and (11) were done in S-Plus 6.1 but were repeated in R 1.9.1 (downloadable from www.r-project.org) using two different pseudo-random number generators different from each other and from the one used by S-Plus. Similar simulations were done in Microsoft Excel with result that suggested the same non-Markovian behavior but were not conclusive due to the smaller numbers of simulations run. A final check included a completely independent programming effort that reproduced the aberrant behavior for the Page Cusum (9) in 16 lines of R code given below.

Even with the unexplained non-Markovian behavior, it seems that (10) and (11) display the correct behavior qualitatively even if their numerical precision is less than what might be desired. Ignoring this question, the standard deviations of residuals for the expressions in (10) and (11) were s_e = 0.004, 0.010, 0.003, and 0.002, respectively; all coefficients were highly significant, with the largest of the individual p-values being 0.009.

This phenomenon can be replicated easily in S-Plus or R using the code for Page's Cusum (9) given below. This code could easily be modified to simulate the jump to $\underline{H}_1$, consider the hazard rate explicitly and also compute Qs.t.d = $\left(Q_t^*/d\right)$ and the bias from the same random numbers; these modifications are not included here as they make it slightly more difficult to see the curious non-Markovian behavior:

```
simCus5 <- function(Ezt.d=-0.1, Qp0.d=3, maxTime=400, nSims=20000){
  Qp.d.mean <- rep(NA, maxTime)
  Qp.t.d <- rep(Page0.d, nSims)
  for(i in 1:maxTime){
    z.t <- (Ezt.d + rnorm(nSims))
    Qp.t.d <- pmax(0, Qp.t+z.t)
    Qp.d.mean[i] <- mean(Qp.t)
  }
  Qp.d.mean
}
plot(simCus5(Qp0=4.5))
```

ACKNOWLEDGEMENTS

This research was supported by the Low Emissions Partnership of Ford, DaimlerChrysler and General Motors. The authors wish to thank John Van Gilder and Thomas Ting of General Motors, Ken Marko and John James of Ford, and Hal Zatorski and Cuiping Wu of DaimlerChrysler for numerous fruitful discussions on various aspects of OBD development and George Cheroff of PDF Solutions for valuable assistance with library research. Anonymous referees also provided valuable suggestions.

REFERENCES

Basseville, M., and Nikiforov, I. V. (1993) *Detection of Abrupt Changes: Theory and Applications* (Englewood Cliffs, NJ: Prentice-Hall).

Bisgaard, S., Graves, S., Kulahci, M., Van Gilder, J., James, J., Marko, K., Ting, T., Wu, C., and Zatorski, H. (2002) "Accelerated Testing of On-Board Diagnostics", technical report available from "www.prodsyse.com".

Bissell, A. F. (1969) "Cusum Techniques for Quality Control", *Applied Statistics*, 18: 1-30.

Box, G. E. P. (1992) "Teaching Engineers Experimental Design with a Paper Helicopter", *Quality Engineering*, 4: 453-459, reprinted as ch. B.2 in Tiao et al. (2000).

Box, G., Graves, S., Bisgaard, S., Van Gilder, J., Marko, K., James, J., Seifer, M., Poublon, M., and Fodale, F. (2000) "Detecting Malfunctions in Dynamic Systems", *Proceedings of the 2000 SAE World Congress & Exposition* (SAE Technical Paper Series number 2000-01-0363).

Box, G. E. P., and Liu, P. (1999) "Statistics as a Catalyst to Learning by Scientific Method, Part I - An Example", *Journal of Quality Technology*, 31: 1-15, reprinted as ch. C.1 in Tiao et al. (2000).

Box, G., and Luceño, A. (1997) *Statistical Control by Monitoring and Feedback Adjustment*, NY: Wiley.

CARB (1997) "Malfunction and Diagnostic System Requirements--1994 and Subsequent Model-Year Passenger Cars, Light-Duty Trucks, and Medium-Duty Vehicles and Engines" (OBD II), with modifications effective as of September 25, 1997, sec. 1986.1 of Title 13,California Code of Regulations. Sacramento, CA: California Air Resources Board; available from "http://www.arb.ca.gov/msprog/obdprog/obdprog.htm".

