RC25029 (W1007-099) July 29, 2010
Mathematics

IBM Research Report

Computational and Monte-Carlo Aspects of Systems for Monitoring Reliability Data

Emmanuel Yashchin
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

Research Division
Almaden - Austin - Beijing - Cambridge - Haifa - India - T. J. Watson - Tokyo - Zurich

Computational and Monte-Carlo Aspects of Systems for Monitoring Reliability Data

Emmanuel Yashchin[1]

IBM, Thomas J. Watson Research Ctr., Box 218,
Yorktown Heights, NY 10598, USA, *yashchi@us.ibm.com*

Abstract. Monitoring plays a key role in today's business environment, as large volumes of data are collected and processed on a regular basis. Ability to detect onset of new data regimes and patterns quickly is considered an important competitive advantage. Of special importance is the area of monitoring product reliability, where timely detection of unfavorable trends typically offers considerable opportunities of cost avoidance. We will discuss detection systems for reliability issues built by combining Monte-Carlo techniques with modern statistical methods rooted in the theory of Sequential Analysis, Change-point theory and Likelihood Ratio tests. We will illustrate applications of these methods in computer industry.

Keywords: Cusum, Lifetime Data, SPC, Warranty, Wearout

1 Introduction

This paper will focus on problems of data monitoring related to so-called time-managed lifetime data streams that are frequently encountered in reliability applications. Specifically, consider a sequence of lifetime tests corresponding to points in time t = 1,2,.., T. In what follows we will refer to these points as "vintages". They could, for example, correspond to dates at which batches of items were produced; as time goes by, these batches generate lifetime data. In other words, data corresponding to a given vintage can be viewed as an outcome of a lifetime test (Fig. 1). The results pertain to a specific point in time (typically, time at which the table has been compiled). For example, in Fig. 1 the table was compiled on 2007-08-02; however, the most recent vintage for which data is available is 2009-07-21. The lifetime tests corresponding to a given vintage typically have a Type-I censoring structure. For example, for the first vintage the number of items on test is 120; of these, 6 items failed, 2 items got right-censored in midstream, and the remaining 112 items survived till the present point in time and thus are type-1 censored. The table shows a distinct triangular structure due to the fact that for very recent vintages only results for relatively short time horizons are available.

In many applications the key problem is one of detection: one is interested in statistical methodology that enables rapid detection of unfavorable process conditions that manifest themselves through the data of type shown in Fig. 1. In essence, the situation here is similar to one handled by conventional

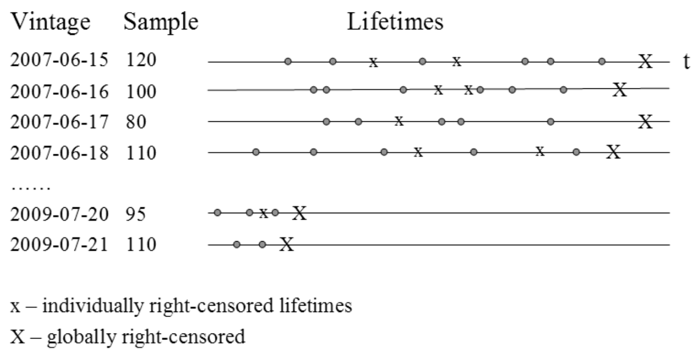

Fig. 1. A sequence of lifetime tests corresponding to range of vintages (2007-06-15, 2009-07-21), as compiled on 2009-08-02.

Statistical Process Control (SPC) methods: one would like to detect rapidly unfavorable changes in parameters of the process driving lifetime data, while keeping the rate of false alarms at some fixed low level. However, the conventional SPC performance measures are not applicable in situations of type described in Fig. 1 because the central premise of the SPC setup (that requires the observations corresponding to a given time to remain unchanged as the data is compiled for subsequent points in time) is no longer valid. For example, for the first vintage in Fig. 1 one could expect to see additional failures and censored points as we compile data for subsequent points in time. So, if a given characteristics of a vintage is represented as a point on a control chart, then this point will continue to change as the new points on the control chart are coming online. Furthermore, this point will continue to change even if there are no new data points (vintages) coming online. For example, if we recompile the data in Fig.1 on the next day (2009-08-03), we could end up with a situation where the last vintage for which the data is available is the same as before, i.e., 2009-07-21 (yet, the data for every vintage would need to be adjusted). Therefore, the concept of Run Length would not be suitable for assessing the performance of this type of control charts. We call such setup as "control charting with *dynamically changing observations (DCO)*" to emphasize the fact that the previous data points on a control chart would continue to change as the new data comes in.

