Statistica Sinica 4(1994), 535-548

A NON-GAUSSIAN KALMAN FILTER MODEL FOR TRACKING SOFTWARE RELIABILITY

Yiping Chen and Nozer D. Singpurwalla

The George Washington University

Abstract: Mainly because of their nice mathematical properties, Gaussian Kalman filter models have been widely used, especially for forecasting. However, in many situations Gaussian models may not do well, and as an alternative, non-Gaussian models may be more appropriate. Unlike Gaussian models, non-Gaussian models are hard to construct. In 1965, Bather developed some methods for constructing "invariant conditional distributions", which can be considered as special forms of non-Gaussian Kalman filter models. In this paper, we propose a non-Gaussian Kalman filter model, based on the family of invariant conditional distributions. This model is suitable for tracking reliability growth, and is applied to a well known set of data on software failures. Implementing the model requires numerical techniques for which the Gibbs Sampling algorithm is used.

Key words and phrases: Software reliability, reliability growth, forecasting, Gibbs sampling, Kalman filtering, non-Gaussian filtering, invariant conditional distributions.

1. Introduction and Overview

Over the past two decades, mainly as a result of an increased demand for credible software, the topic of software reliability has attracted the attention of computer scientists, software engineers, and statisticians. In a recent article, Littlewood and Strigini (1992) give a perspective on how critical the reliability of software has become, and the role of probability and statistics in assessing it.

A key stage in the development of reliable software is the "test-debug stage". Here the software is subjected to testing with the aim of determining whether the developed product meets specifications. When an input is run on the software and its outcome differs from what is expected, it is said that a "software failure" has occurred. Those parts of the code that cause the failure are referred to as "bugs". Whenever a software failure occurs, an attempt is made to identify the bug, and to remove it. This process is called "debugging the software". Debugging is intended to improve the quality of the software which is often expressed in terms of the reliability of the software. Many definitions for the reliability of software have been proposed, (see, for example, Nelson (1978), Brown and Lipow (1979), and Goel (1985)), but the one most commonly used is, "the probability of failure-free operation of a computer program for a specified time in a specified environment" (cf. Jelinski and Moranda (1972), Musa (1989)).

Since an attempt at identifying and removing a bug may not always result in a success, a question that arises is whether the overall reliability of the software is growing as the result of the debugging, and if so, how reliable is the software at its current stage? Numerous models have been proposed to describe the growth of reliability during the test-debugging stage. Many of these models take the approach of specifying the failure rate functions for the times between software failures. Others use non-homogeneous Poisson Processes to model the number of bugs that are detected and eliminated up to a certain time. The former have been referred to by Singpurwalla and Wilson (1992) as *models for times between successive failures*, and the later as *models for the number of failures.* Recently, Singpurwalla and Soyer (1992) have proposed a Gaussian Kalman filter model to describe reliability growth. This model, when applied to a well known set of software failure data, outperforms the predictive ability of competing models.

Since failure data are generally skewed, a use of Gaussian models for analyzing such data is not meaningful. In this paper we propose a non-Gaussian Kalman filter model, and discuss its application to the set of data alluded above. The theory underlying our model is based on a paper by Bather (1965) entitled "Invariant Conditional Distributions" (ICD); thus we refer to our model as the ICD model. In Section 2, we describe a non-Gaussian Kalman filter model in its general form; in Section 3, we discuss the relevance of this model to assess the growth of reliability and illustrate its application. An application of our ICD model entails expressions which, due to a lack of conjugacy of non-Gaussian terms, cannot be expressed in closed form. Thus its implementation calls for numerical techniques, and for this we resort to the Markov-chain simulation technique of "Gibbs Sampling"; see Gelfand and Smith (1990). The performances of the ICD model and the Gaussian Kalman filter model are then compared. Our conclusion is that a consideration of the non-Gaussian model results in an improved predictive performance over the Gaussian model. Thus, our recommendation is to use models of the type discussed here for assessing reliability growth.

2. A Non-Gaussian Kalman Filter Model

Largely due to their mathematical tractability, Gaussian Kalman filter models and their applications have been extensively reported in the statistical literature (cf. Meinhold and Singpurwalla (1983, 1986), and West and Harrison (1989)). However, in an important, but hitherto unnoticed paper, Bather (1965) introduced the concept of "Invariant Conditional Distributions", developed some theory, and gave four illustrative examples. Bather's paper is generally difficult to read; however, once decoded, this paper and its examples provides an appropriate framework for both Gaussian and non-Gaussian filtering. Smith (1979) uses Bather's result for a generalization of the Bayesian steady forecasting model, and Smith and Miller (1986) use it to address a problem in the prediction of records. In this section we use the ICD to develop the updating mechanism for a non-Gaussian Kalman filter model that is relevant to assessing reliability growth; in the sequel, we generalize some aspects of the work of Smith and Miller (1986).