Carlin, B. P., and Louis, T. A. (1998) *Bayes and Empirical Bayes Methods for Data Analysis* (NY: Chapman and Hall)

Chang, J. T., and Fricker, R. D., Jr. (1999) "Detecting When a Monotonically Increasing Mean has Crossed a Threshold", *Journal of Quality Technology*, 31: 217-234.

Eide, P. K., and Maybeck, P. S. (1996) "An MMAE Failure Detection System for the F-16", *IEEE Transactions on Aerospace and Electronic Systems*, 32: 1125-1136.

Gertsbakh, I. B. (1977) *Models of Preventive Maintenance* (NY: North-Holland).

Girshick, M. A., and Rubin, H. (1952) "A Bayes Approach to a Quality Control Model", *Annals of Mathematical Statistics*, 23: 114-125.

Graves, S. B. Bisgaard, S., Kulahci, M., Van Gilder, J., Ting, T., Marko, K., James, J., Zatorski, H., and Wu, C. (2001) *Foundations of Monitoring Dynamic Systems*, technical report available from "www.prodsyse.com".

Graves, S. B. Bisgaard, S., and Kulahci, M. (2002a) "Designing Bayesian EWMA Monitors Using Gage R & R and Reliability Data", technical report available from "www.prodsyse.com".

__________ (2002b) "A Bayesian EWMA for Mean and Variance", technical report available from "www.prodsyse.com".

Gunderson, B. (2000) "Monitoring Methods in Implantable Medical Devices", invited presentation at the Spring Research Conference, Seattle, WA, June 27, 2000.

Hawkins, D. M., and Olwell, D. H. (1998) *Cumulative Sum Charts and Charting for Quality Improvement* (NY: Springer).

Kalbfleisch, J. D., and Prentice, R. L. (1980) *The Statistical Analysis of Failure Time Data* (NY: Wiley).

Kardon, B., and Fredendall, L. D. (2002) "Incorporating Overall Probability of System Failure into a Preventive Maintenance Model for a Serial System", *Journal of Quality in Maintenance*, 8: 331-345.

Kenett, R., and Zacks, S. (1998) *Modern Industrial Statistics: Design and Control of Quality and Reliability* (San Francisco, CA: Duxbury).

Lai, T. L. (1995) "Sequential Changepoint Detection in Quality Control and Dynamical Systems", *Journal of the Royal Statistical Society, Ser. B*, 57: 613-658.

Lawless, J. F. (1982) *Statistical Models and Methods for Lifetime Data* (NY: Wiley)

Lehmann, E. L. (1986) *Testing Statistical Hypotheses*, 2nd ed. (NY: Wiley)

Lucas, J. M., and R. B. Crossier (1982) "Fast Initial Response for CUSUM Quality-Control Schemes", *Technometrics*, 24: 199-205.

Menke, T. E., and Maybeck, P. S. (1995) "Sensor / Actuator Failure in the VISTA F-16 by Multiple Model Adaptive Estimation", *IEEE Transactions on Aerospace and Electronic Systems*, 31: 1218-1229.

Mondt, R. J. (2000) *Cleaner Cars* (Warrendale, PA: Society of Automotive Engineers).

Neyman, J., and Pearson, E. S. (1966) *Joint Statistical Papers of J. Neyman & E. S. Pearson* (Berkeley, CA: University of California Press).

Padgett, W. J., and Spurrier, J. D. (1985) "On Discrete Failure Models", *IEEE Transactions on Reliability*, R-34, 253-256.

Page, E. S. (1954) "Continuous Inspection Schemes", *Biometrika*, 41, 100-115.

Productivity Development Team (1999) *OEE for Operators* (Portland, OR: Productivity).

Roberts, C. P., and Casella, G. (1999) *Monte Carlo Statistical Methods* (NY: Springer)

Roberts, S. W. (1966) "A Comparison of Some Control Chart Procedures", *Technometrics*, 8: 411-430.

Salvia, A. A., and Bollinger, R. C. (1982) "On Discrete Hazard Functions", *IEEE Transactions on Reliability*, R-31, 458-459.

Shiryaev, A. N. (1963) "On Optimal Methods in Quickest Detection Problems", *Theory of Probability and Its Applications*, 8, 22-46.

