



INSTITUTE FOR DEFENSE ANALYSES

**Bayesian Hierarchical Models for
Common Components Across Multiple
System Configurations**

Kassandra Fronczyk
Rebecca Dickinson
Alyson Wilson
Caleb Browning
Laura Freeman

June 2015

Approved for public release;
distribution is unlimited.

IDA Document NS D-5514

Log: H 15-000545



The Institute for Defense Analyses is a non-profit corporation that operates three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.

About This Publication

In this paper, we use a Bayesian hierarchical model to assess the reliability of the Joint Light Tactical Vehicle (JLTV), which is a family of vehicles. The proposed model effectively combines information across three phases of testing and across common vehicle components. The analysis yields estimates of failure rates across failure modes and vehicles as well as an overall estimate of the failure rate for the family of vehicles. We are also able to obtain estimates of the Fix Effectiveness Factor (FEF), a measure of how well repairs between test phases improve failure rates. In addition to using all data to improve on current assessments of reliability and reliability growth, we illustrate how to leverage the information learned from the three phases to determine the length of testing needed to demonstrate a given reliability threshold in subsequent testing.

Copyright Notice

© 2015 Institute for Defense Analyses
4850 Mark Center Drive, Alexandria, Virginia 22311-1882 • (703) 845-2000.

This material may be reproduced by or for the U.S. Government pursuant to the copyright license under the clause at DFARS 252.227-7013 (a)(16) [Jun 2013].

Bayesian Hierarchical Models for Common Components Across Multiple System Configurations

**Kassandra Fronczyk*, Rebecca Dickinson, Alyson
Wilson, Caleb Browning, Laura Freeman**

Mathematical Methods in Reliability

Tokyo, Japan

June, 2015

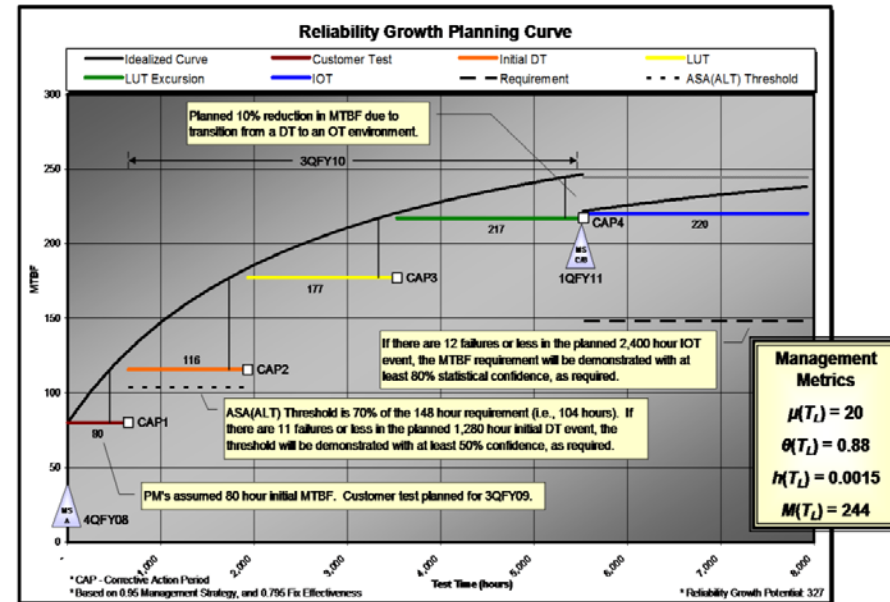
**NC STATE
UNIVERSITY**

IDA

- **Background:**
 - The JLTV Family of Vehicles (FoV) is replacing the legacy Humvee vehicle fleet.
 - » Eight prototypes have gone through a series of three testing phases.
 - Reliability and Reliability Growth are a high priority in testing – the test plan required vendors conduct 160,000 miles of reliability testing!
 - » It is important that the capabilities and limitations of the system be understood.
 - **Purpose of our Case Study:**
 - How do we leverage all data to assess the reliability/reliability growth?
 - How do we use the observed data to scope a future test plans?
 - **Methods & Results:**
 - Bayesian Hierarchical Model
 - » Combines the data from all test phases, vehicles, and failure modes to obtain data driven estimates of reliability and reliability growth at multiple levels simultaneously.
 - Assurance Testing
 - » Leverages all information about reliability and growth of the FoV to reasonably size a test while accounting for both consumer and producer risk.
 - **Future Directions:**
 - Unknown number of failure modes
 - Exponential distribution assumption
 - Incorporate meaningful Covariates
-

- Reliability Growth – models changes and improvement (we hope!) in the reliability of a system as it moves through testing phases and undergoes corrective action periods (CAP)
- Popular Growth Models used in the DoD:
 - Duane Model
 - Crow-AMSAA
 - AMSAA PM2
- Planning Parameters directly influenced by Program Management
 - Initial system MTBF
 - Management Strategy
 - Goal system MTBF
 - Average fixed effectiveness factor (FEF) of corrective actions
 - Duration of developmental testing

AMSAA Projection Methodology (PM2)



Reliability growth planning models used in the DoD are not based on test data!

- **Family of Vehicles designed to replace the Legacy Humvee Fleet.**
 - Tasked with an Impressive Mission:
 - » System should provide ground mobility that is deployable worldwide and capable of operating across the range of military roles (i.e. combat, sustainment, police action, peace-keeping, and security patrol) , in all weather and terrain conditions.