Another practical aspect one has to deal with is the "time-management" of the lifetime data: the data corresponding to older vintages is likely to become either gradually underpopulated (especially with respect to items with longer lifetimes) or become unavailable altogether because of the administrative constraints of the data management. For example, if a machine carrying a 3-year warranty is introduced into the field at some point in time (corresponding to its "vintage"), then one could expect to find in the database only lifetimes occurring in the first 3 years of service. Information on subsequent failures would typically be either unavailable or unreliable, in light of missing infrastructure for keeping track of censoring and failure events beyond a 3-year horizon. As the vintage becomes too "old", it could also disappear from the "rolling" database completely - another artifact of a standard warranty data management policy. For example, after recompiling the data in Fig.1 on the next day (2007-08-03), one could find that the data for the first vintage 2007-06-15 is no longer available.

The computational challenges in the phases of design, analysis and implementation of this type of control schemes are substantial. In conventional control charting one can typically convert the sequence of observations to a sequence of control schemes (e.g., Cusum or EWMA) that obey some form of a Markov property. This enables one to design and analyze monitoring procedures by taking advantage of the theory of Markov Chains (though more complex cases still require Monte Carlo approach). In contrast, in the case of schemes with DCO, no such Markovian representation is apparent and we have to rely almost exclusively on simulation.

A number of articles and books have been published that deal with various aspects of monitoring lifetime data. For example, likelihood ratio methods for monitoring parameters of lifetime distributions in non-DCO setting were discussed in Olteanu and Vining (2009) and Sego et al. (2009). Several methods for monitoring warranty data by using Shewhart-type procedures are discussed in Dubois et al. (2008), Wu and Meeker (2002). Steiner and McKay (2000, 2001) discuss methods and applications related to monitoring of type I censored data. This type of data (in conjunction with an EWMA monitoring procedure and Weibull observations) was considered in Zhang and Chen (2004). Analysis of warranty claims data is discussed in Blischke and Murthy (2000), Doganaksoy et al. (2006), Kalbfleisch et al. (1991), Lawless (1998) and Lawless and Kalbfleisch (1992). Methods for analysis of failure data based on marginal counts of warranty claim (and under incomplete information about items introduced into the field) are discussed in Karim et al. (2001). Finally, methods based on change-point analysis for hazard curves have also be considered by a number of authors, e.g., Patra and Dey (2002).

In the next section we will present the basic approach to design and analysis of non-DCO control schemes. In Sec. 3 we will focus on computational and Monte-Carlo issues. In Sec. 4 we discuss the problem of detecting wearout conditions.

2 Basic Approach

The base methodology is based on using a version of the weighted Cusum approach, e.g. see Yashchin (1993). The key steps are as follows:

1. Sort data in accordance with vintages of interest. This will make sure that the control scheme is tuned to detect unfavorable changes that happen on

the time scale of these vintages. Typically, several representations of type shown in Fig. 1 are available for the same data collection. For example, a given component of a PC can be associated with a component manufacturing vintage, machine ship vintage or calendar time vintage. If one is interested, for example, in detection of a problem at the component manufacturer's then the vintages in Fig. 1 should correspond to component manufacturing dates. If one is interested in detection of problems at the PC assembly plant then one should organize Fig. 1 so that the vintages correspond to machine ship dates. To detect changes related to introduction of a new version of an operating system one should construct Fig. 1 data with vintages corresponding to calendar time (i.e., a row in Fig. 1 represents machines that were in the field at the respective date).