Srivastava, M. S., and Wu, Y. (1993) "Comparison of EWMA, Cusum, and Shiryayev-Roberts Procedures for Detecting a Shift in the Mean", *Annals of Statistics*, 21: 645-670.

Steiner, S. H., Cook, R. J., Farewell, V. T., and Treasure, T. (2000) "Monitoring Surgical Performance Using Risk Adjusted Cumulative Sum Charts", *Biostatistics*, 1: 441-452.

Tiao, G. C., Bisgaard, S., Hill, W. J., Peña, D., and Stigler, S. M. (2000) *Box on Quality and Discovery* (NY: Wiley).

Wald, A. (1947) *Sequential Analysis*, NY: Wiley; reprinted 1973 by Dover.

West, M., and Harrison, J. (1999) *Bayesian Forecasting and Dynamic Models*, 2nd ed., 2nd corrected printing (NY: Springer)

Figure 1. Bayesian Sequential Updating

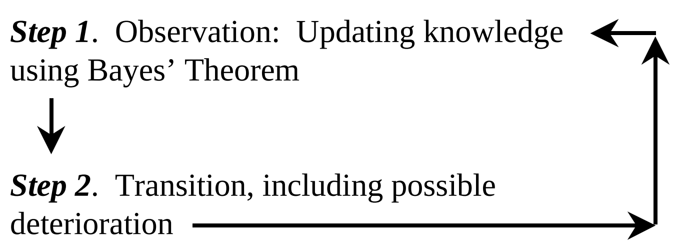

Figure 2. One-Sided and a Bayes-Adjusted Cusum Simulations

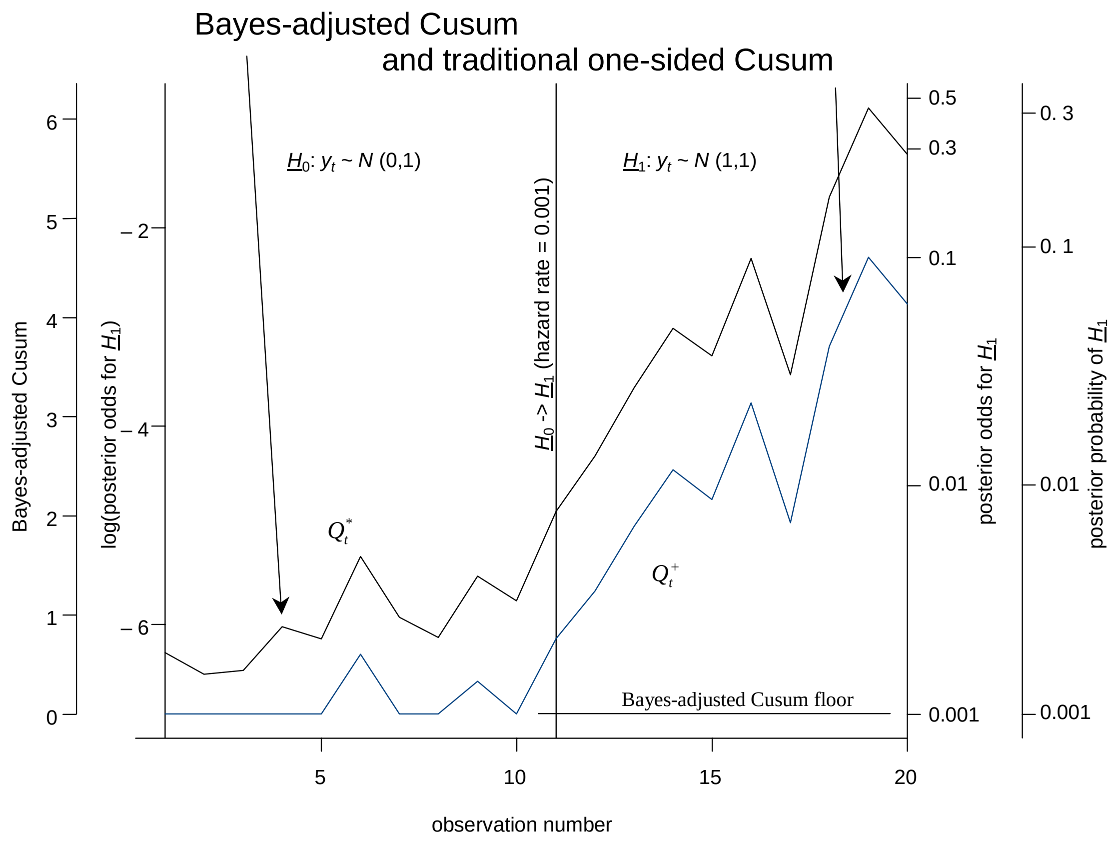

Figure 3. Bayes-Adjusted and Traditional Cusum Iteration

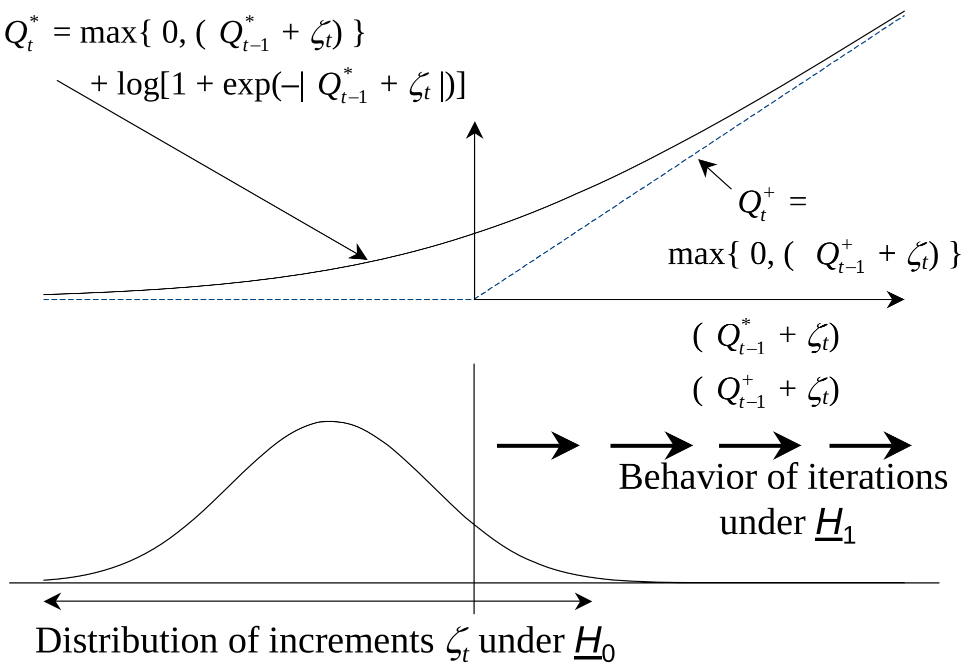

Figure 4. Steady-State Excess of Bayes-Adjusted over Page's One-Sided Cusum
(mean of 20,000 simulated series for each of 17 scenarios; see text)

