LA-UR- 09-04667

Approved for public release; distribution is unlimited.

Title:	A Random Onset Model for Degradation of High-Reliability Systems
Author(s):	Scott Vander Wiel, Alyson Wilson, Todd Graves and C. Shane Reese
Intended for:	Compatibility, Aging and Stockpile Stewardship Conference, Sandia National Lab, Albuquerque

Los Alamos NATIONAL LABORATORY EST. 1943

A Random Onset Model for Degradation of High–Reliability Systems

Scott Vander Wiel[1], Alyson Wilson[2], Todd Graves[1] and C. Shane Reese[3]
[1] Los Alamos National Laboratory, MS F600, Los Alamos, NM 87545
[2] Iowa State University, Dept. of Statistics, 2105 Snedecor, Ames, IA 50011
[3] Brigham Young University, Dept. or Statistics, 208 TMCB, Povo, UT 84602

Introduction

Both the U. S. Department of Defense (DoD) and Department of Energy (DOE) maintain weapons stockpiles: items like bullets, missiles and bombs that have already been produced and are being stored until needed. Ideally, these stockpiles maintain high reliability over time. To assess reliability, a surveillance program is implemented, where units are periodically removed from the stockpile and tested. The most definitive tests typically destroy the weapons so a given unit is tested only once. Surveillance managers need to decide how many units should be tested, how often they should be tested, what tests should be done, and how the resulting data are used to estimate the stockpile's current and future reliability. These issues are particularly critical from a planning perspective: given what has already been observed and our understanding of the mechanisms of stockpile aging, what is an appropriate and cost-effective surveillance program?

Surveillance programs are costly, broad, and deep, especially in the DOE, where the U. S. nuclear weapons surveillance program must ``ensure, through various tests, that the reliability of nuclear weapons is maintained" in the absence of full-system testing (General Accounting Office, 1996). The DOE program consists primarily of three types of tests: nonnuclear flight tests, that involve the actual dropping or launching of a weapon from which the nuclear components have been removed; and nonnuclear and nuclear systems laboratory tests, which detect defects due to aging, manufacturing, and design of the nonnuclear and nuclear portions of the weapons. Fully integrated analysis of the suite of nuclear weapons surveillance data is an ongoing area of research (Wilson et al., 2007). This paper introduces a simple model that captures high-level features of stockpile reliability over time and can be used to answer broad *policy* questions about surveillance programs. Our intention is to provide a framework that generates tractable answers that integrate expert knowledge and high-level summaries of surveillance data to allow decision-making about appropriate trade-offs between the cost of data and the precision of stockpile reliability estimates.

A Random Onset Degradation Model

Though expectations are that DoD and DoE stockpiles will maintain high reliability over time, it is nevertheless important to sample and test units regularly to maintain confidence that aging effects have not begun to degrade reliability. Weapon systems surveillance plans have traditionally been designed to provide a specified level of confidence in detecting problems that affect certain proportions of the stockpile within a certain period of time. For example a "90/10/2" plan assesses 11 units per year with a justification that this provides a 90% chance of detecting, within a two year period, a problem that affects 10% of the stockpile. The 90/10/2 rationale is a justification for certain sampling rates but it does not attempt to anticipate when stockpile reliability might begin to degrade or how quickly aging might progress and it is not a method for estimating stockpile reliability at a given point in time or for measuring the uncertainty of a given reliability estimate. However, these are precisely the questions that a decision-maker needs answered in an era of cost-cutting to determine, for example, how reliability estimates will be affected if surveillance sampling is curtailed for a few years or ramped down to lower levels. Pressure to lower sampling rates often relies on the rationale that if no defects have been found after many years of testing then a few more years of data will do little to improve one's confidence that reliability is high. This reasoning implicitly relies on a static model of the stockpile and ignores the potential that aging effects will eventually degrade reliability.

We consider a statistical model for stockpile reliability with the following features.

- Reliability in the year of manufacture is high.
- Reliability in one year is typically the same as the previous year, but certainly no better.
- In any given year there is a small chance that reliability will begin to degrade.
- Once degradation starts, it continues at a certain rate.
- Units are routinely drawn from the stockpile for destructive testing; individual units either pass or fail.

Taken together, these features describe a stockpile that will probably maintain high reliability from one year to the next under an ever-present possibility that reliability could begin to decline at any time. Under these conditions, it is clear that ongoing sampling and testing are imperative to maintain confidence that a stockpile has high reliability The random onset model described herein provides a means for quantitatively assessing confidence in stockpile reliability where the potential impact of aging explicitly contributes to uncertainty.