2. Introduce time scale transformation that corresponds to *reference hazard* behavior, so as to reduce the number of monitored parameters. For example, let us suppose that one anticipates a hazard curve that is proportional to some known function $h_0(t)$ exhibiting complex behaviour (for example, U-shaped). Then it may be beneficial to switch from the natural time scale t to a scale determined by $\int_0^t h_0(z)dz$. Data presented on such a scale would generally be easier to model parametrically: for example, if the anticipated hazard pattern indeed holds exactly then the lifetimes on the transformed scale are exponentially distributed. If, on the other hand, the anticipated pattern holds only approximately then there is a good possibility that one could model lifetimes through one of numerous extensions of the exponential law, e.g., Weibull family. In what follows we will assume that we are already working on a scale where a relatively simple parametric model applies. For the sake of simplicity, we will assume that the lifetime distribution is Weibull.
3. For every lifetime distribution parameter (or a function of parameters of interest, say, λ), establish a sequence of statistics $X_i, i = 1, 2, \ldots$ to serve as a basis of monitoring scheme. In general, one should try to use control sequences that represent unbiased estimates of the underlying parameters (typically, any loss of statistical power due to this strategy will be minimal - and it will be more than offset by the ensuing visualization and design advantages). For example, such a parameter could represent the expected lifetime or an expected rate of failures. Another parameter of interest could be a measure of wearout (for example, under the Weibull assumption one can use the shape parameter as such a measure). Yet another one could be the scale parameter of the lifetime distribution. One will generally choose parameters that are meaningful to the users and facilitate problem diagnostics.
4. For control sequences of interest, obtain corresponding weights. These weights determine the impact of individual vintages based on inherent properties of control sequence members (such as number of failures observed for the vintages or number of Machine-Months (MM) for the vintages; we will refer to such weights as $w_i \geq 0, i = 1, 2, \ldots$) or based on

desired properties of the monitoring scheme (such as weights that enhance the importance of more recent vintages).

5. For every parameter establish acceptable and unacceptable regions. For example, if λ represents the expected failure rate (say, expected number of fails per 1000 MM of service), then we could set the acceptable level of failure rate to λ_0 and the unacceptable rate to $\lambda_1 > \lambda_0$.
6. For every parameter establish the acceptable rate of false alarms. As noted earlier, in the context of schemes with DCO we cannot use the criteria related to Run Length (such as ARL); one reasonable criterion appears to be related to the *probability of flagging* a parameter. One can control the rate of false alarms in the monitoring system by setting this probability to a low level that reflects the desirable degree of trade-off between false alarms and sensitivity.
7. Deploy the designed scheme to every relevant data set; flag this data set if out-of-control conditions are present. For some monitoring systems the deployment will involve massive re-computing of control sequences, thresholds and control schemes on a regular basis. For example, for PC manufacturing operation it was considered suitable to activate the system on a weekly basis (note, however, that the vintages in Fig. 1 were still being summarized on a daily basis). For other systems re-computing could be, at least in part, event-driven.

3 Computational and Monte-Carlo Aspects

For every monitored parameter we convert the control sequence $X_i, i = 1, 2, \ldots, N$ to the values of a control scheme $S_i, i = 1, 2, \ldots, N$ via the weighed Cusum algorithm. For example, for a parameter λ we can use the *Weighted Geometric Cusum* defined by

$$S_0 = 0, S_i = max[0, \gamma S_{i-1} + w_i(X_i - k)], i = 1, 2, \ldots, N, \quad (1)$$

where γ is typically chosen in [0.7, 1] and the *reference value* $k \approx (\lambda_0 + \lambda_1)/2$ ("optimal" values for k are obtained based on behavior of likelihood ratios for X_i, see Hawkins and Olwell (1998)).

For schemes with DCO we define $S = max[S_1, S_2, \ldots, S_N]$ and refer to the current point in time as T. The data set (i.e., sequence of lifetime tests) is flagged at time T if $S > h$, where the threshold h is chosen based on

$$Prob[S > h | N, \lambda = \lambda_0] = \alpha_0, \quad (2)$$

where typically $\alpha_0 < 0.01$. Thus, test(1) is a series of repeated Cusum tests in the sense that at every new point in time the whole sequence (1) is re-computed from scratch.

We should note immediately that in many cases the scheme (1) alone is not sufficient for efficient detection of change in the process level, and so supplemental tests may be needed. For example, consider the situation where X_i is the replacement rate (number of replacements per 1000 MM for parts of vintage i) and w_i is the number of MM accumulated by parts of vintage i. As data for new vintages continues to come in, the value of the threshold h will be moving up (note that threshold violation could occur not only at the last vintage, but for some earlier vintages as well) - and this could desensitize the scheme with respect to more recent vintages, especially given the fact that weights w_i will tend to be large for early vintages and small for recent vintages. Use of $\gamma < 1$ will typically help somewhat to ameliorate the extreme manifestations of this problem - however, supplemental tests specifically geared towards detection of unfavorable events for later vintages are sometimes needed. Such tests will be discussed later in the section. Analogously to the case of conventional (non-DCO) schemes, use of $\gamma < 1$ is advisable when the primary mode of change in the process level is in the form of drifts rather than shifts.