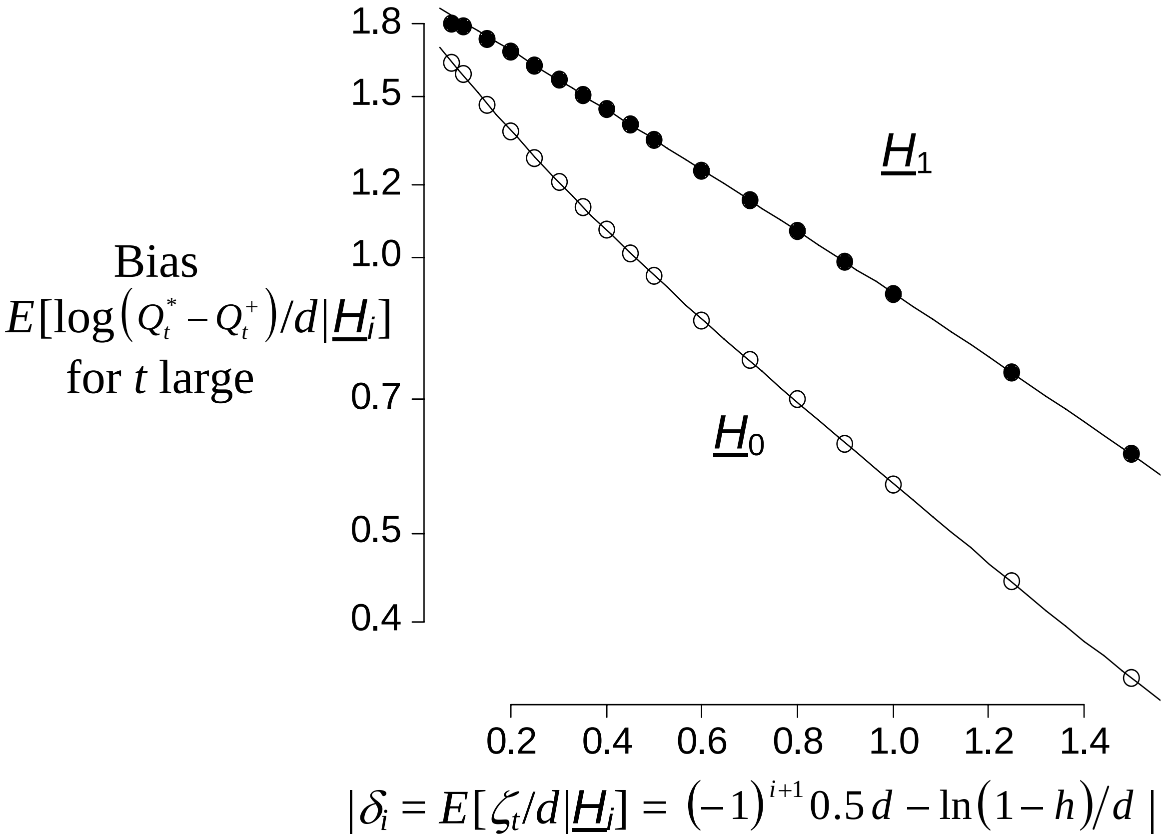

Figure 5. Predicted Fall Time vs. Wing Loss, Height, and Number of Staples

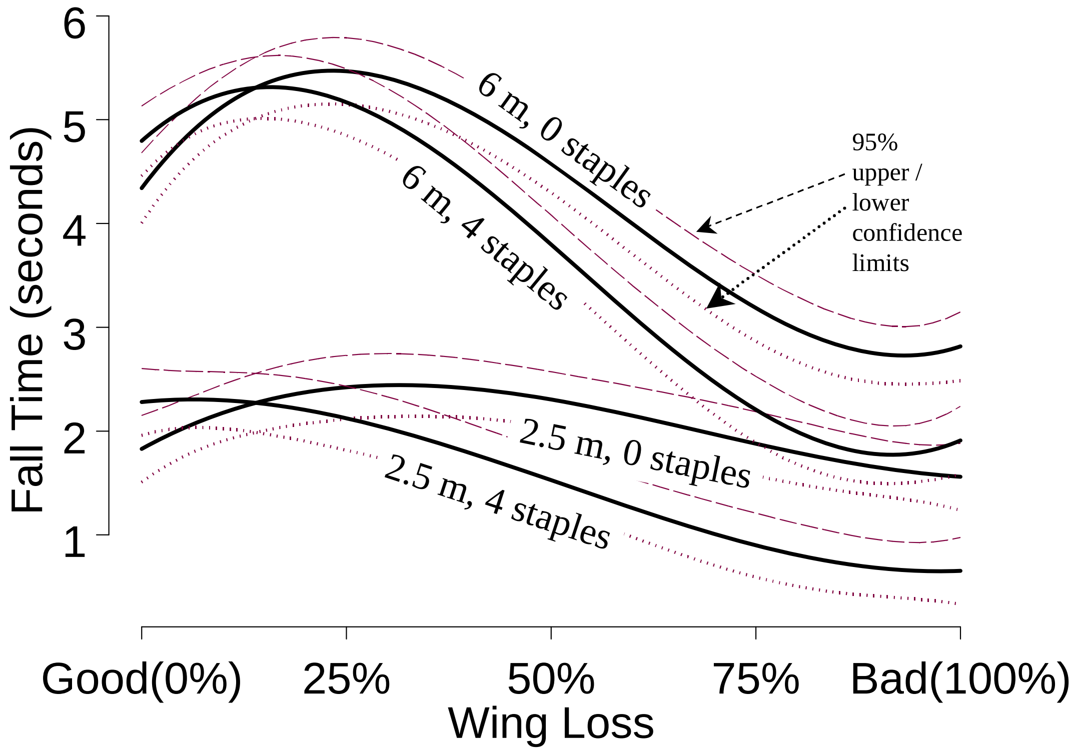

Figure 6. Bayes-Adjusted Cusum with Constant and Weibull (Shape = 3) Hazard; Characteristic Life = 100 Drops