AM General JLTV



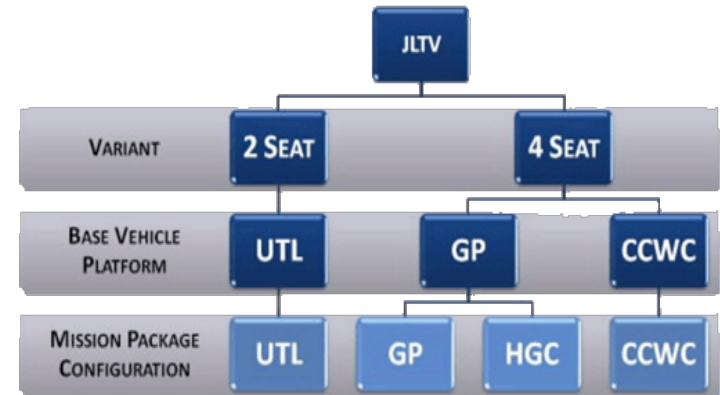
Lockheed Martin systems JLTV



Oshkosh Corporation JLTV

- **The Family of Vehicles is Comprised of Two Variants:**

- Two-seat variant Combat Support Vehicle
 - » Utility Vehicle (UV)
- Four-seat variant Combat Tactical Vehicle.
 - » Close Combat Weapons Carrier (CCWC)
 - » General Purpose Vehicle (GP)
 - » Heavy Guns Carrier (HGC)



- **What do we know about the Vehicles?**

- Vehicle Variants were designed to have a very high degree of commonality.
 - » Example: Brakes and Radios

- **Purpose of Testing**
 - Discover failure modes, implement corrective actions, and assess whether the vendor's vehicles could meet the required Mean Miles Between Failure (MMBF)
- **Three Phases of Developmental Testing (DT1, DT2, DT3)**
 - For every vehicle, each failure encountered during testing was recorded and attributed to a specific failure mode.
 - » There are 26 observed failure modes across the three phases of testing.
- **Test Plan Called for 160,000 Miles of Testing (20,000 per vehicle)**
 - Phase 1: 6,000 miles
 - Phase 2: 6,800 miles
 - Phase 3: 8,700 miles
- **Testing**
 - Aberdeen Proving Ground, Maryland
 - Yuma Proving Ground, Arizona



- **Operational Mission Failure (OMF):** failure discovered during mission execution that result in an abort or termination of a mission in progress
 - » Reliability requirements are typically written in terms of OMFs.
- **Essential Function Failures (EFF):** failures of mission essential components. By definition all OMFs are EFFs
 - » EFFs include a large portion of the failure modes that drive maintenance costs and reduce system availability
 - » Across all eight vehicles tested and the three test phases, there are only 91 OMFs compared to 1,321 EFFs.
- **Comparing EFFs and OMFs**
 - » Engine: temporary power failure vs. not starting at all
 - » Steering: excessive pulling in one direction vs. vehicle rolling
 - » Brakes: brake fluid leak/line worn vs. brake lock up

Combining failures provides a more robust reliability estimate.

- Each test phase (and each vehicle type) independently and uses the exponential distribution to model the miles between failures. Failure mode is ignored.
- Requirements are written at the FOV level 2,400 MMBOMF. To assess if the FoV meets the requirement, the miles from all vehicles and the number of OMFs across vehicle and failure mode are pooled together.
- Reliability is expressed in terms of the mean number of miles between an operational mission failure (MMBOMF):

$$\widehat{\text{MMBOMF}} = \frac{\text{Total Miles Driven}}{\# \text{ of OMF Failures}}$$

An Overly Simplistic Analysis!

Modeling vehicle failure miles

Estimating a Fixed Effectiveness Factor for the two observed CAPs

$$t_{DT_1} \sim \exp(\lambda_{ij}), \quad t_{DT_2} \sim \exp(\lambda_{ij}(1 - \rho_{1j})), \quad t_{DT_3} \sim \exp(\lambda_{ij}(1 - \rho_{1j})(1 - \rho_{2j}))$$

Estimating a failure rate for each vehicle and failure mode

$i = 1, 2, \dots, 8$ (vehicle variants) $j = 1, 2, \dots, 26$ (failure modes)