The immediate computational problem spawned by (1) is related to obtaining h by solving (2). The algorithm needs to be efficient because in massive data monitoring systems (such as a warranty data system) the number of schemes run in parallel can easily reach 100,000 - and the computing operation typically needs to complete within a narrow time window. For procedures of type (1) Monte Carlo methods have proven to be effective, provided that certain measures are taken to enhance their efficiency. For example:

- Use parallel computation and elemental procedures (i.e., procedures designed to be applied to every element of an array simultaneously). In the PC warranty data application we use a set of K (typically, about 2000) simulated replications of the trajectory (1). For every replication a value of S is computed and h is obtained by solving (2) based on the empirical distribution of the K values of S. To reduce the computing time, the values of K replications for a given point in time are treated as a K-dimensional vector; its values are computed for $i = 1, 2, \ldots, N$, and the corresponding vector of S values is computed progressively in time. This is much more efficient than computing the statistics for every trajectory (1) and repeating the process K times. What helps us here is that the characteristics S of every trajectory (1) can be computed recursively in time i.

  A key element enhancing the efficiency of this vector-based operation is also related to generation of random variables X_i. Special algorithms of random variable generation, optimized for simultaneous production of K variables simultaneously, enable one to complete the computation of trajectories for each time i in an efficient manner.
- Take advantage of asymptotic properties of statistics S that can be derived on the basis of theory of stochastic processes. For example, one can expect that for distributions of X_i that have first two moments, the

distribution of S can be approximated based on the tail property

$$Prob[S > h|\lambda = \lambda_0] \sim A \times exp[-ah], h \to \infty, \quad (3)$$

where A is a constant and a is a function of the first two moments and γ. When $\gamma = 1$, relations of type (3) can typically be justified (for analysis of in-control situations only) based on the approximation of (1) by a Brownian motion with reflecting barrier at zero (e.g., see Cox and Miller (1977), Bagshaw and Johnson (1975)). Our experience suggests that (3) continues to hold even when $\gamma < 1$, though we have no proof of it at this time.

If for a given in-control situation it is known a-priori that (3) holds, one can reduce substantially the number of replications K that is needed to achieve the required level of accuracy. In particular, one can use a relatively low K, obtain the replications and obtain a non-parametric estimate of the upper 25-th percentile $\hat{q}_{75}$ of S. Subsequently, one can fit an exponential distribution to the excess points above $\hat{q}_{75}$, taking advantage of (3). Finally, the equation (2) is solved by using the pair of estimates $(\hat{q}_{75}, \hat{a})$. This approach can be used not only for threshold derivation, but also for computing the *severity* of an observed trajectory, expressed in terms of a *p-value* corresponding to S observed for the data set at hand. Note that test based on (1) effectively triggers an alarm if the p-value of the test falls below α_0.

- When *ancillary* statistics are available, try to *condition* on them in the course of Monte-Carlo replications, in line with usual recommendations of statistical theory. For example, suppose that X_i are rates per 1000 MM. Then the overall MM observed for a vintage, though a random quantity, can typically be assumed to have a distribution that does not depend on λ. It is, therefore, advisable to resample X_i based on the MM in period i that was actually observed in the data.

Supplemental Tests. As noted above, tests based on (1)-(2) could turn out to be insensitive to recent unfavorable changes - and such changes could be of high importance to the users. In the PC warranty data system we introduce the concept of *active component.* In particular, we set a threshold of, say, D_a days; a part is considered active if there are vintages present in the data within the last D_a days from the current point in time T. A warranty system will typically contain many inactive parts that are no longer produced (even though these parts are still in the field, continue to contribute data and could present a risk to the manufacturer). For inactive parts there is little benefit from emphasis on recent vintages because all the parts are anyway out of manufacturer's reach and so there is no longer an opportunity to prevent an escape of unreliable parts into the field. In contrast, for *active* parts such an opportunity does exist. Therefore, supplemental tests are applied to active parts only.

The first supplemental criterion for currently active part calls for testing the hypothesis that the collection of X_i observed within the last D_a days correspond to the in-control parameter not exceeding λ_0. The corresponding p-value can sometimes be computed numerically. The criterion triggers an alarm if this p-value is smaller than a pre-specified threshold.

The second supplemental criterion is based on the final value S_N of the scheme (1). The corresponding p-value is computed via Monte-Carlo method, by using the same runs that were used for the analysis of (1). Techniques described earlier for enhancing efficiency of the computations fully apply to this criterion; in many situations the relationship of type (3) can also shown to be relevant for it.