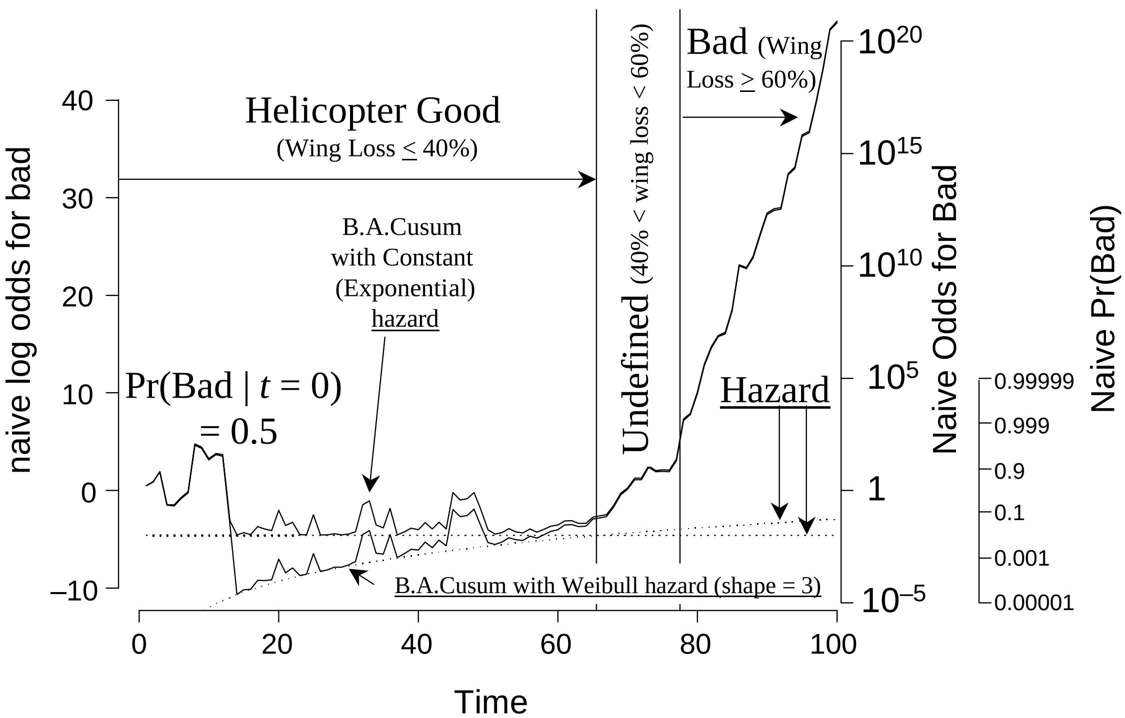

Table 1. Equivalent Thresholds between a Bayes-Adjusted Cusum and the Bayes' Posterior

Bayes-Adjusted	Bayes Posterior with Hazard 0.01		
Cusum	log(odds)	odds	probability
3	– 1.60	0.20	0.17
4	– 0.60	0.55	0.36
5	0.40	1.50	0.60

Figure Captions

Figure 1. Bayesian Sequential Updating

Figure 2. One-Sided and a Bayes-Adjusted Cusum Simulations

Figure 3. Bayes-Adjusted and Traditional Cusum Iteration

Figure 4. Steady-State Excess of Bayes-Adjusted over Page's One-Sided Cusum (mean of 20,000 simulated series for each of 17 scenarios; see text)

Figure 5. Predicted Fall Time vs. Wing Loss, Height, and Number of Staples

Figure 6. Bayes-Adjusted Cusum with Constant and Weibull (Shape = 3) Hazard; Characteristic Life = 100 Drops