2.1. Review of a Gaussian Kalman filter Model

For developing a non-Gaussian Kalman filter Model we shall adopt the format used by West and Harrison (1989) for Gaussian models. Accordingly, let X_n represent the observation at time n, θ_n the state of nature at n, $D_n = \{X_1, \ldots, X_n\}$ the collection of observations until n, and m_0 and B_0 the starting values of the Kalman filter. The notation "$X \sim N[\mu, \sigma^2]$" denotes the fact that X has a Gaussian distribution with mean μ and variance σ^2.

The Gaussian Model

The observation equation:	$(X_n \mid \theta_n) \sim N[F_n\theta_n, V_n]$.
The system equation:	$(\theta_n \mid \theta_{n-1}) \sim N[G_n\theta_{n-1}, W_n]$.
The starting distribution:	$(\theta_0 \mid D_0) \sim N[m_0, B_0]$.

The Results (Updating Equations)

The posterior of θ_{n-1}:	$(\theta_{n-1} \mid D_{n-1}) \sim N[m_{n-1}, B_{n-1}]$.
The prior of θ_n:	$(\theta_n \mid D_{n-1}) \sim N[G_n m_{n-1}, R_n]$.
The 1-step ahead prediction:	$(X_n \mid D_{n-1}) \sim N[f_n, Q_n]$.
The posterior of θ_n:	$(\theta_n \mid D_n) \sim N[m_n, B_n]$.

The appropriate expressions for m_n, B_n, R_n, Q_n and f_n can be found in West and Harrison (1989, p.110). Observe that we have used B_n in place of the C_n used by West and Harrison (1989); this is to avoid confusion with the C_n's of the following section.

2.2. Development of a non-Gaussian Kalman filter Model

By rewriting Example 2 of Bather (1965) in a format that is parallel to the one above, we are led to a model for non-Gaussian filtering. The notation "$X \sim \text{Gamma}(\upsilon, \theta)$" denotes the fact that X has a gamma distribution with shape parameter υ and scale parameter θ. Similarly, "$X \sim \text{Beta}(\sigma, \theta)$" denotes the fact that X has a beta distribution with parameters σ and θ. The coefficient C, which for us plays a key role for assessing reliability growth, is a scaling constant.

A Non-Gaussian Model

The observation equation: $(X_n|\theta_n) \sim \text{Gamma}(v, \theta_n)$.
The system equation: $(C\theta_n/\theta_{n-1}|\theta_{n-1}) \sim \text{Beta}(\sigma, v)$,
where v, C and σ are assumed to be known.
The starting distribution: $(\theta_0|D_0) \sim \text{Gamma}(\sigma + v, u_0)$.

The Results (Updating Equations)

The posterior of θ_{n-1}: $(\theta_{n-1}|D_{n-1}) \sim \text{Gamma}(\sigma + v, u_{n-1})$.
The prior of θ_n: $(\theta_n|D_{n-1}) \sim \text{Gamma}(\sigma, cu_{n-1})$.
The 1-step ahead prediction: $(X_n/(Cu_{n-1})|D_{n-1}) \sim$ Pearson Type VI (with $p = v$ and $q = \sigma$); (see Johnson and Kotz (1970), p.51).
The posterior of θ_n: $(\theta_n|D_n) \sim \text{Gamma}(\sigma + v, u_n)$, where $u_n = Cu_{n-1} + x_n$.

The restrictions on the model parameters indicated above, ensure that the set up remains within the framework of the ICD's. For a brief overview of the ICD's see Appendix A. However, for many applications the above restrictions are limiting, and need to be relaxed. This is done below.

2.3. A generalization of the non-Gaussian Model

The observation equation: $(X_n|\theta_n) \sim \text{Gamma}(\omega_n, \theta_n)$.
The system equation: $(C_n\theta_n/\theta_{n-1}|\theta_{n-1}) \sim \text{Beta}(\sigma_{n-1}, v_{n-1})$.
The initial information: $(\theta_0|D_0) \sim \text{Gamma}(\sigma_0 + v_0, u_0)$,

where C_n, ω_n, v_n and σ_n are assumed to be known and non-negative; furthermore, they are required to satisfy the condition

$$\sigma_{n-1} + \omega_n = \sigma_n + v_n, \qquad \text{for } n = 2, 3, \ldots.$$

Observe that our generalization allows the C's and σ's to change from stage to stage, and does not require that the observation equation and the system equation have a common fixed parameter. Even though our application does not call for the above generalizations, they are presented here for completeness.

The Results (Updating Equations)

The posterior of θ_{n-1}: $(\theta_{n-1}|D_{n-1}) \sim \text{Gamma}(\sigma_{n-1} + v_{n-1}, u_{n-1})$.
The prior of θ_n: $(\theta_n|D_{n-1}) \sim \text{Gamma}(\sigma_{n-1}, C_nu_{n-1})$.
The 1-step forecast: $(X_n/C_nu_{n-1}|D_{n-1}) \sim$ Pearson Type VI $(p = \omega_n, q = \sigma_{n-1})$.

The posterior for θ_n: $(\theta_n|D_n) \sim \text{Gamma}(\sigma_{n-1} + \omega_n, u_n)$,
where $u_n = C_n u_{n-1} + X_n$.

