Bayesian Hierarchical Models for Common Components Across Multiple System Configurations

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

October 21, 2015

NC STATE UNIVERSITY

IDA

Big Picture

- **Background:**
  - A future combat Family of Vehicles (FoV) is being designed to be deployable worldwide for many mission types with a high degree of commonality
    - Four prototypes have gone through a series of three testing phases.
  - Reliability and Reliability Growth are a high priority in 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
    - Leverages 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

- **Reliability is an essential component of the assessment of operational suitability of major defense systems.**

- **In the DoD, reliability is typically expressed in terms of the mean number of miles between an operational mission failure (MMBOMF):**

  – Operational Mission Failure (OMF): failure discovered during mission execution that result in an abort or termination of a mission in progress
    » 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

- **Comparing EFFs and OMFs**
    » 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.

Testing the Family of Vehicles

- **The Family of Vehicles is comprised of four types, built with a high degree of commonality**
  - Utility Vehicle (UV), General Purpose Vehicle (GP)
  - RAM testing at two test locations
- **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.
  - Between each DT phase, there is a Corrective Action Period (CAP) for the program to make fixes and (hopefully!!) improve the vehicles' reliability

Bayesian Hierarchical Model

Modeling vehicle failure miles

Estimating a change in the failure rate after the two CAPs

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

Estimating a failure rate for each vehicle and failure mode

$i = 1,2,\ldots,4$ (vehicles) $j = 1,2,\ldots,26$ (failure modes)

Not Common vs. Common

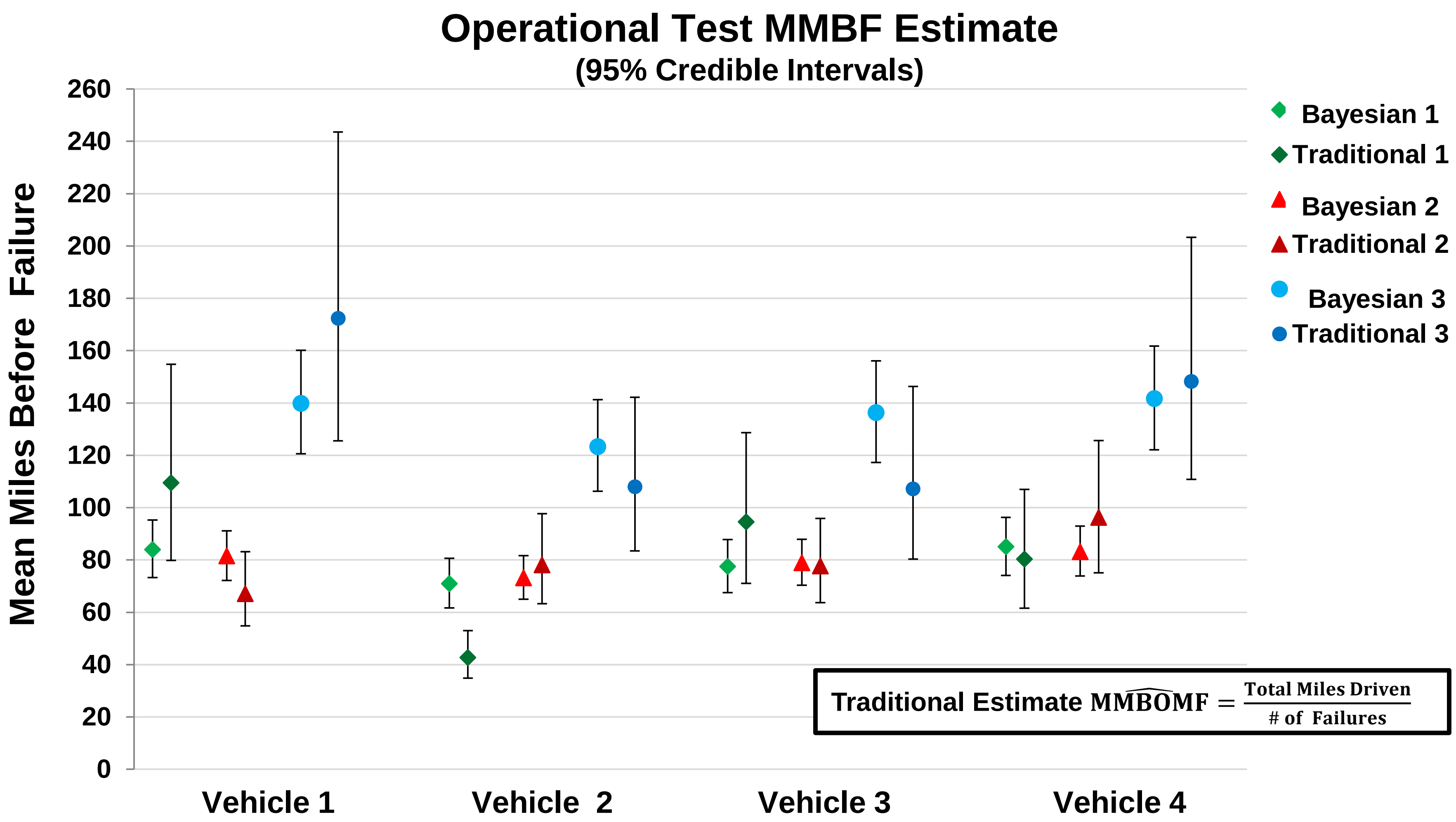
$$\lambda_{ij} \sim gamma(a,b)$$
$$\lambda_{ij} = \lambda_j \sim gamma(a,b)$$