Figure 1 illustrates the main characteristic of the random onset model: *Uncertainty grows large once you stop looking.*

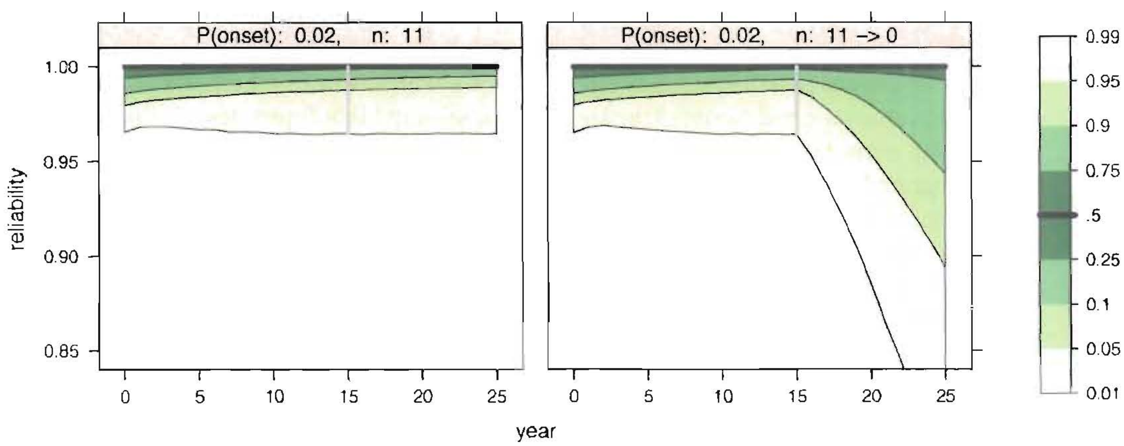

Figure 1: Reliability estimates over time when no failures are found in surveilled units. In both panels 11 units per year are surveilled from year zero through year 15. This continues through year 25 in the left panel but sampling ceases at year 15 in the right panel. Curtailment of sampling does not impact the median estimate (thick black line at reliability 1.0) but dramatically affects the uncertainty bounds (shaded regions). Shading shows quantiles of the reliability distribution as indicated by the key. Quantiles above the median are pegged at 1.0 and are therefore not visible.

The figure plots median reliability estimates (black lines at 1.0) and various uncertainty bounds (boundaries of shaded regions) that would be obtained under a nominal sampling plan of 11 units per year (left panel) and in a situation when sampling ceases completely after year 15 (right panel). All units sampled are assumed to pass when tested. Complete curtailment of the surveillance plan has no impact on the median estimate of reliability (which remains at 1.0

throughout the 25 year period) but it has an immediate and progressive impact on the uncertainty about reliability. If no units are sampled after the 15th year uncertainty bounds begin to drop at rates of approximately 1% per year.

Input specifications used to produce Figure 1 can be summarized as follows: (a) initial reliability of the stockpile is most likely greater than 99%; (b) detrimental aging could occur at any time but is not expected for 50 years; and (c) once degradation begins, it will continue at an unknown rate that is expected to be around 1% of the current reliability level each year.

Figure 2 shows 100 sample paths (left panel) of plausible stockpile reliabilities over a 25 year period generated by the random onset model with the given input specifications. Various quantiles (right panel) of the trajectory distribution are also show, based on a large number of simulated paths. Each sample path in the left panel represents a possible trajectory for the probability that a random unit will function when tested. Incorporating successful test results into the analysis converts the input assumptions illustrated in Figure 2 into the posterior assessments of reliability shown Figure 1. The need for surveillance testing to control uncertainty is obvious.

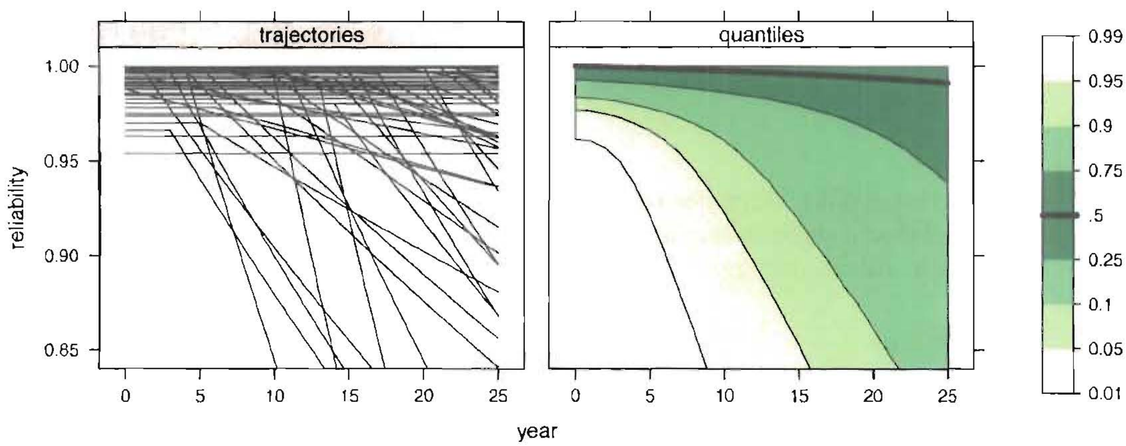

Figure 2: One hundred simulated reliability trajectories (left) and quantiles (right) of the reliability distribution used as input to the random onset model.