In the next section, we argue that when the above model is used for assessing reliability, the parameter C_n provides inference about the growth or decay of reliability. Thus it is meaningful to treat the C_n's as being unknown. However, when such is the case, and even if C_n were to be a constant C, the C_n and θ_n are correlated, and no close form updating formulae are available. Thus inference about C_n and θ_n would call for computer intensive technologies, such as Gibbs Sampling, which involves an extensive simulation, or the numerical solution of optimization problems, as is done by control engineers.

3. Relevance of Model for Tracking Reliability Growth

The model for software failures proposed here is based on the development of Section 2.3. In what follows, we replace the X_n's of the observation equation by T_n's, where T_n denotes the time between the $(n-1)$th and the nth failure, for $n \geq 1$. $T_0 \geq 0$ is chosen arbitrarily; it reflects our best assessment about the inter-failure times prior to observing any data. For predictions, we use the mean of the distribution of the 1-step ahead forecasts. To simplify the model specifications, we let $\omega_n = v_n = \sigma_n = 2$, and set $C_n = C$. The latter simplification is restrictive, but as will be argued soon, does provide advantages of interpretation vis-á-vis the overall growth or decay of reliability. When the above have been done, $\hat{t}_n$, our prediction for T_n, given the collection of inter-failure times $D_{n-1} = \{t_1, \ldots, t_{n-1}\}$, and the parameter C, is of the form

$$\hat{t}_n = E(T_n|D_{n-1}, C) = 2Cu_{n-1},$$

where

$$u_n = t_n + Ct_{n-1} + C^2 t_{n-2} + \cdots + C^n t_0.$$

The above relationships suggest that C plays a key role in describing the growth or decay of reliability. Specifically, for $C \geq 1, E(T_n|D_{n-1}, C)$ is always increasing in n, implying a *substantial* growth in reliability, whereas C equal to zero would suggest a *drastic* decay in reliability. Values of C that are intermediate to the above would, depending on D_{n-1}, imply a growth or decay in reliability. As a general rule, we should expect a growth (decay) in reliability as C tends towards 1 (0); for details about this, see Appendix B. Consequently, a suitable prior for C, assuming that there are no previous notions about the growth or the decay of the reliability of the software, is a Uniform on $(0, 1)$, and this is what we shall use. Other prior distributions on $(0, 1)$ may also be meaningful, depending on the prior information that is available.

As an interesting aside, the expression for u_n suggests that when $C < 1$, its role is identical to that of the weight constant in *exponential smoothing* (see Box and Jenkins (1976), Section 4.3). Thus it appears that there must be an analogy between the ARIMA $(0, 1, 1)$ model of Box and Jenkins (1976) and a non-Gaussian version of the ICD model of Section 2.3. For $C \geq 1$, the corresponding ARIMA process would not be invertible, and so this case will be of little interest to pursue.

3.1. Application to real data on computer software failures

Singpurwalla and Soyer (1992) used the "System 40" data of Musa (1979) to investigate the predictive performance of their model which we refer to as the S&S Model. To illustrate the workings of our ICD Model, and also to see how well it performs in relation to the S&S Model, we apply the former to the "System 40" data. Our application entails the following algorithm:

Step 1: Discretize the prior distribution of C at m points to obtain $\{c_i | i = 1, 2, \ldots, m\}$, and $P(C = c_i | D_0)$, for $i = 1, 2, \ldots, m$.

Step 2: Obtain the conditional distributions $(\theta_n | C, D_{n-1})$ and $(C | \theta_n, D_{n-1})$; the former is via the ICD model, and the latter by an application of Bayes law.

Step 3: For each c_i, perform a Gibbs Sampling on $(\theta_n | C, D_{n-1})$ and $(C | \theta_n, D_{n-1})$ s times, to obtain $c_i^{(s)}$, $i = 1, 2, \ldots, m$.
(See Appendix C for details).

Step 4: Predict T_n, the time between next failure via $\hat{t}_n$.
(See Appendix D for details).

Step 5: Observe $T_n = t_n$, and compute the posterior distribution $(C | t_n, D_{n-1}) = (C | D_n)$, and go to Step 2.

The number of points m for the discretization is chosen to be 200, and s, the number of iterations for the Gibbs Sampler, is set to 500. For an expository overview on Gibbs Sampling, see Casella and George (1990).

In Figure 1, we plot the mean values of the posterior distribution of C from stage 1 to 101; in Figure 2, we plot the initial prior and the final posterior distributions of C. Figure 1 shows that the mean of the posterior of C stabilizes after the first 15 observations at around 0.425, and Figure 2 shows the sharpness of the final posterior distribution of C. All of these indicate that the ICD Model is quite stable in assessing the reliability growth (decay), and that it is quite robust to the choice of the initial uniform distribution of C.