The overall severity (p-value) of the battery of (1) and two supplemental tests can be approximated by a function of the individual p-values $\psi(p_1, p_2, p_3)$ of these tests. Typically, correlation between the supplemental test statistics and S is negligible and can be ignored. However, the supplemental tests do tend to be correlated - and, therefore, one important issue here is obtaining a suitable form for ψ.

4 Monitoring of Wearout

One of issues of primary concern to engineers is onset of wearout. Such an event can substantially damage the reputation of a company - but when wearout occurs within the warranty period, this could lead to substantial additional losses and even to loss of the whole operation. Organizing an effective system for lifetime data monitoring that detects onset of wearout should, therefore, be a key priority for many manufacturing operations, especially those involved in mass manufacturing.

The computational strategy for wearout monitoring could be developed along the lines of that in Sec. 2. It is important to note, however, that in many situations it is unpractical to compute the wearout indicator for every vintage, for example, because of issues of parameter estimability. Therefore, data is typically grouped by vintages: for example, the rows of Fig. 1 are consolidated so as to yield just one row per month.

One can use several characteristics that are sensitive to wearout; any given characteristic can be used as a basis for a monitoring scheme of type (1). Given that (possibly after time scale transformation as mentioned in Sec. 1) the data under acceptable conditions is likely to show behavior that is "similar" to exponential, we could select an index that represents the estimate of the Weibull shape parameter. Note that we do not really have to believe that the data is Weibull: the estimated shape parameter can be used as a form of "wearout index" even in many situations where the lifetime distribution is non-Weibull, as it retains a substantial graphical and analytical appeal.

Denoting by c the shape parameter of the Weibull distribution we can specify the acceptable and unacceptable levels as c_0 and $c_1 > c_0$ (in many practical situations, $c_0 = 1$ is a good choice). One way to proceed is to compute consecutive *unbiased* estimates $\hat{c}_1, \hat{c}_2, \ldots, \hat{c}_M$, where M is the number of months for which data is available. These values are then used in the Cusum test:

$$S_{0w} = 0, S_{iw} = max[0, \gamma_w S_{i-1,w} + w_{iw}(\hat{c}_{iw} - k_w)], i = 1, 2, \ldots, M, \qquad (4)$$

where γ_w is the geometric parameter (typically close to 1) and $k_w \approx (c_0 + c_1)/2$. The weight w_{iw} is the number of failures observed in the period i (and not the overall number of MM for period i, as in (1)). This is because a period with large MM but very few failures does not contain much information about c). Now we can see why we want the sequence of estimates $\hat{c}_i$ to be bias-corrected: with such a choice we can use the same reference value k_w for every period. Since we operate under the DCO conditions, the decision statistics is $S_w = max[S_{1w}, S_{2w}, \ldots, S_{Mw}]$ and we flag the data set at time T if $S_w > h_w$, where h_w is chosen from the equation:

$$Prob[S_w > h_w | M, c_0] = \alpha_0. \qquad (5)$$

As before, Monte-Carlo approach is used to derive the threshold h_w and p-value of the test. In the course of replications we assume that the scale parameter β of the Weibull law can change from period to period, and we focus our attention entirely on c. Under such an assumption the number of fails w_{iw} in period i can be treated as an ancillary statistic for c and we condition the replications for this period on the number of fails being equal to w_{iw}. For period i we then estimate the scale parameter $\hat{\beta}_i$ under the hypothesized assumption $c = c_0$ and produce replications of w_{iw} failures under the assumption that lifetimes for this period are distributed Weibull $(\hat{\beta}_i, c_0)$, taking into account the censoring times. Processing such replications for our collection of periods enables one to evaluate the null distribution of S_w and obtain the corresponding threshold h_w and a p-value for the observed value of S_w. In some situations one may want to consider using a supplemental test (for active components) similar to the second supplemental test described above, i.e.: flag a component on wearout if the last value S_{Mw} of the sequence (4) exceeds a suitably chosen threshold (or if the associated p-value becomes sufficiently small). Such a test could help in catching a relatively recent onset of wearout somewhat earlier. One shought weigh the additional benefit of such a test against the necessity to raise thresholds for other tests, in order to maintain the target protection against false alarms. Such a supplemental test can be designed in parallel with the main test, by using the same simulation runs.

