A BAYESIAN ON-LINE CHANGE DETECTION ALGORITHM WITH PROCESS MONITORING APPLICATIONS

Pradipta Sarkar psarkar@iastate.edu

William Q. Meeker wqmeeker@iastate.edu

Center for Nondestructive Evaluation

and

Department of Statistics

Iowa State University

Ames, Iowa 50011

1 Introduction

1.1 Motivation: Casting Example

This work has been motivated by some issues of process monitoring in casting applications. In die casting, as time passes, material is deposited on the inside wall of the die. This foreign material can cause an unacceptable level of porosity in the output of the casting process. After accumulated material reaches a certain level, some of the output of the casting process no longer meets the specifications and the process must be stopped to clean the die. Sudden changes in output quality are also common in casting processes. The degree of porosity of the output of the casting process changes suddenly if, for example, one opens the door of the production facility thereby changing the laboratory temperature, causing an immediate larger accumulation of foreign material. Also day-to-day variation in weather, changes in humidity, and other factors can be responsible for this kind of change.

1.2 On-line vs. Off-line Change Detection Algorithm

Let $y_1, y_2, \ldots$ be a sequence of observed process output values (such as the degree of porosity) with conditional density $f(y_k|y_{k-1}, \ldots, y_1; \theta)$. Let the unknown time of change be t_0. In on-line change-detection algorithms, the objective is to detect a change as soon as possible, if not immediately after it occurs. Detection of the change-time is performed by considering all past data and stopping at the first point in time at which there is sufficient evidence that a change of importance has occurred.

Mathematically, such a stopping-time rule can be expressed as:

$$t_a = \min_n \{\Psi_n(y_1, \ldots, y_n) \geq \lambda\}$$

where $\Psi_1, \Psi_2, \ldots$ is a family of functions of n co-ordinates indexed by time and λ is a threshold specified to control the probability of detection errors. For example, Ψ_n may be the probability of exceeding the critical level at the n th inspection. Another alternative would be to define Ψ_n as $\max_{1 \leq j \leq n} \log \Lambda_j^n$, where Λ_j^n is the likelihood ratio for the observations from time j to time n (The likelihood ratio is generally defined as the ratio of the probability of the data maximized over all process states to the probability when the process is operating correctly.)

Off-line algorithms, on the other hand, can be looked upon as postmortem analyses. They are designed to collect data and test if there had been any change in the level of the process sometime during the past. If it is suggested that there was a change, these algorithms try to estimate the unknown change time. In this paper we will focus mainly on the on-line algorithms.

1.3 Overview

Section 2 provides a brief survey of literature and describes the basic problem of detecting a change of *known* magnitude. This basic problem is important for understanding the subsequent developments in the following sections. In section 3 we will generalize to more realistic situations of multiple changes of unknown magnitude, and trend in the process mean. Section 4 will be devoted to the discussion of rejection sampling and its use in Bayesian updating of the distributions used in making decisions. In section 5 we will use a simulated example to illustrate the behavior of the methodologies developed in sections 2 and 3. In section 6 we will formulate a decision theory-based method to determine an optimum inspection interval. Section 7 illustrates the methodology of section 6 with an example. Finally, in section 8, we discuss the possibility of extending the methodology described in this paper to a hierarchical Bayes formulation and provide some concluding remarks.

2 Brief Review of Literature

2.1 Bayes Type Algorithm for Detecting Changes in Process Monitoring

A Bayesian change detection approach has several advantages over non-Bayesian approaches. It is technically simple and is easier to implement and allows incorporation of available engineering information. There is a large amount of literature on Bayesian changepoint detection problems. Smith (1975) presents the Bayesian formulation for a finite sequence of independent observations. In particular he gives details for binomial and normal distribution models. Some other works in this area include

- Changepoint in univariate time series : Booth and Smith (1982), West and Harrison (1986).
- Gamma type random variables and Poisson processes : Diaz (1982), Hsu(1982), Raftery and Akman (1986).

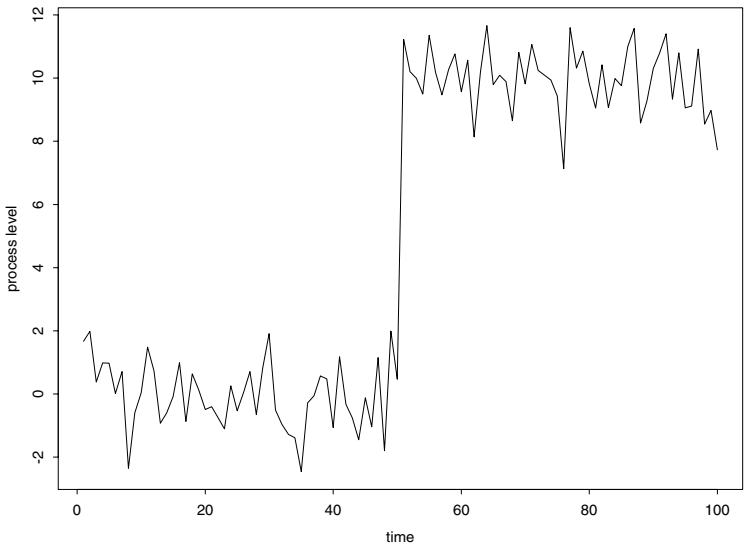

Figure 1: Increase in process mean, process variance remaining the same .