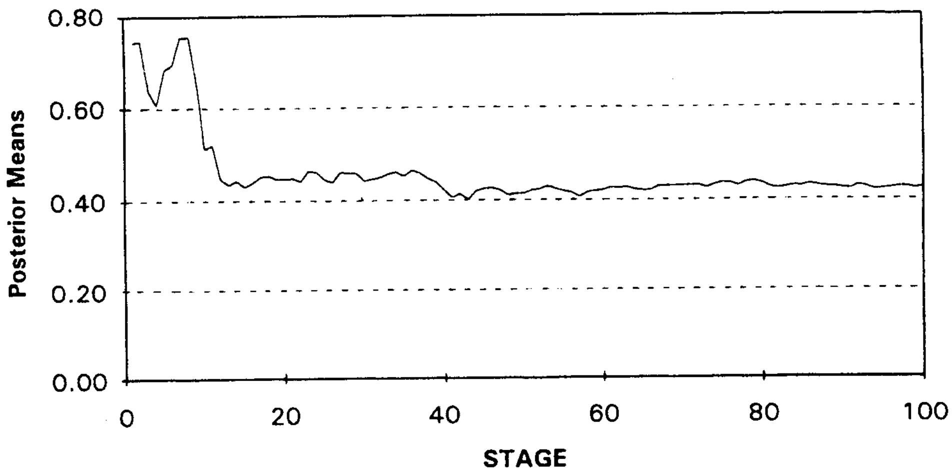

Figure 1. Mean of the posteriors of C from Stage 1 to 101.

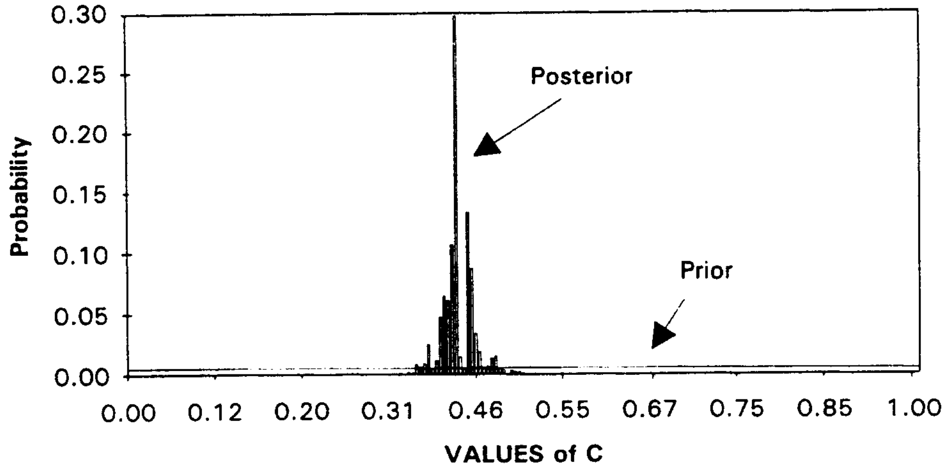

Figure 2. The prior and the final posterior of C.

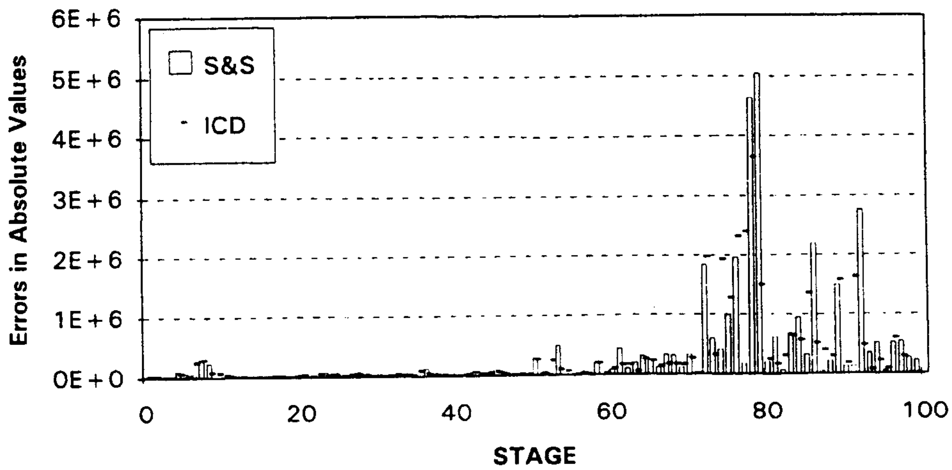

Figure 3. Absolute values of the errors of prediction of the ICD Model and the S&S Model.

3.4. Comparison of the results given by ICD Model and S&S Model

With the S&S Model, predictions of $Y_n = \log(T_n)$, the logarithms of the times between failures T_n are made. For this Model, the posterior distribution of Y_n is a Gaussian, and therefore T_n has the lognormal distribution. Thus to compare the ICD Model with the S&S Model, we exponentiate $\hat{y}_n$ the predictions given by the of S&S Model, to obtain $\hat{t}_n = \exp(\hat{y}_n)$. Observe that $\exp(\hat{y}_n)$ is an estimate of the median of the lognormal distribution of T_n. The predictions given by the S&S Model, together with the real data, and the predictions given by the ICD Model are given in Appendix E. The absolute value of each prediction error from both the ICD Model and the S&S Model are plotted in Figure 3. It shows that the predictions from the ICD Model outperforms the S&S Model most of time, with the sum of the absolute errors reduced by as much as 3,860,608.