Not Common
vs.
Common

$$\lambda_{ij} \sim \text{gamma}(a, b)$$

$$\lambda_{ij} = \lambda_j \sim \text{gamma}(a, b)$$

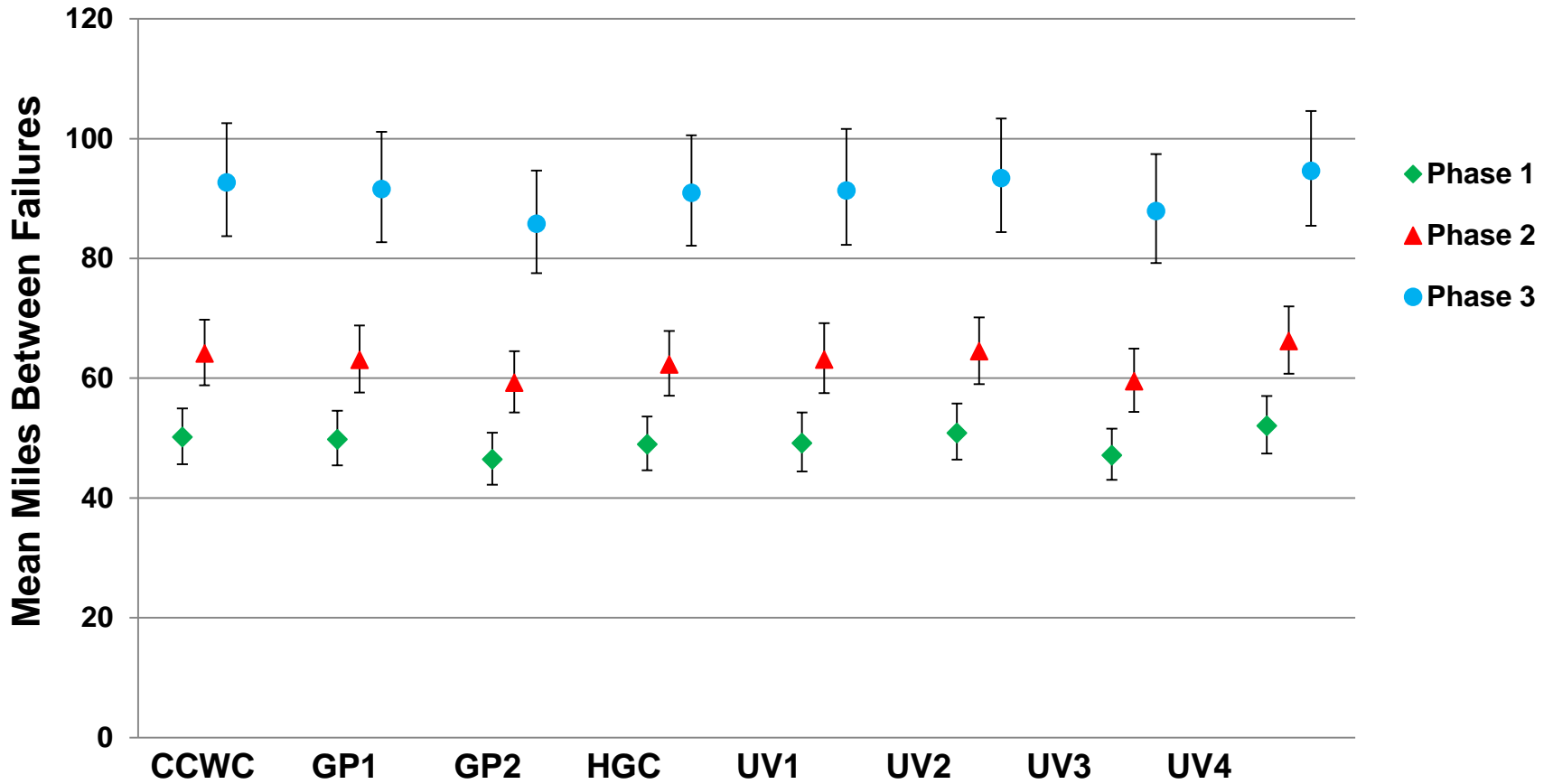
$$\rho_1 \sim \text{beta}(1, 1) \quad \rho_2 \sim \text{beta}(1, 1)$$

$$a \sim \text{gamma}(.001, .001), \quad b \sim \text{gamma}(.001, .001)$$

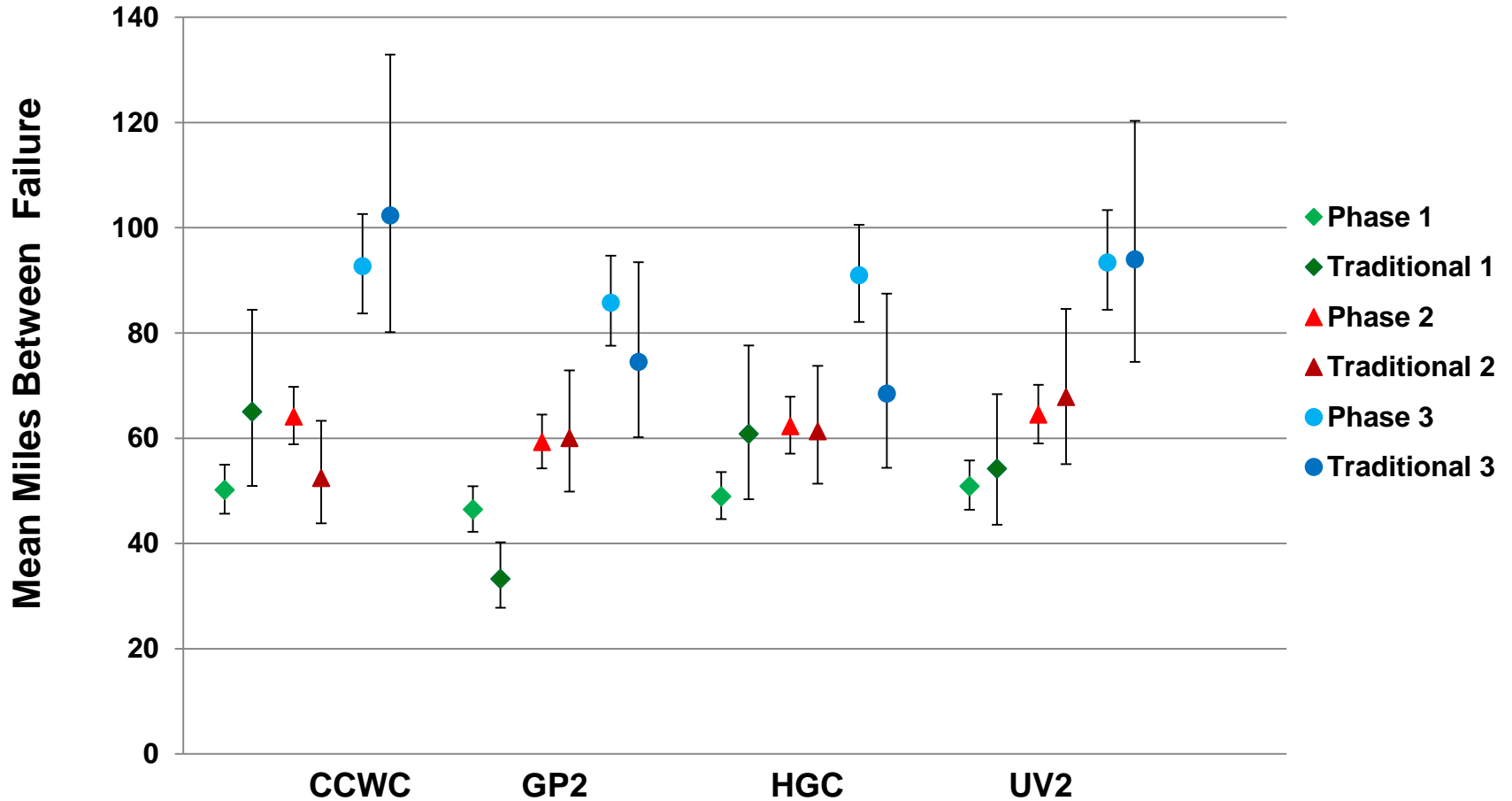
The number of failure modes is assumed fixed and known *a priori*