- Changepoint in linear models : Bacon and Watts (1971), Ferreira (1975), Holbert and Broemeling (1977), Choy and Broemeling (1980), Moen, Salazar and Broemeling (1985), Smith and Cook (1990).
- Hierarchical Bayes formulation of changepoint model : Carlin, Gelfand and Smith (1992).

Most of these references provide methodology for off-line problems. We are interested in process control applications where on-line algorithms are needed. The next subsection describes the simplest possible on-line scenerio.

2.2 Single Jump - Before and After-Change Parameter Values Known

In the simplest case, the parameter values (e.g., mean process level) before and after the change, $\theta = \theta_0$ and $\theta = \theta_1$ respectively, are both known. Figure 1 shows a process where the mean level changes after 50 cycles, but process variance remains the same. Figure 2 shows a process where the mean is constant but variance increases after 50 cycles. These kinds of changes are typical in many industrial processes. Detection of such changes has been discussed in detail by several authors in different contexts. Here we describe briefly the solution to this problem to motivate our following developments. Some of the pioneering work for this model is due to Girshick and Rubin(1952) and Shirayaev(1961) and is described nicely in Basseville and Nikiforov(1993). Our notation will follow closely that in Basseville and Nikiforov(1993). Assume that the prior distribution of (discrete) change time t_0 is Geometric(ρ) and is given by,

$$P(t_0 = k) = \rho(1-\rho)^{k-1}, k = 1, 2, \ldots$$

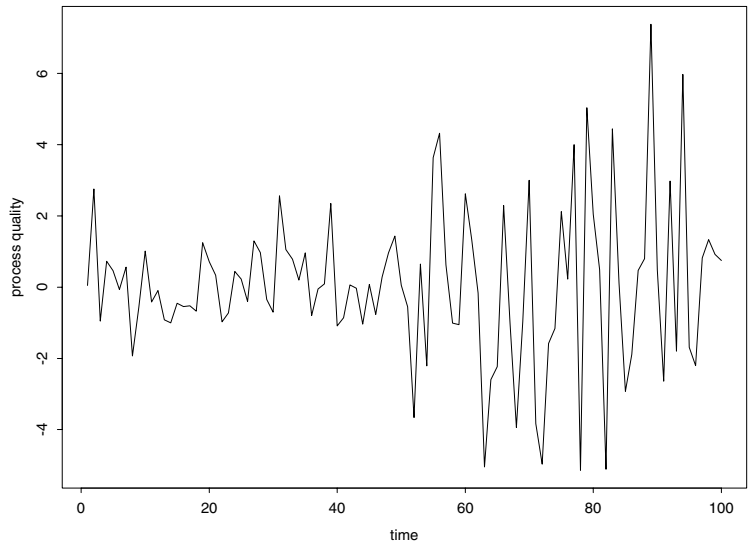

Figure 2: Increase in process variance, process mean remaining the same .

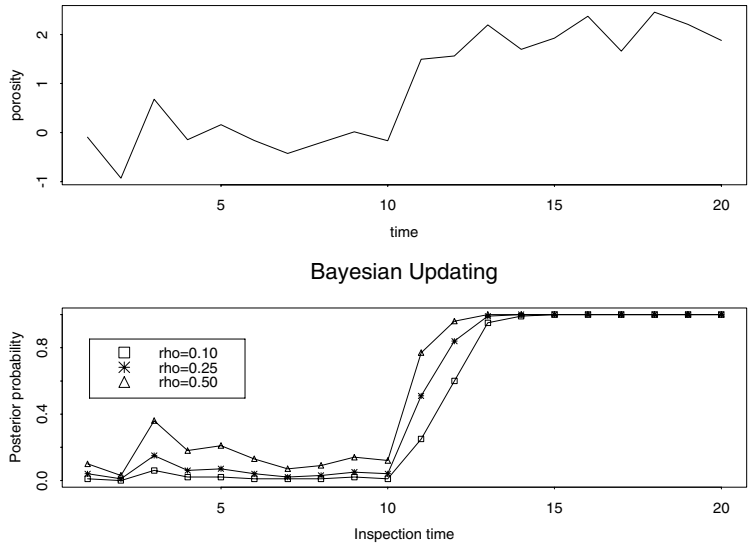

Figure 3: Bayesian updating when process mean changes. Top : Porosity level, Bottom: Updated probability that the process parameter has changed