To further support our claim of the superiority of the ICD Model over the S&S Model, we analyze their prequential ratio of likelihoods. Figure 4 shows the likelihood ratios $f(t_n|T(n-1), \text{ICD})/f(t_n|T(n-1), \text{S\&S})$, for $n = 1, \ldots, 100$. Here $T(n-1) = \{t_1, t_2, \ldots, t_{n-1}\}$, and $f(t_n|T(n-1), \text{ICD})$ and $f(t_n|T(n-1), \text{S\&S})$ are the predictive distributions of the ICD Model and the S&S Model, respectively. Most of the ratios are greater than 1, suggesting the superiority of the ICD Model over the S&S Model.

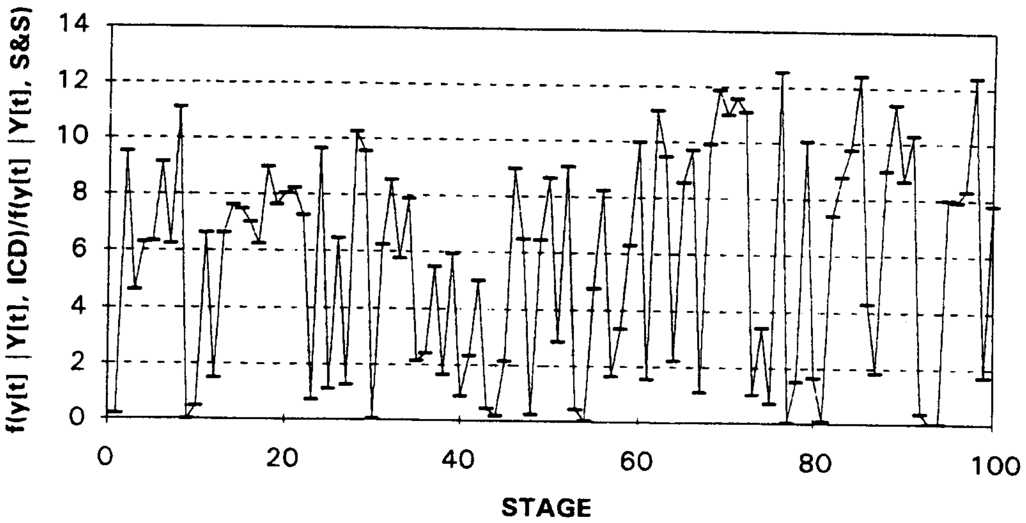

Figure 4. Prequential ratios of the likelihood.

The cumulative of the logarithm of the prequential likelihood ratio is plotted in Figure 5; this ratio is defined as

$$\prod_{i=1}^{n} \frac{f(t_i \mid T(i-1), \text{ICD})}{f(t_i \mid T(i-1), \text{S\&S})} \qquad n = 1, 2, \ldots, 100.$$

Though this ratio starts with a negative number, it soon becomes positive and increases with n to a very large number. The superiority of the ICD Model over S&S Model is, again, evident.

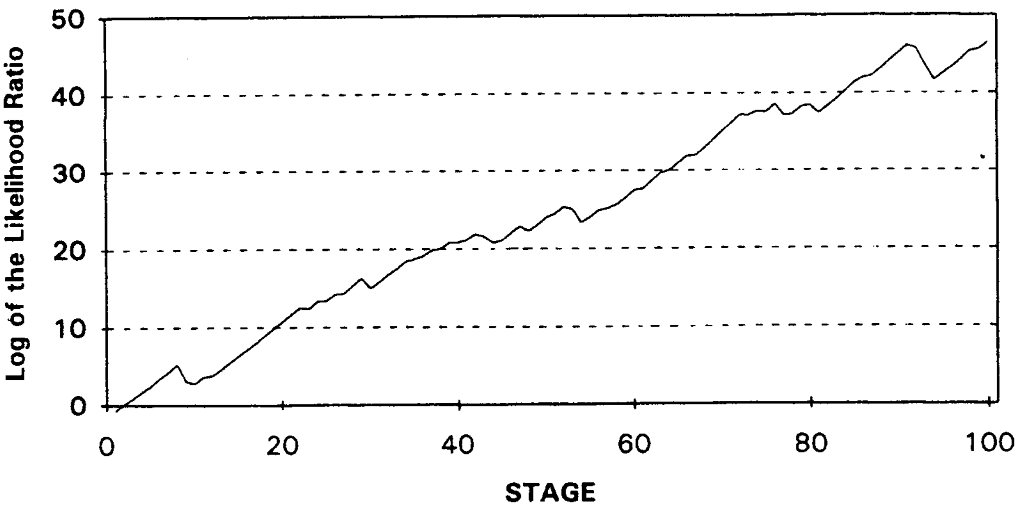

Figure 5. Cumulative prequential likelihood ratios.