Note that a process similar to that described above can also be used to deploy schemes for monitoring the scale parameter β.

5 Conclusions

Design and analysis of systems for monitoring reliability, especially in the DCO environment is a highly complex task, both from the methodological and computational points of view. The main technical challenges include (a) establishing "on the fly" the thresholds for a large number of tests and efficient use of Monte Carlo techniques (b) establishing "newsworthiness" of the detected conditions based on p-values and similar indices of severity and (c) deployment in the field that satisfies the requirements for low rate of false alarms, detection capability, user interface and report generation. In this article we discuss a possible approach that was deployed several years ago at the IBM PC company. Our impression, based on user feedback, was that this approach can lead to usable and powerful system for monitoring massive streams of reliability and warranty data.

6 Acknowledgements

I would like to thank David L. Jensen, Ildar Khabibrakhmanov and Prince Stanley (IBM Thomas J. Watson Research Center) for help in development of the Early Detection Tool (EDT) that incorporates many of of the methods discussed in the paper. I am also thankful to Andrew Dubois, Stephen Restivo, Christopher Ross (IBM Personal Systems Division) for help with implementation and for their valuable input, and to Vaugh Robert Evans (IBM Server Group) for his generous feedback and advice in relation with deployment of the EDT and the methodology.

References

BAGSHAW, M. and JOHNSON, R. (1975): The Effect of Serial Correlation on the Performance of CUSUM TESTS II, *Technometrics, 17, 73-80.*

BLISCHKE, W.R. and MURTHY, D.N.P. (2000): *Reliability Modeling, Prediction and Optimization.* Wiley, New York.

COX, D. and MILLER, H. (1977): *The Theory of Stochastic Processes.* Chapman and Hall/CRC, New York.

DOGANAKSOY, N., HAHN, G.J. and MEEKER, W.Q. (2006): Improving reliability through warranty data analysis. *Quality Progress, November, 63-67.*

DUBOIS, A., EVANS, V.R., JENSEN, D., KHABIBRAKHMANOV, I., RESTIVO, S. ROSS, C. and YASHCHIN, E. (2008): System and method for early detection of system component failure. *US Patent 7401263*, US Patent Office.

HAWKINS, D. and OLWELL, D. (1998): *Cumulative Sum Charts and Charting for Quality Improvement.* Springer, New York.

KALBFLEISCH, J.D., LAWLESS, J.F. and ROBINSON, J.A. (1991): Methods for the analysis and prediction of warranty claims. *Technometrics 33, 273-285.*

KARIM, M.R., YAMAMOTO, W. and SUZUKI, K (2001): Statistical analysis of of marginal count failure data. *Lifetime Data Analysis, 7, 173-186.*

LAWLESS, J.F. (1998): Statistical Analysis of Product Warranty Data. *International Statistical Review 66, 41-60.*

LAWLESS, J.F. and KALBFLEISCH, J.D. (1992): Some issues in the collection and analysis of field reliability data. *Survival Analysis: State of the Art, J.P. Klein and P.K. Goel (eds.). Kluwer Academic Publishers, Netherlands, 141-152.*

OLTEANU, D. and VINING, G.G. (2009): CUSUM Charts for Censored Lifetime Data. *Paper at the 2009 Joint Statistical Meetings, Washington, DC.*

PATRA, K. and DEY, D.K. (2002): A general class of change point and change curve modeling for lifetime data. *Ann. Inst. Statist. Math., 54(3) 517-530.*

SEGO, L., REYNOLDS, M. and WOODALL, W. (2009): Risk-adjusted monitoring of survival times. *Statist. Med., 28, 1386-1401.*

STEINER, S.H. and MACKAY, R.J. (2000): Monitoring processes with highly censored data. *J. Qual. Technol., 32, 199-208.*

STEINER, S.H. and MACKAY, R.J. (2001): Detecting Changes in the Mean from Censored Lifetime Data. *Frontiers in Statistical Quality Control 6, H.J. Lenz and P.T. Wilrich eds., Physica-Verlag, 275-289.*

WU, H. and MEEKER, W.Q. (2002): Early detection of reliability problems using information from warranty databases. *Technometrics 44(2), 120-133.*

YASHCHIN, E. (1993): Statistical control schemes: methods, applications and generalizations. *International Statistical Review 61, 41-66.*

ZHANG, L. and CHEN, G. (2004): EWMA charts for monitoring the mean of censored Weibull lifetimes *J. of Quality Technology 36(3), 321-328.*