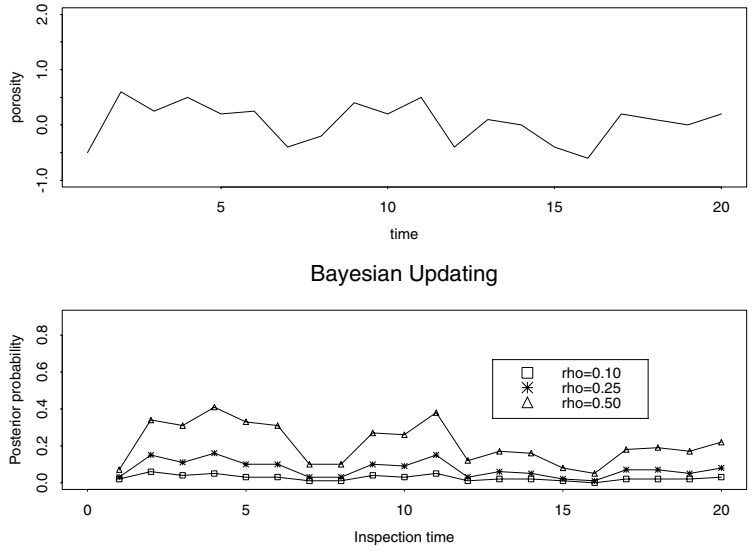

Figure 4: Bayesian updating when there is no change in the process mean.Top : Porosity level, Bottom: Updated probability that the process parameter has changed

where ρ is the probability of change in each production cycle, an estimate of which is available from process history. Let 0 and 1 denote, respectively, the process states before and after the change. The associated transition matrix is given by

$$P = \begin{pmatrix} p(0|0) & p(1|0) \\ p(0|1) & p(1|1) \end{pmatrix} = \begin{pmatrix} 1-\rho & \rho \\ 0 & 1 \end{pmatrix}.$$

Here $p(i|j)$ is the transition probability from state j to state i at each $t_i, i = 1, 2, \ldots$. We will denote by $p(0) = 1 - \pi$ and $p(1) = \pi$ the prior probabilities associated with the states 0 and 1 respectively at time 0. Then posterior probability that the process is at state 1 at the k th inspection/production cycle can be calculated using the Bayes rule and is given by,

$$\pi_k(1) = \frac{\pi_{k-1}(1) f(y_k;\theta_1) + (1-\pi_{k-1}(1)).\rho.f(y_k;\theta_1)}{\pi_{k-1}(1) f(y_k;\theta_1) + (1-\pi_{k-1}(1)).\rho.f(y_k;\theta_1) + (1-\pi_{k-1}(1))(1-\rho) f(y_k;\theta_0)},$$

where y_k is the observation at the k th inspection/production cycle and $f(y_k;\theta)$ is the likelihood at the k th cycle , $\theta = \theta_0, \theta_1$.

The stopping-time rule is :

$$t_a = \min_k \{\pi_k(1) \geq \lambda\}.$$

That is, a change is detected when the posterior probability of a change $\pi_k(1)$ exceeds λ for the first time. The value of λ is chosen to satisfy certain criteria, such as to have a specified probability of having a false detection (i.e., a detect when the $\theta = \theta_0$).

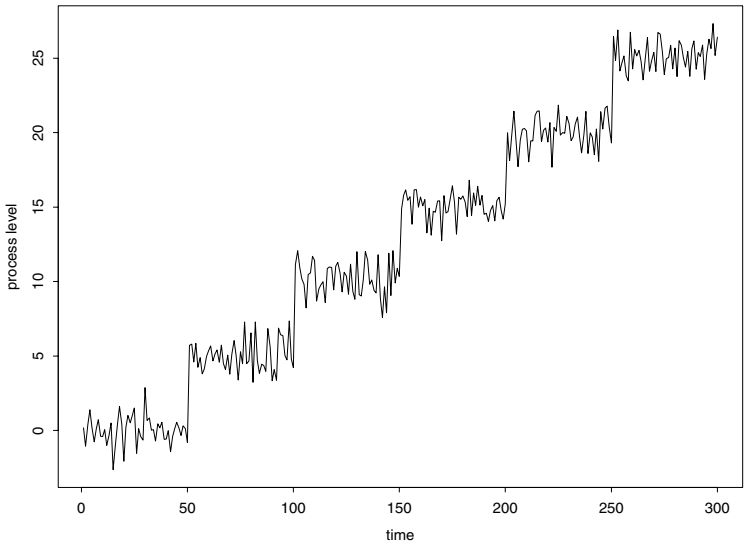

Figure 5: Multiple jumps in the process level.