Acknowledgements

The authors would like to thank Mr. Jingxian Chen for helpful discussion on the implementation of the Gibbs Sampling technique, and the referee for several comments which have resulted in the present version of the paper. Also acknowledged is the help of Ms. Sylvia Campodonico on producing the illustrations. This research was supported by Contract N00014-85-K0202 Office of Naval Research, Grant DAAH04-93-G-0020 Army Research Office, and Grant AFOSR-F49620-92-J-0030 Air Force Office of Scientific Research.

Appendix A: An Overview of the ICD Framework

A general parametric mathematical model can be stated as follows:

θ_n, the state, is of interest to us;

$X^n = (x_1, \ldots, x_n)$, all observations up to time n;

$(X_n | \theta_n, X^{n-1}) = (X_n | \theta_n)$, that is, given θ_n, X_n is independent of the history X^{n-1}.

There are two major difficulties that need be overcome to use this model. These two problems are:

(i) in general $(\theta_n | X^{n-1})$ becomes very complicated as n increases, and

(ii) the accumulation of $X^n = (x_1, \ldots, x_n)$ causes computer storage problem, as n becomes larger and larger.

To address this problem, Bather (1965) studied the conditions under which the relevant information $(\theta_n|X^{n-1})$ can be transformed through some sufficient statistics of fixed dimension. In particular, under what conditions, could we have

$$(\theta_n|X^{n-1}) = (\theta_n|u_n), \quad \text{where } u_n = u_n(x_n, u_{n-1})?$$

Here u_n, which depends on x_n and u_{n-1}, is a two-dimensional sufficient statistic.

We need to introduce some notation before proceeding further. Under the assumption that u_n is a sufficient statistic,

$p(\theta_n|\theta_{n-1})$, denotes the transition probability density function of θ_n given θ_{n-1};

$f(x|\theta)$, the observation probability density function of an observation;

$h_n(\theta_n|u_{n-1})$, the prior distribution of θ_n;

$g_n(\theta_n|u_n)$, the posterior distribution of θ_n;

$g_0(\theta_0)$, the initial distribution of θ.

Under suitable regularity conditions, it is shown that all the densities concerned have exponential forms as follows:

$$f(x|\theta) = a(\theta)b(x)\exp(\theta x),$$

$$g_n(\theta|u) = m_n(\theta)\exp(u\theta - \lambda_n(u)),$$

$$h_n(\theta|u) = s_n(\theta)\exp\{c_n(u)\theta - \rho_n(c_n(u))\};$$

or equivalently:

$$f(x|\theta) \propto b(x)\exp(\theta x),$$

$$g_n(\theta|u) \propto m_n(\theta)\exp(u\theta),$$

$$h_n(\theta|u) \propto s_n(\theta)\exp\{c_n(u)\theta\};$$

for some functions $a(\cdot), b(\cdot), c_n(\cdot), m_n(\cdot)$ and $s_n(\cdot)$, where $s_n(\theta) = m_n(\theta)/a(\theta)$ and $u_n = c_n(u_{n-1}) + x_n$.

The above results provide the same forms for the densities of our concern, together with the updating formula for u_n. However what is not good about the above results is the dependence of the $m_n(\cdot)$ and $c_n(\cdot)$ on n.

Bather (1965) investigated some additional conditions under which the $m_n(\cdot)$ and $c_n(\cdot)$ are invariant with respect to n. That is, $\exists\ m(\cdot)$ and $c(\cdot)$ such that

$$m_n(\cdot) = m(\cdot), \qquad c_n(\cdot) = c(\cdot), \quad \forall n.$$

When this is the case, we have

$$g_n(\theta_n|u_n) = m(\theta_n)\exp(u_n\theta_n - \lambda(u_n)) = g(\theta_n|u_n), \text{ and}$$
$$h_n(\theta_n|u_{n-1}) = s(\theta_n)\exp\{c(u_{n-1})\theta_n - \rho(c(u_{n-1}))\} = h(\theta_n|u_{n-1}).$$

The above are called the *Invariant* ($wrt\ n$) *conditional densities* (*ICDs*), and the ICD framework that we are referring to is as follows:

$$f(x|\theta) = a(\theta)b(x)\exp(\theta x),$$
$$g_n(\theta_n|u_n) = m(\theta_n)\exp(u_n\theta_n - \lambda(u_n)) = g(\theta_n|u_n), \text{ and}$$
$$h_n(\theta_n|u_{n-1}) = s(\theta_n)\exp\{c(u_{n-1})\theta_n - \rho(c(u_{n-1}))\} = h(\theta_n|u_{n-1}),$$

where the sequence of distributions $\{g_n\}$ ($\{h_n\}$) stays within the same family g (h) with only its parameter as a function of u_n (u_{n-1}).

Appendix B: The Effect of C on Reliability Growth

Observe that when $C = 0$,

$$\hat{t}_n = E(T_n|D_{n-1}, C) = 2Cu_{n-1} = 0,$$

and so the prediction is always 0. This implies that the next failure is experienced immediately after debugging, suggesting that the software reliability clearly has decayed.