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

$$\rho_1 \sim gamma(c,d) \quad \rho_2 \sim gamma(c,d)$$

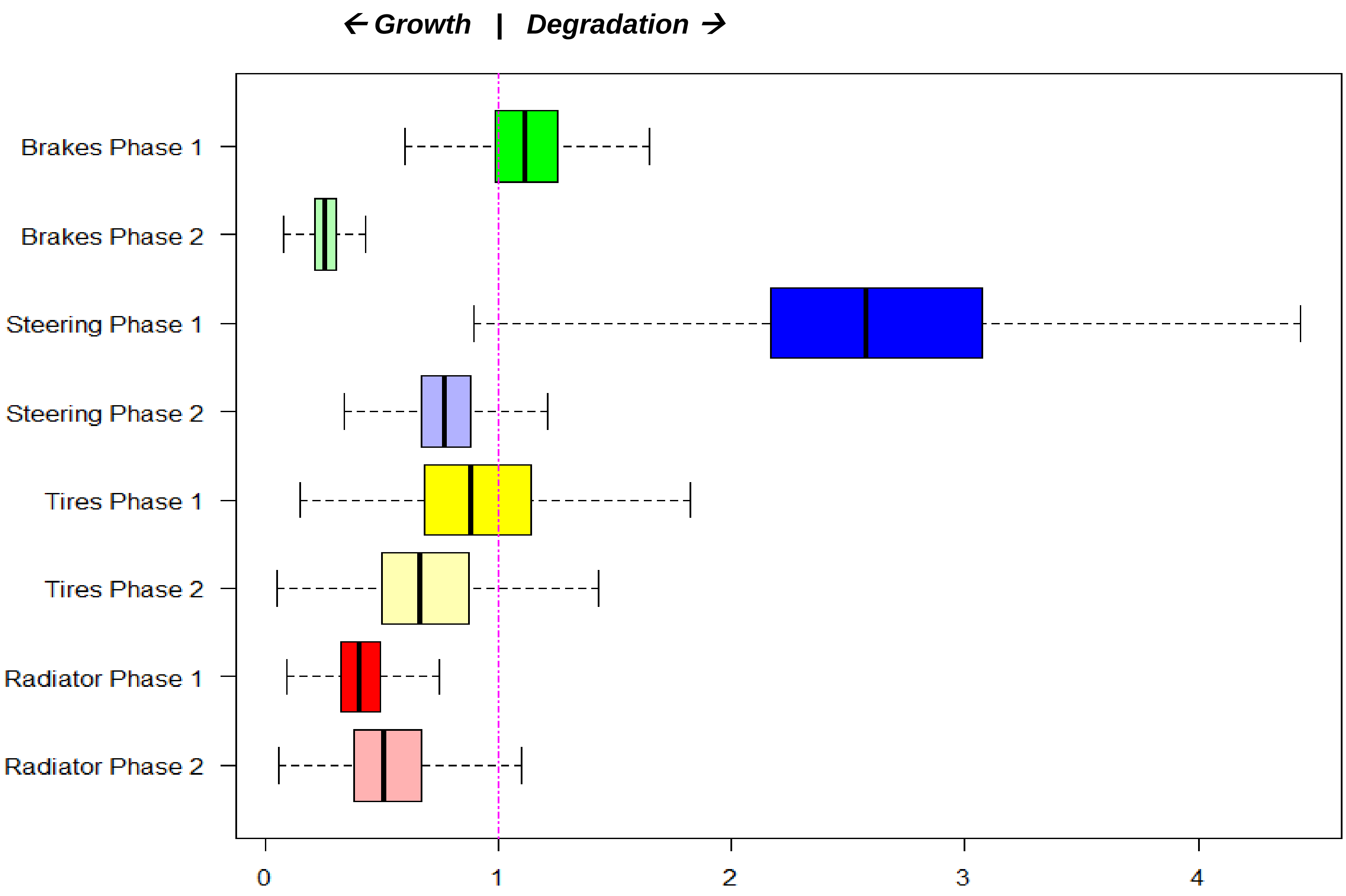
$$a \sim gamma(.001,.001)\,,\; b \sim gamma(.001,.001)$$
$$\mathrm{c} \sim gamma(.001,.001)\,,\; d \sim gamma(.001,.001)$$