3 Models for Process Degradation

Generally prior information about the distribution of the change time is available from the past history of the process. In the following description the process characterization parameter θ is a scalar. The theory, however, can be extended easily to the case when the parameter is a vector of two or more quantities.

3.1 Multiple-Jump Case

In practice, the multiple-jump situation is more common than the single jump situation. Although it is usual in the industrial process control applications to know the before-change parameter value (for example, when a process starts, typically the machine parameters are set to some fixed ideal values), once a process is running, more than one jump might occur between two consecutive inspection points. Also the process is generally not inspected at every production cycle, but it is done at some fixed time interval. Thus, it is necessary to modify and extend the simple model of the previous section. When a process jumps several times, in an on-line approach, our goal is to detect a change of an important magnitude as soon as possible after the underlying process characterization parameter has reached a critical level.

Let us denote the production cycle times by $t_1, t_2, \ldots$. Suppose that the actual measurements of process output are taken at $t^{(1)}, t^{(2)}, \ldots$. Thus $t^{(k)} = t_j$ for some $j \geq k$, say $t^{(k)} = t_{i_k}, k \geq 1$. Let θ_0 (known) be the target for the process. Also, when the process starts or it is readjusted, we assume that its level is reset to θ_0.

At each production cycle the process level either jumps (with probability q) or remains at the same

level (with probability $1-q$), and if there is a jump, the magnitude of the jump is a random variable with cumulative distribution function $H(.;\underset{\sim}{\alpha})$. So the jump at the i th production cycle is $Y_i = X_i\delta_i$ where the δ_i's are independent Bernoulli random variables with $P(\delta_i = 1) = q$, $P(\delta_i = 0) = 1 - q$, and

$$X_i \overset{iid}{\sim} \quad H(.;\underset{\sim}{\alpha}),$$

The level of the process after the kth measurement is denoted by $\theta^{(k)}$. This quantity can be calculated recursively as

$$\theta^{(k)} \equiv \theta_{i_k} = \theta_0 + \sum_{j=1}^{i_k} Y_j = \theta^{(k-1)} + \sum_{j=i_{k-1}+1}^{i_k} Y_j.$$

The posterior density of θ (expressing the state of knowledge given the available past data) after k th measurement is given by

$$\pi^{(k)}(\theta) \equiv \pi^{(k)}(\theta|y_k) = \frac{\int \pi^{(k-1)}(\theta^{(k-1)}) \,.\, g(\theta|\theta^{(k-1)}) \,.\, f(y_k;\theta)\, d\theta^{(k-1)}}{\int\int \pi^{(k-1)}(\theta^{(k-1)}) \,.\, g(\theta|\theta^{(k-1)}) \,.\, f(y_k;\theta)\, d\theta^{(k-1)}\, d\theta}, \tag{1}$$
$$= \frac{f(y_k;\theta)g^*(\theta)}{\int f(y_k;\xi)g^*(\xi)d\xi}$$

where $g(\cdot|\theta^{(k-1)})$ is the pdf of $\theta^{(k)}$ given $\theta^{(k-1)}$ and $g^*(\theta) = \int \pi^{(k-1)}(\theta^{(k-1)})g(\theta|\theta^{(k-1)})d\theta^{(k-1)}$, is the new prior just before the kth measurement. The stopping rule is formulated as in section 2.2.

3.2 Several Jumps Approximation to Trend

Now we consider the case when the process has a stochastic trend. In such a situation we are interested in detecting the change before the process deviates too much from the target, typically when the process level goes beyond a critical level. This case can be thought of as a combination of a number of small jumps with random size. This is a special subcase of the several jumps case with $q = 1$. We model each jump as a gamma(α, λ) increment. Then the sum of several jumps also follows a gamma distribution:

$$\left(\sum_{j=i_{k-1}+1}^{i_k} Y_j\right) \sim \text{Gamma}\left((i_k - i_{k-1})\,\alpha, \lambda\right), \tag{2}$$

$$g\left(\theta|\theta^{(k-1)}\right) = \frac{\lambda^{[i_k - i_{k-1}]\alpha}}{\Gamma\left([i_k - i_{k-1}]\alpha\right)} \,.\, \exp\left(-\lambda(\theta - \theta^{(k-1)})\right) \,.\, \left[\theta - \theta^{(k-1)}\right]^{(i_k - i_{k-1})\alpha - 1}; \theta > \theta^{(k-1)}.$$

The posterior density can be updated as usual using Bayes rule. The stopping rule is as before. Although the availability of a closed form expression for the conditional density of $\theta^{(k)}$ given $\theta^{(k-1)}$ makes the calculation relatively simpler, obtaining an analytical solution for the posterior density is impossible and even a numerical solution is computationally demanding. This has led us to use a simulation based approach using rejection sampling to compute the posterior distribution of θ. This approach will be described in the next section.