- **Estimate of a vehicle's failure rate**
 - Phase 1: $\sum_j \lambda_{ij}$
 - Phase 2: $\sum_j \lambda_{ij} (1 - \rho_{1j})$
 - Phase 3: $\sum_j \lambda_{ij} (1 - \rho_{1j})(1 - \rho_{2j})$
- **Estimates of a failure mode rates**
 - Common failure modes: λ_j
 - Not Common failure modes: λ_{ij}
- **Estimate of the fixed effectiveness factor (FEF) between phases 1 and 2.**
 - How does a data driven estimate actually compare to the assumed average FEF of 0.70?

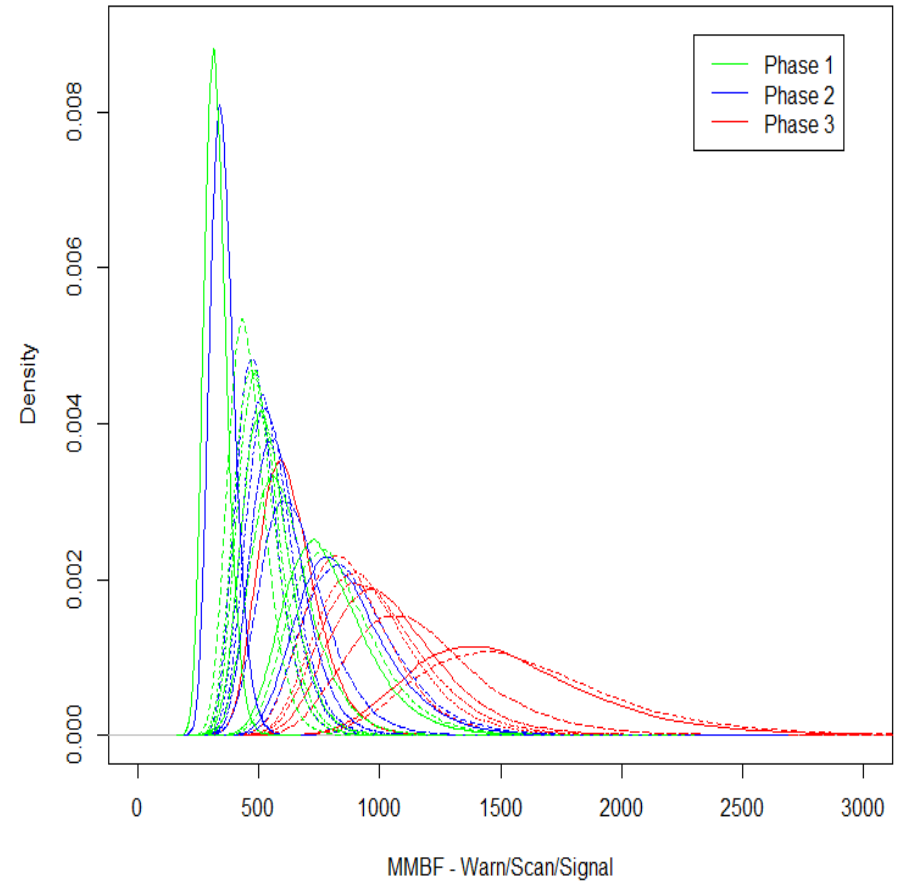
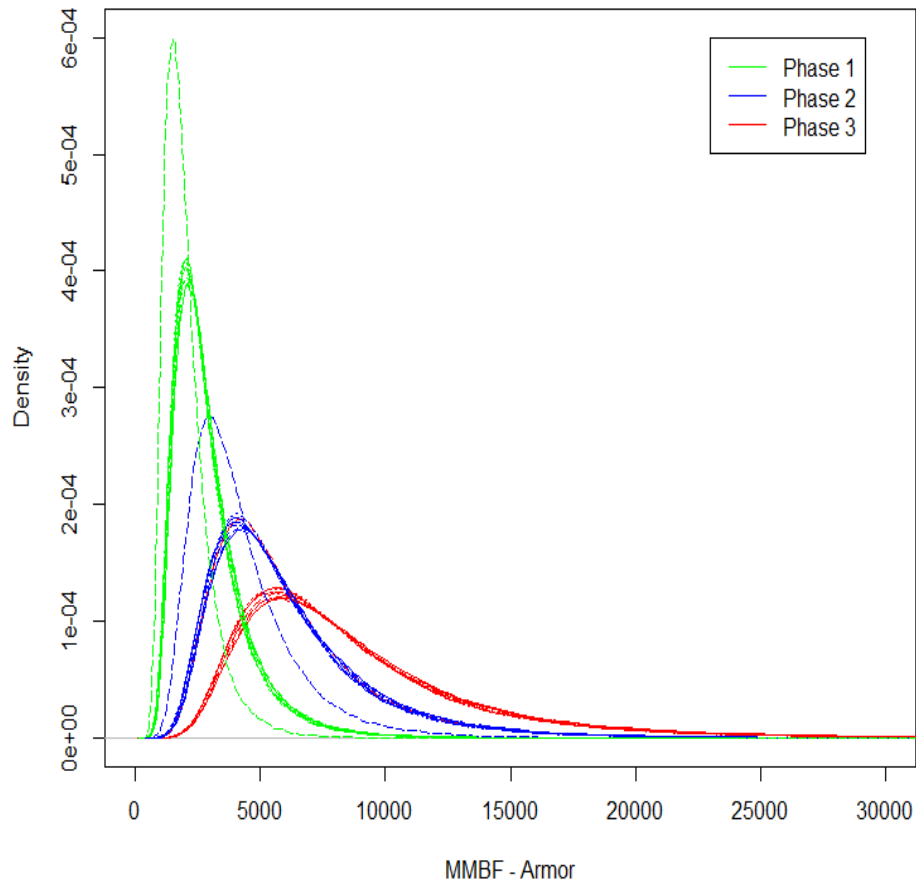
MMBF Estimates by Phase
(95% Credible Intervals)