On the other hand, when $C = 1$,

$$\hat{t}_n = E(T_n|D_{n-1}, C) = 2Cu_{n-1} = 2u_{n-1},$$

where

$$\begin{aligned} u_{n-1} &= t_{n-1} + Ct_{n-2} + C^2t_{n-3} + \cdots + C^{n-1} \\ &= t_{n-1} + t_{n-2} + t_{n-3} + \cdots + t_0. \end{aligned}$$

This implies that the predicted value of the next time between failure is twice the cumulative of all previous times between failures, suggesting that the reliability of the software has increased substantially.

Appendix C: The implementation of Step 3 of Section 3.1

Suppose that the conditional distributions $(\theta_n|C, D_{n-1})$ and $(C|\theta_n, D_{n-1})$ are given. Then for each (fixed) c_i, using computer simulation, we first sample

$\theta_n^{(1)}$ from $(\theta_n|c_i^{(0)}, D_{n-1})$, where $c_i^{(0)} = c_i$, and

using the result from the above sampling, we obtain

$c_i^{(1)}$ by sampling from $(C|\theta_n^{(1)}, D_{n-1})$.

In general, assuming that we have obtained $\theta_n^{(j)}$ and $c_i^{(j)}$, the $\theta_n^{(j+1)}$ and $c_i^{(j+1)}$ are obtained by first sampling

$\theta_n^{(j+1)}$ from $(\theta_n | c_i^{(j)}, D_{n-1})$,

and then sampling

$c_i^{(j+1)}$ from $(C | \theta_n^{(j+1)}, D_{n-1})$.

This process continues until the $c_i^{(s)}$ is obtained for s, a pre-determined number of iterations.

When we finish the above exercise for c_i, we move on to c_{i+1}, and so on. Step 3 terminates when the above simulation is performed on every c_i in $\{c_i | i = 1, 2, \ldots, m\}$ and the sequence $\{c_i^{(s)} | i = 1, 2, \ldots, m\}$ is produced.

Appendix D: The Prediction of T_n, in Step 4 of Section 3.1

Based on the theory of Gibbs Sampling, we have the marginal distribution of θ_n as

$$f(\theta_n | D_{n-1}) \simeq \frac{1}{m} \sum_{i=1}^{m} f(\theta_n | c_i^{(s)}, D_{n-1}),$$

which is a linear combination of the prior for θ_n, given $c = c_i^{(s)}, i = 1, 2, \ldots, m$. Applying the results from Section 2.2, we obtain the density function of the predictive distribution of T_n

$$f(t_n | D_{n-1}) \simeq \frac{1}{m} \sum_{i=1}^{m} f(t_n | c_i^{(s)}, D_{n-1}).$$

The prediction of T_n, the next time between failures is then taken as

$$\begin{aligned} \hat{t}_n = E(T_n | D_{n-1}) &\simeq \frac{1}{m} \sum_{i=1}^{m} E(T_n | c_i^{(s)}, D_{n-1}) \\ &= \frac{1}{m} \sum_{i=1}^{m} 2 \cdot c_i^{(s)} \cdot u_{n-1}(c_i^{(s)}), \end{aligned}$$

where $u_{n-1}(c) = t_{n-1} + ct_{n-2} + c^2 t_{n-3} + \cdots + c^{n-1} t_0$, is a function of c and $T(n-1) = \{t_1, t_2, \ldots, t_{n-1}\}$.