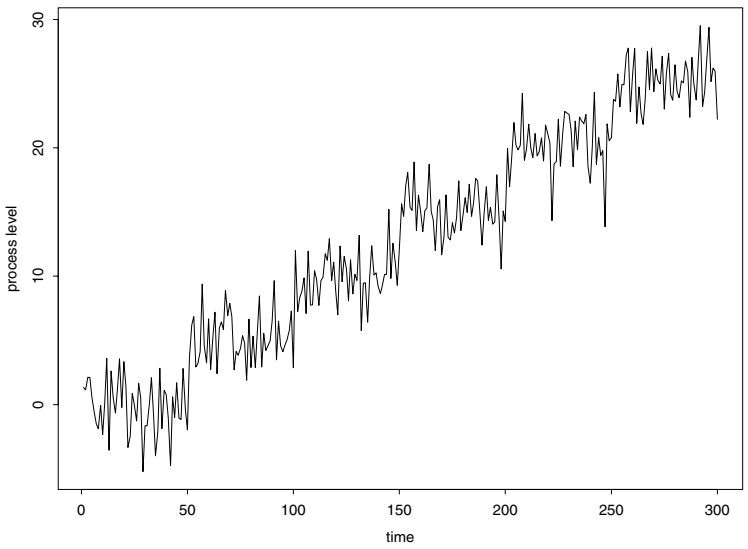

Figure 6: Multiple jumps approximation to trend.

4 Simulation Based Approach for Updating the Prior Distribution

The calculation of the posterior density of θ involves several integrals, as shown in equation (1). Computational needs will be considerably higher if the parameter θ is a vector. Such multidimensional numerical integration is difficult to carry out, requiring close attention to numerical methods and substantial amounts of computer time. Also since our problem is an on-line algorithm, finding the range of numerical integration at a later stage is not an easy task. This has led us to use the simulation based technique of rejection sampling, described by Smith and Gelfand(1992), to evaluate such integrals. Needed calculations are easy to program and can be carried out quite efficiently by using any standard mathematical/numerical software package. The method of rejection sampling used here is described in detail in the appendix.

We initially generate a large number of observations from the prior distribution $g(\theta)$. Then we propagate those observations according to our random gamma-distributed increment model given by equation (2), providing a large sample from the propagated prior density, say $g^*(\theta)$, at the next inspection. Then we combine this propagated prior with the likelihood to obtain a large sample from the posterior. For the trend case $q = 1$, so in order to get the propagated prior we add a random gamma increment to each observation at each production cycle.

To obtain the posterior distribution, we use rejection sampling. Rejection sampling can be viewed as a random filter. Points pass through the filter with probability equal to the relative likelihood of the point, as determined by the likelihood function of the data. The result is a random sample from the posterior distribution that can be used to approximate this distribution. Details are given

in the appendix.

We have written a FORTRAN program using NAG routines (The Numerical Algorithms Group Limited, 1993) to calculate the probability distribution updated at each inspection and the probability that the process is beyond the critical level. We have used S-Plus (Statistical Sciences, 1995) to obtain graphical output. The program requires the following inputs:

1. An initial prior distribution for θ,
2. The inspection interval,
3. Parameters for the gamma increment distribution,
4. The distribution of measurement errors,
5. The observations at the actual inspections,
6. The critical level for the process output level.

The initial prior distribution for θ can be estimated from the knowledge about the precision of the process (eg., how precise the machine is) and the measurement error distribution can be estimated from the past measurement errors. The parameters of the gamma increment model can also be estimated from past process data history.

5 Example of Process Monitoring

We consider a process that jumps by a random amount at every production cycle. The size of the jump at each cycle follows a gamma distribution with mean 0.2 and variance 0.1. The initial level of the process is 0. The critical level is $\theta_c = 5.1$. When the process is reset, the process output variable has a normal distribution with mean 0 and standard deviation 0.25. This variation is expected to be low and may be caused, for example, by the errors in adjusting the process equipment. We describe the measurement error with a normal distribution with mean 0 and standard deviation 0.5. In our illustration the process is inspected at every 10th production cycle. Suppose that the observed values were $y_{t^{(1)}} = 2.8$, $y_{t^{(2)}} = 3.8$, $y_{t^{(3)}} = 3.9$, $y_{t^{(4)}} = 4.1$ respectively, as marked by small squares in Figure 7. After each observed value of y is reported, the Bayesian updating formula is used to obtain the current posterior distribution. If the posterior probalitity of $\theta \geq \theta_c$ exceeds 0.1, the process is stopped and readjusted.