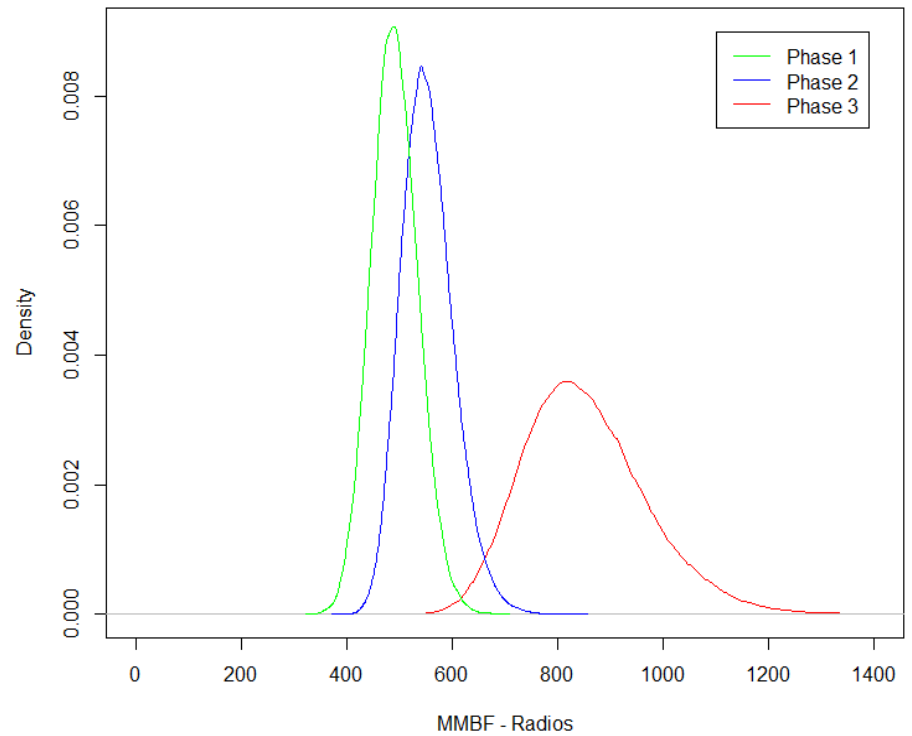
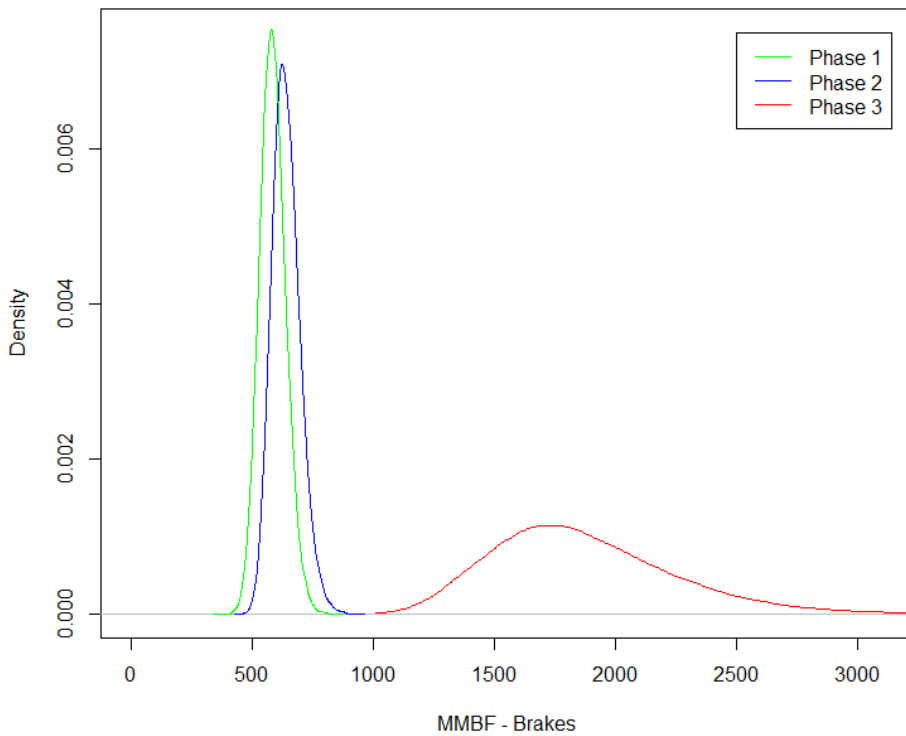
MMBF Estimates by Phase
(95% Credible Intervals)