Appendix E: Musa's Data and Predictions by the ICD Model and the S&S Model

n	$t(n)$	ICD prediction	S&S prediction
0	320	510.00	–
1	14390	1204.36	319.99
2	9000	22058.89	49418.76
3	2880	23924.20	22236.62
4	5700	20255.50	4927.09
5	21800	25632.41	10413.51
6	26800	48750.13	48004.90
7	113540	88110.80	56586.91
8	112137	244956.63	286440.66
9	660	290003.82	261373.89
10	2700	110755.53	818.43
11	28793	65888.62	3886.44
12	2173	46894.56	54929.40
13	7263	21388.95	3029.24
14	10865	16430.52	11316.04
15	4230	16066.28	17231.77
16	8460	11269.56	5990.43
17	14805	13728.46	12564.66
18	11844	19914.57	22720.07
19	5361	19629.38	17370.18
20	6553	13669.69	7305.34
21	6499	12744.48	8953.28
22	3124	11043.04	8763.92
23	51323	8872.73	3962.57
24	17010	51652.02	80314.23
25	1890	38217.43	24194.46
26	5400	17364.48	2319.21
27	62312	14942.98	7038.58
28	24826	64362.37	95854.78
29	26335	53669.84	35438.79
30	363	45508.51	37330.68
31	13989	21629.86	410.53
32	15058	22453.09	20047.29
33	32377	24529.41	21467.89
34	41362	42373.38	48051.63
35	4160	56608.08	61709.57
36	82040	31062.48	5391.53
37	13189	89922.38	127928.87
38	3426	52177.42	18268.28
39	5833	24836.98	4378.32
40	640	14455.37	7629.25
41	640	5929.40	749.67
42	2880	3246.33	747.07
43	110	3563.77	3611.59
44	22080	1699.38	120.35
45	60654	19390.86	33445.11
46	52163	59924.55	97102.96
47	12546	69002.42	81958.45
48	784	37764.68	17991.99
49	10193	16920.65	969.08
50	7841	15401.10	14528.02
51	31365	13336.63	10941.90
52	24313	32765.46	47112.43
53	298890	35197.25	35781.24
54	1280	267522.66	506611.63
55	22099	114274.11	1628.10
56	19150	65709.57	32894.16
57	2611	41151.91	28092.17
58	39170	20532.78	3436.01
59	55794	42022.25	59716.70
60	42632	64982.47	86167.72
61	267600	65464.91	64446.04
62	87074	254521.68	445675.81
63	149606	182153.15	135423.75
64	14400	204574.95	238005.20
65	34560	96365.12	20226.06
66	39600	69572.82	50497.13
67	334395	65522.52	57947.17
68	296015	316144.10	544305.44
69	177355	390860.27	475570.03
70	214622	329747.45	275861.63
71	156400	330965.95	334991.66
72	166800	280318.57	238957.23
73	10800	254832.05	254217.33
74	267000	123497.40	14485.34
75	2098833	291172.69	416760.75
76	614080	1935391.90	3611422.25
77	7680	1351849.40	988926.88
78	2629667	629965.61	10155.20
79	2948700	2591260.15	4666077.00
80	187200	3667195.44	5226017.50
81	18000	1677352.01	288869.81
82	178200	731019.82	24878.79
83	487800	477826.72	273314.97
84	639200	642735.98	780934.00
85	334560	813990.24	1030877.13
86	1468800	651416.14	520881.38
87	86720	1555295.16	2418596.75
88	199200	729633.94	126358.53
89	215200	482663.53	299971.44
90	86400	381221.35	323716.69
91	88640	235176.38	124444.38
92	1814400	182481.61	127318.24
93	4160	1634926.93	2971369.00
94	3200	678584.40	5319.34
95	199200	287712.72	4039.84
96	356160	298391.67	300655.94
97	518400	431429.47	549136.63
98	345600	626867.43	809060.63
99	31360	551257.58	527703.38
100	265600	266743.05	43242.98

References

Bather, J. A. (1965). Invariant conditional distributions. *Ann. Math. Statist.* **36**, 829-846.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control.* Holden-Day, Oakland, California.

Brown, J. R. and Lipow, M. (1979). Testing for software reliability. *Proc. Internat. Conf. Reliable Software*, Los Angeles, CA, Apr. 1979, 518-527.

Casella, G. and George, E. I. (1990). An introduction to Gibbs sampling. Technical Report 96, Statistics Research Center, University of Chicago.

Gelfand, A. E. and Smith. A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.* **85**, 398-409.

Goel, A. L. (1985). Software reliability models: Assumptions, limitations, and applicability. *IEEE Trans. Software Engrg.* **6**, 1411-1423.

Jelinski, Z. and Moranda, P. (1972). Software reliability research. *Statistical Computer Performance Evaluation* (Edited by W. Freiberger), 465-484, Academic Press, New York.

Johnson, N. and Kotz, S. (1970). *Continuous Univariate Distributions-2.* Houghton Mifflin Company.

Littlewood, B. and Strigini, L. (1992). The risks of software. *Scientific American*, November 1990, 62-65.

Meinhold, R. J. and Singpurwalla, N. D. (1983). Understanding the Kalman filter. *Amer. Statist.* **37**, 123-127.

Meinhold, R. J. and Singpurwalla, N. D. (1986). Introduction to Kalman filtering — A Bayesian view. In *Theory of Reliability* (Edited by A. Serra and R. E. Barlow), 208-214, Course XCIV, Italy.

Musa, J. D. (1979). Software reliability data. *IEEE Comput. Soc. Repository.*

Musa, J. D. (1989). Faults, failures, and a metrics revolution. *Quality Time*, March 1989.

Nelson, E. (1978). Estimating software reliability from test data. *Micro-Electron. Rel.* **17**, 67-74.

Singpurwalla, N. D. and Soyer, R. (1992). Non-homogeneous autoregressive processes for tracking (software) reliability growth, and their Bayesian analysis. *J. Roy. Statist. Soc. Ser. B* **54**, 145-156.

Singpurwalla, N. D. and Wilson, S. P. (1992-3). Models for assessing the reliability of computer software. *Naval Research Reviews*, Joint **44** - **45**.

Smith, J. Q. (1979). A generalization of the Bayesian steady forecasting model. *J. Roy. Statist. Soc. Ser. B* **41**, 375-387.

Smith, R. L. and Miller, J. E. (1986). A non-Gaussian state space model and application to prediction of records. *J. Roy. Statist. Soc. Ser. B* **48**, 79-88.

West, M. and Harrison, J. (1989). *Bayesian Forecasting and Dynamic Models.* Springer-Verlag, New York.

Department of Operations Research, The George Washington University, Washington, DC 20052, U.S.A.

(Received February 1993; accepted January 1994)