At each inspection, the dotted and the solid curves in Figure 7 denote the prior and posterior distributions respectively. At the first inspection our observation is towards the right tail of the prior suggesting that the process is a bit higher than predicted by the prior. The Bayesian updating scheme accounts for this and we get a posterior that is centered to the right of the prior mean. At the second inspection our observed value is a little lower than the center of the prior. This is also reflected in the posterior: although the prior probability of exceeding the critical level is greater than 0.1, the posterior distribution has a negligible probability above the critical level. This implies that in the light of our observations at the first two inspections, we do not yet need to adjust the process. After the third inspection, the posterior distribution again does not suggest that the process should be adjusted. Note that in this case the prior probability of exceeding the critical level is more than 0.5. When combined data, however, there is not much evidence (probability = 0.06) that the

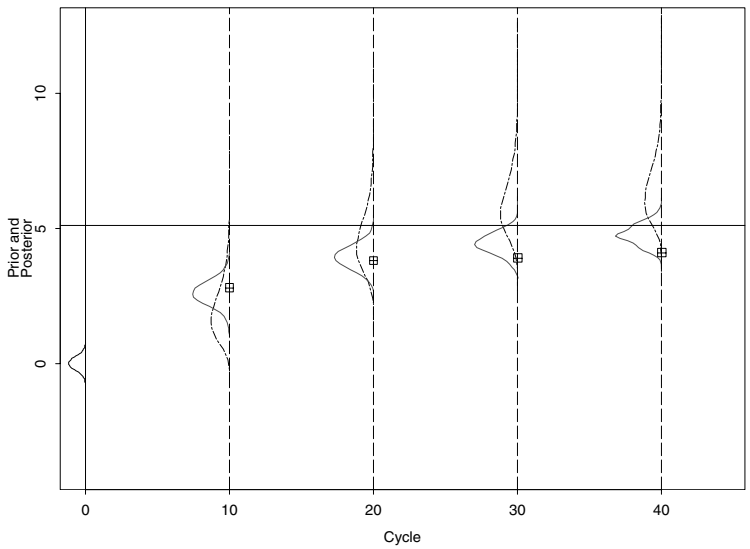

Figure 7: Bayesian Updating for Trend Case : Gamma Jump at Each Production Cycle.

process output level has reached the critical limit. At the fourth inspection our observation is at the left tail of the prior, but the posterior indicates a probability of 0.26 that θ has crossed the critical level. At this point the decision is to stop the process for adjustment.

6 Optimization of the Inspection Interval

In this section, for a given cost structure, we investigate the effect that the inspection interval has on total cost.This allows choosing the inspection interval to minimize the average cost over a long period of time.

For a given critical level θ_c we calculate the long run average cost for different inspection intervals. The optimum inspection interval is the one that minimizes the long run average cost of production. We also use a simulation-based approach for this calculation.

First we fix an inspection interval, say k. We generate 1000 values from the prior distribution and let the process degrade for the first k production cycles. At the kth cycle one of these degraded numbers, say θ^* is randomly selected and a number, say y is generated from the likelihood $f(.;\theta^*)$. This y is treated as the pseudo observed value at the kth production cycle (or equivalently, the first inspection) and is used for updating the prior to get a posterior. At each inspection we have two options: continue the process or adjust the process. If the updated probability of exceeding the cutoff value θ_c is more than the (predetermined) threshold level (e.g., 0.1) the process is readjusted. Thus we have a prior corresponding to a process restart instead of the usual updated posterior. Then the posterior is considered as the new prior and is allowed to degrade for the next k cycles after which another pseudo observation is generated by the above mechanism. Updating is performed

again and the procedure is repeated for a large number (say N) of production cycles.

Three types of costs are associated with the process :

- C_d: cost per unit of (squared) deviation between θ and the target θ_0.
- C_a: cost of a single adjustment of the process.
- C_i: cost of a single inspection.

Total cost over N cycles is calculated as,

Total cost = Cost of deviation + Cost of adjusting the process + Cost of inspection.

Cost of deviation at a particular cycle is C_d times the sum of the squared deviations of the (possibly) degraded numbers from the target. Summing these over all N cycles gives the total cost of deviation. Total cost of inspection is $C_i \times [N/k]$, where $[z]$ denotes the largest integer not exceeding z. Total cost of adjustment = $C_a\times$ number of adjustments (in N cycles). The long run average cost per unit time is the ratio of the total cost over N cycles to N, the total number of cycles, provided N is sufficiently large. We calculate the long run average costs for different values of k and choose the value of k that minimizes total cost.

7 Examples of Inspection Optimization

The long run expected cost based decision criteria, as discussed in the previous section, can be illustrated with the following example. We generated 10,000 observations from the process described in the previous example. But this time we varied the inspection intervals. The costs of deviation (per unit time), adjustment and inspection were fixed at 10, 200, and 100 respectively.