The model allows for a variety of reliability degradation patterns including early or late onset and slow or fast progression. In the large majority of realizations the reliability remains above 99% for at least a 10 years, but in a few cases, it begins to degrade rather steeply within that period. If these patterns represent a credible description of the various ways in which stockpile reliability might degrade, then fresh tests with no failures are required to assure that reliability remains high over time.

Failure Scenario

The random onset model can be used to play out a variety of "what if" scenarios. For example, suppose that following 10 years of sampling 11 units per year with no trouble found, a single failure is observed among the 11 units tested in year 11. Figure 3 shows a snapshot of reliability estimates over time with projections shown after year 11 assuming that no further surveillance is conducted. With a single failure, the median of the predictive posterior drops to about 0.99 but then remains nearly flat through year 25, consistent with age-related degradation of reliability remaining relatively unlikely, even though a defective unit has been observed. This demonstrates that the dynamic effects of aging in the random onset model have limited impact on the median of the predictive posterior distribution, similar to what would be obtained from a standard analysis method that assumes reliability is constant over time. On the other hand, the lower tails of the reliability distribution degrade at rates between 1% and 3% per year, consistent with the possibility that the defective test result may be indicative of age related reliability degradation, even if such degradation has not explicitly been demonstrated by understanding the root cause of the failure.

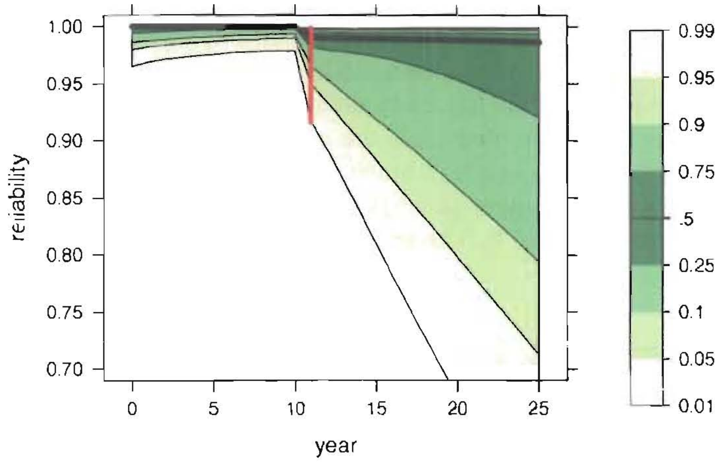

Figure 3: Reliability estimates when 11 units are surveilled each year with no failures for 10 years and then a single failure in year 11. Projections are shown after year 11 assuming no further surveillance testing.

Conclusions

We have described a random onset model for stockpile reliability and shown that "what-if" scenarios can be used to demonstrate how confidence is affected by changes in sampling rates and failures in testing. This model captures the main features of uncertainty in stockpile reliability and is useful to answer broader policy questions about surveillance programs. It is clear that ongoing sampling and testing are imperative to maintain confidence that a stockpile has high reliability. The random onset model described herein provides a means for quantitatively assessing confidence in stockpile reliability. In addition, we believe that this framework is flexible enough to allow for extensions to more complex situations such as aggregation of component reliabilities to system-level reliability and mixed age populations.

Acknowledgement.

The authors would like to thank the Enhanced Surveillance Campaign for support to conduct this research and Mike Hamada for helpful discussions and interfacing with the surveillance community at LANL.

References

General Accounting Office (1996). Report to the Committee on Armed Services, U. S. Senate. Nuclear Weapons: Improvements Needed to DOE's Nuclear Weapons Stockpile Surveil lance Program. Washington, DC. GAO/RCED-96-216.

Wilson, A., Hamada, M., Anderson-Cook, C., Graves, T., Vander Wiel, S., Sims, B., and Wilson, G. (2007). Statistical models for stockpile health assessment. In ADTSC Nuclear Weapons Highlights 2007. Los Alamos National Laboratory. LA-LP-041, http://www.lanl.gov/source/orgs/adtsc/nw highlights 2007.php.