Results: MMBF of Failure Modes

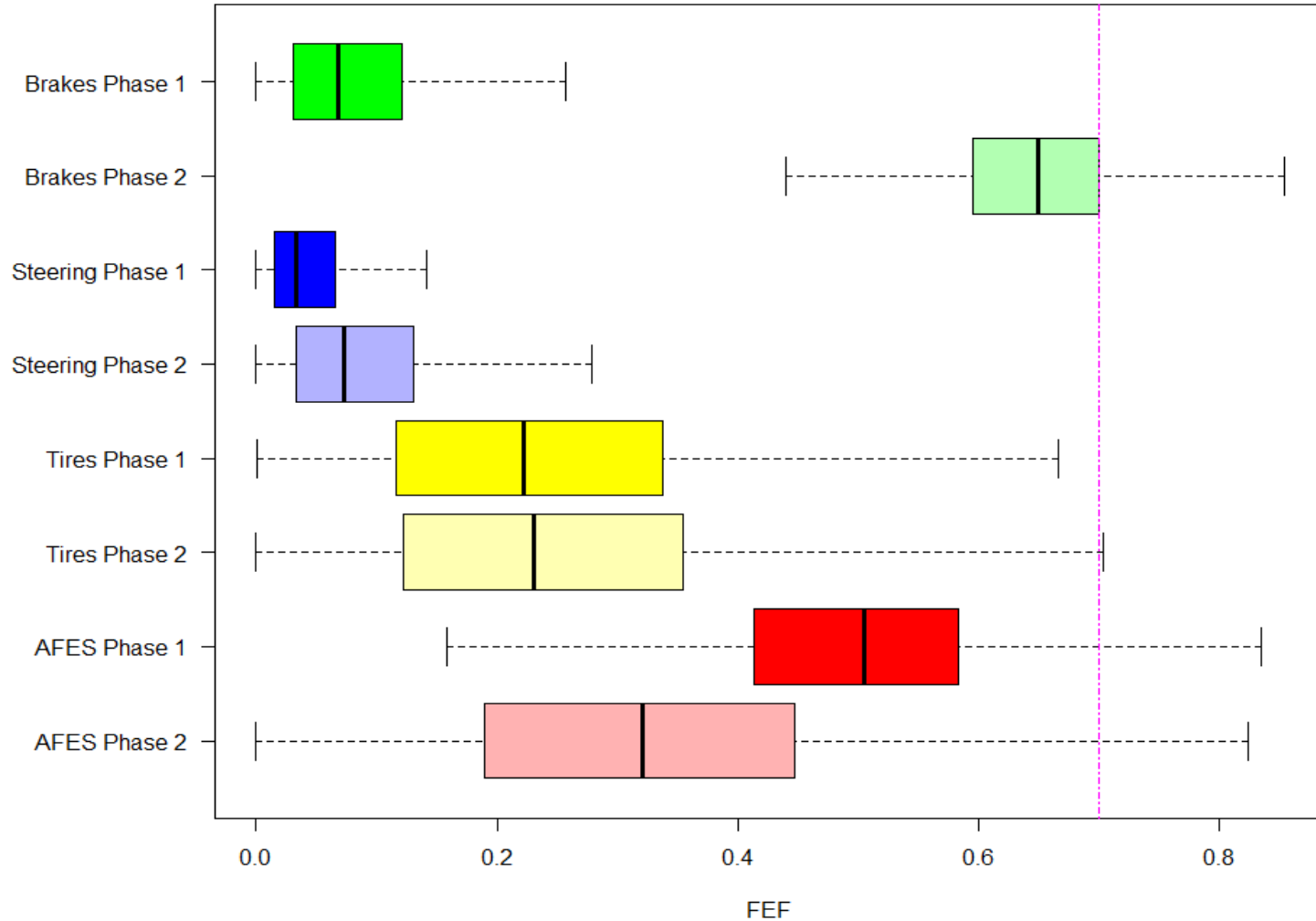


Results: MMBF of Common failure modes



Results: Fixed Effectiveness Factor

Commonly Assumed Average FEF is 0.7 !!!



Objective

- Scope an appropriately sized Operational Test (OT) using the demonstrated reliability and growth of the JLTV FoV in the three DT phases.
- If our reliability-quantity of interest is mean time between failures (MTBF) then
 - » How many miles do we need to drive?
 - » And how many failures are allowable for a successful test?

Reliability Demonstration or Reliability Assurance?

Demonstration Test

- A classical hypothesis test, which *uses only data from the test* to assess whether reliability requirements are met - often requires an exorbitant amount of testing!

Assurance Test

- *Combines information from various sources* to reduce the amount of testing required to meet a requirement.

- Two Risk Criteria in Determining a Test Plan

- Consumer's Risk

$$P(\text{Test is Passed} | MTBF_c) \leq \beta_c$$

Classical Risk Criteria

$$P(MTBF \leq MTBF_c | \text{Test is Passed}) \leq \beta_c$$

Bayesian Risk Criteria

- Producer's Risk

$$P(\text{Test is Failed} | MTBF_p) \leq \alpha_p$$

$$P(MTBF \geq MTBF_p | \text{Test is Failed}) \leq \alpha_p$$

Considering one Failure Mode:

- Automatic Fire Extinguishing System (AFES) - a common failure mode across all vehicles.

$$\eta_{AFES} = \underbrace{\lambda_{AFES}(1 - \rho_1)(1 - \rho_2)}_{\text{draws from the posteriors}} \theta$$

draws from the posteriors

$$MTBF_{AFES} = \frac{1}{\eta_{AFES}}$$

A degradation factor. It is common to see a 10-30% reduction in reliability from DT to OT. We put a beta prior on θ with most of the mass between 0.10 and 0.30.