Table 1: Comparison of Long Run Average Costs for Different Inspection Intervals

Inspection Interval	Long Run Average Cost
1	185.92
2	141.24
3	123.64
4	118.91
5	119.29
6	121.98
7	122.06
8	124.80
9	124.88
10	130.29

We can see from the table 1 that the long run expected cost is minimized when the inspection interval is 4 (the figure for 5 is very close). So the optimum inspection interval should be 4 or 5.

Table 2: Comparison of Short Run Average Costs for Different Inspection Intervals : 50 Cycles Per Hour, 8 Hrs. Shift

Inspection Interval	Short Run Average Cost	s.d.	25 percentile	75 percentile
1	182.57	9.39	179.52	187.81
2	137.30	6.67	132.74	142.35
3	121.78	8.08	117.16	126.75
4	118.81	10.24	114.42	123.92
5	115.97	8.32	109.82	120.91
6	114.69	10.29	107.60	121.06
7	117.93	9.35	110.38	123.92
8	120.01	11.85	114.24	127.31
9	123.53	10.57	116.62	130.65
10	126.55	10.37	119.44	133.08

If we inspect too often, cost of inspection will dominate the figure for average cost. On the other hand, if we do not inspect frequently enough, the cost of deviation will be dominant. Thus optimum inspection interval is the one that balances these costs.

Table 2 is a comparison of short run average costs for different inspection intervals. In many practical situations a process runs for a shift of 8 or 9 hours and then the process is adjusted before the next shift starts. This means that the long run average cost may not be a feasible idea in many practical situations. To illustrate this situation we consider a process that runs for 8 hours with 50 cycles per hour. For each fixed inspection interval, we generate 200 such realizations of the process and calculate the average cost over 8 hour period. Table 2 shows that the optimum inspection interval is 6, but the average costs when the inspection intervals are 4, 5, 6 or 7 are very close. So although the results of table 1 and table 2 are not same, they are close. We do not lose much by approximating the short run average cost by long run average cost to make decision about the inspection interval.

8 Concluding Remarks And Directions for Future Developments

Although we have illustrated the multiple jump case for "jumps in mean level" of the process, the simulation-based algorithm is general enough to handle changes in other parameters of the process as well. Some other types of changes that one may be interested in include :

- Multiple changes in the process variance, process level (mean) remaining the same.
- Trend in the process level (mean) associated with an increase in the process variance. This is also of interest in modeling some financial time series.

We have so far assumed the process parameters to be either known or estimated from the past history. If we estimate them from the past history the analyses become empirical Bayesian in nature. If no information about the gamma parameters and/or the likelihood parameters is available, our model can be readily extended to a hierarchical Bayesian one by assigning diffuse prior distributions to these parameters. See Gelman, Carlin, Stern and Rubin(1995).

Also note that the equation (1) is very general, it does not involve any particular density for $g(.|\theta^{(k-1)})$. In the trend example we have assumed gamma increments, but there is nothing special to this distribution. We could have modeled the jumps as any other sequence of independent random variables.

Acknowledgements

This work was supported by the program for Integrated Design, NDE and Manufacturing Sciences under U.S. Department of Commerce, National Institute of Standards and Technology Cooperative Agreement No. 70NANB4H1535.

Appendix : Rejection Sampling

Suppose that we have random variates from an absolutely continuous density $m_1(x)$. But what we really want is a sample from $m_2(x)$, a probability density that is absolutely continuous with respect to $m_1(x)$. Rejection sampling is a useful tool in such situation. This methodology can be extended to the case when $m_2(x)$ is not known completely, but is known to be proportional to some function $m_3(x)$, [i.e., $m_2(x) = m_3(x)/\int m_3(x)$].

The algorithm proceeds as follows :

1. Find a positive constant M such that $m_3(x)/m_1(x) \leq M$, for all x.
2. Generate x from $m_1(x)$.
3. Generate u from Uniform(0,1).
4. If $u \leq m_3(x)/(M \times m_1(x))$ accept x, otherwise go to step 2.

For details see Smith and Gelfand(1992).

The rejection sampling algorithm can be used very effectively to sample from the posterior distribution (Meeker and Escobar, 1998). When formulated in terms of the relative likelihood, the algorithm has more intuitive appeal. Let $L(\theta; y)$ denote the likelihood of θ, and suppose that the maximum of the likelihood is attained at $\hat{\theta}$. Then the relative likelihood is defined as

$$R(\theta) = \frac{L(\theta; y)}{L(\hat{\theta}; y)}.$$

We note that the posterior distribution of θ can be written as

$$h_1(\theta|\text{data}) = \frac{L(\text{data}|\theta)h_0(\theta)}{\int L(\text{data}|\xi)h_0(\xi)d\xi} = \frac{R(\theta)h_0(\theta)}{\int R(\xi)h_0(\xi)d\xi},$$

where $h_0(\theta)$ is the prior pdf of θ. So $m_1 = h_0(\theta), m_3 = R(\theta)h_0(\theta), m_3/m_1 = R(\theta) \leq 1 (= M)$. Hence the rejection sampling algorithm in terms of relative likelihood can be formulated as follows :

1. Generate θ from the prior.
2. Generate u from Uniform(0,1).
3. If $u \leq R(\theta)$, accept θ, otherwise go to step 1.

We use this approach to generate a large sample from the posterior distribution $\pi^{(k)}(\theta|y_k)$.

References

Bacon, D. W. and Watts, D. G. (1971), Estimating the Transition between Two Intersecting Straight Lines. *Biometrika,* 58, 525-534.

Basseville, M. and Nikiforov, I.V. (1993), *Detection of Abrupt Changes, Theory and Applications,* Prentice Hall, Englewood, New Jersey.

Booth, N. B. and Smith, A. F. M. (1982), A Bayesian Approach to Retrospective Identification of Change Points, *J. Econ.,* 19, 7-22.

Carlin, B. P., Gelfand, A. E. and Smith, A. F. M. (1992) Hierarchical Bayesian Analysis of Changepoint Problems, *Applied Statistics,* 41, 389-405.

Choy, J. C. and Broemeling, L. (1980), Some Bayesian Inferences for a Changing Linear Model. *Technometrics,* 22, 71-77.

Diaz, J. (1982), Bayesian Detection of a Change of Scale Parameter in Sequences of Independent Gamma Random Variables. *Journal of Econometrics,* 19, 23-29.

Ferreira, P. E. (1975), A Bayesian Analysis of a Switching Regression Model : A Known Number of Regimes. *Journal of the American Statistical Association,* 70, 370-374.

Gelman, A., Carlin, J., Stern, H. and Rubin, D. (1995), *Bayesian Data Analysis,* Chapman and Hall, New York.

Girshick, M. A. and Rubin, H. (1952), A Bayes approach to a quality control model.*Annals of Mathematical Statistics,* 23, pp.114-125.

Holbert, D. and Broemeling, L. D. (1977), Bayesian Inference Relate to Shifting Sequences and Two-Phase Regression. *Communications in Statistics, Part A–Theory and Methods,* 6, 265-275.

Hsu, D. A. (1982), A Bayesian Robust Detection of Shift in the Risk Structure of Stock Market

Returns. *Journal of the American Statistical Association,* 77, 29-39.

Meeker, W. Q. and Escobar, L. A. (1998), *Statistical Methods for Reliability Data,*To be published by John Wiley & Sons, New York.

Moen, D. H., Salazar, D. and Broemeling, L. D. (1985), Structural Changes in Multivariate Regression Models. *Communications in Statistics, Part A–Theory and Methods,* 14, 1757-1768.

The Numerical Algorithms Group Limited (1993), *The NAG Fortran Library Manual, Mark 16.*

Raftery, A. E. and Akman, V. E. (1986) Bayesian Analysis of A Poisson Process with a Change Point. *Biometrika,* 73, 85-89.

Shirayaev, A. N. (1961), The problem of the most rapid detection of a disturbance in a stationary process. *Soviet Math. Dokl.,* no 2, pp.795-799.

Smith, A. F. M. and Cook, D. G. (1990), Straight Lines with a Change- Point. *Applied Statistics,* 29, 180-189.

Smith, A. F. M. and Gelfand, A. E. (1992), Bayesian Statistics Without Tears: A Sampling-Resampling Perspective. *The American Statistician,* 46, 84-88.

Statistical Sciences (1995), *S-PLUS Guide to Statistical and Mathematical Analysis, Version 3.3,* Seattle : Stat Sci, a Division of MathSoft Inc.

West, M. and Harrison, P. J. (1986) Monitoring and Adaptation in Bayesian Forecasting Models. *Journal of the American Statistical Association,* 81, 741-750.

Zacks, S. (1983) A Survey of Classical and Bayesian Approaches to the Change-Point Problems : Fixed Sample and Sequential Procedures of Testing and Estimation, *Recent Advances in Statistics,* 245-269, Academic Press, New York.

Pradipta Sarkar is a PhD (Statistics) candidate at Iowa State University. He received his Masters degree in Statistics with specialization in Statistical Quality Control and Operations Research from Indian Statistical Institute. William Meeker is Professor of Statistics and Distinguished Professor of Liberal Arts and Sciences at Iowa State University. He is a Fellow of the American Statistical Association and past Editor of Technometrics.