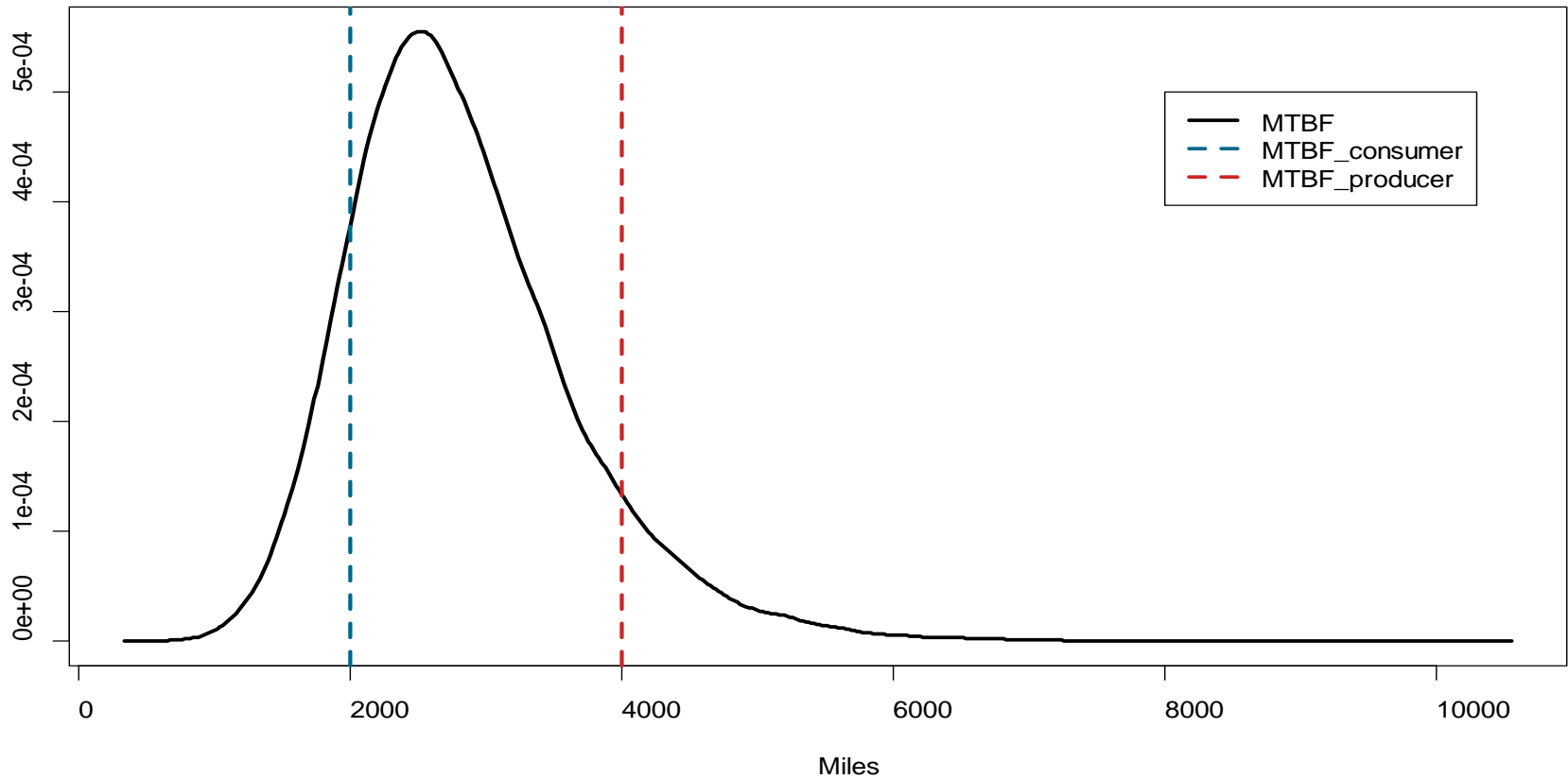
Posterior Consumer Risk

$$P(MTBF \leq MTBF_c | \text{Test is Passed}) \approx \frac{\sum_{j=1}^N \left[1 - \sum_{y=0}^c (t_0 \eta^{(j)})^y e^{-t_0 \eta^{(j)}} \right] I \left(\eta^{(j)} \leq \frac{1}{MTBF_c} \right)}{\sum_{j=1}^N \left[1 - \sum_{y=0}^c (t_0 \eta^{(j)})^y e^{-t_0 \eta^{(j)}} \right]} \leq \beta_c$$

Posterior Producer Risk

$$P(MTBF \geq MTBF_p | \text{Test is Failed}) \approx \frac{\sum_{j=1}^N \left[1 - \sum_{y=0}^c (t_0 \eta^{(j)})^y e^{-t_0 \eta^{(j)}} \right] I \left(\eta^{(j)} \leq \frac{1}{MTBF_p} \right)}{\sum_{j=1}^N \left[1 - \sum_{y=0}^c (t_0 \eta^{(j)})^y e^{-t_0 \eta^{(j)}} \right]} \leq \alpha_p$$

Density of MTBF: Automatic Fire Extinguishing System



IDA Results: Comparison to Traditional Test Plan

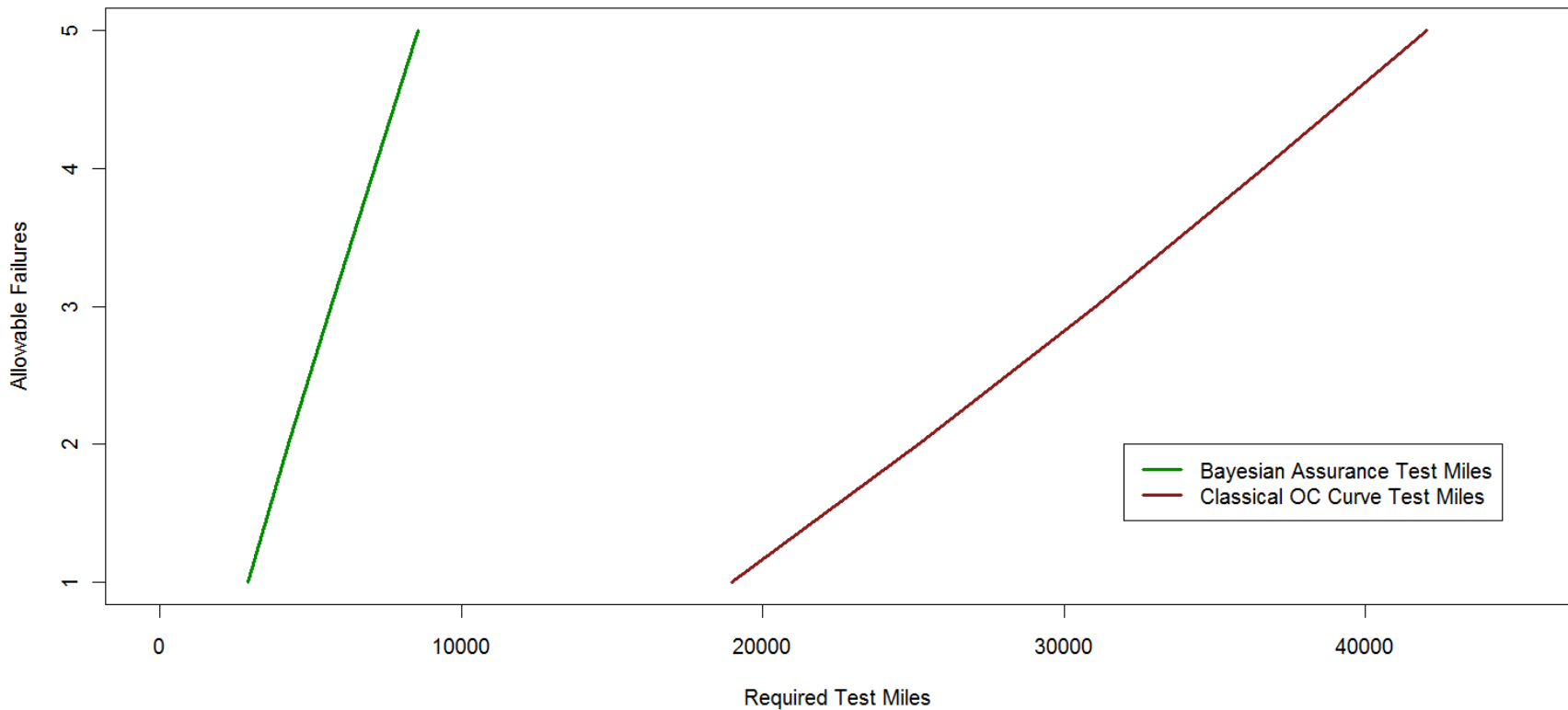
- A traditional test plan approach in the DoD fixes consumer risk at the requirement (e.g., $\beta = 0.05$ for a MTBF = 2000)
 - Plans the minimum test around a fixed number of failures:
 - $Test\ Duration = Req * \left(\chi^2_{1-\alpha, N_f+2}\right)^{-1} / 2$
 - Ignores producer risk

Failures Allowed	Bayesian Assurance Test Miles ($\beta_c = 0.10, \alpha_p = 0.05$)	Classical OC Curve Miles ($\beta_c = 0.10, \alpha_p = ?$)
1	2,940	7,780 ($\alpha_p = 0.58$)
2	4,280	10,645 ($\alpha_p = 0.50$)
3	5,680	13,362 ($\alpha_p = 0.43$)
4	7,120	15,988 ($\alpha_p = 0.37$)
5	8,580	18,550 ($\alpha_p = 0.32$)

Compared to traditional methods – the Assurance based approach reduces test duration and controls producers risk.

- For matching risk criteria, the Bayesian assurance methodology provides a defensible method for justifying significantly shorter tests.

Comparing Test Length: $\beta = 0.10, \alpha = 0.05$



- **Presented a proof of concept ... more work to be done.**
 - Constructed a reliability assurance test using a single failure mode (common across vehicles).
 - Next Steps: Construct an assurance test for the JLTV FoV
 - » For each vehicle?
 - » For a variant (2 seat and 4 seat) ?
 - » For a location?

- **Unknown number of failure modes**
 - In a given test phase, every vehicle is not guaranteed to have a failure of all 26 failure modes
 - A new failure mode code be discovered in a future test phase
- **Exponential distribution assumption**
 - Assess the fit of our distributional assumptions
- **Incorporate Covariates**
 - Vehicle Variant
 - » Two Seat vs Four Seat
 - Test Site
 - » Difficulty of Terrain
 - » Weather Conditions

Hamada, Michael S., et al. *Bayesian reliability*. Springer Science & Business Media, 2008.

Meeker, William Q., and Luis A. Escobar. "Reliability: the other dimension of quality." *Quality Technology and Quantitative Management* 1.1 (2004): 1-25.