Results: Comparing Vehicle MMBF Estimates



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

Results: Fix Effectiveness Factor



Building a Test Plan for OT

Objective

- Scope an appropriately sized Operational Test (OT) using the demonstrated reliability and growth of the FoV in the three DT phases.
- If our reliability-quantity of interest is mean miles between failures (MMBF) 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

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

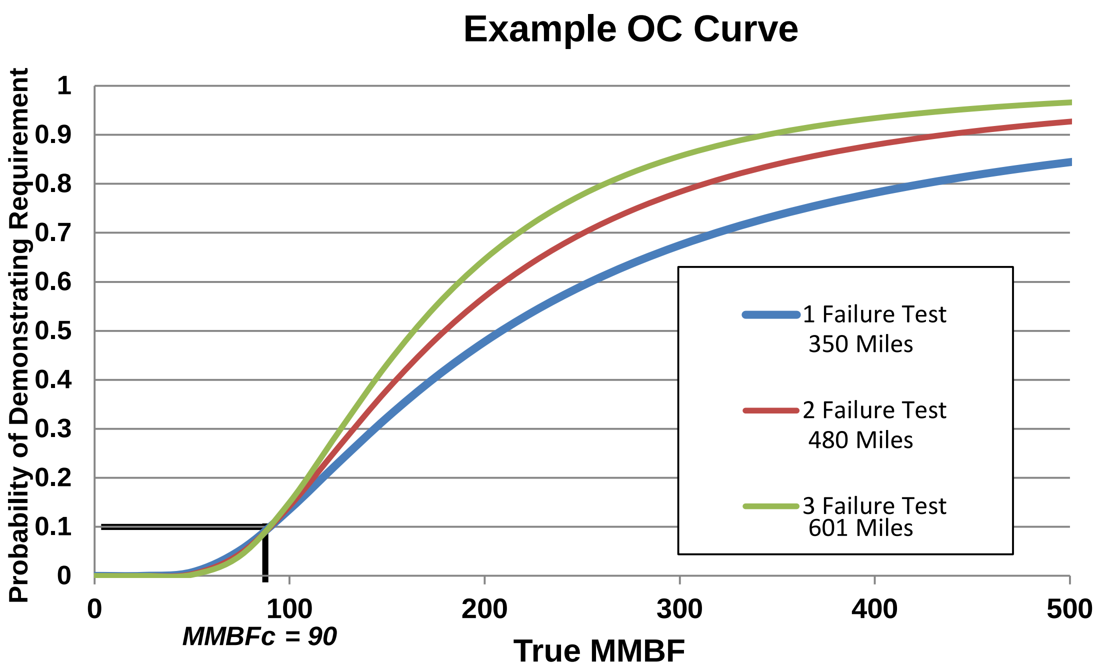
Traditional Test Plan

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!
- A traditional test plan approach in the DoD fixes consumer risk at the requirement (e.g., $\boldsymbol{\beta}_c = 0.10$ for a MMBFc = 90) and plans the minimum test around a fixed number of failures

Consumer Risk: *probability that a system with a MMBF as low as MMBFc will pass the demonstration test.*

Producer Risk: *probability that a system with a MMBF as high as MMBFp will not pass the demonstration test*



Assurance Testing: A Few Details

- **Two Risk Criteria in Determining a Test Plan**
  - **Consumer's Risk**

$$P(\text{Test is Passed} \mid MMBF = MMBF_c) \le \beta_c$$ *Classical Risk Criteria*

$$P(\text{MMBF} \le MMBF_c \mid \text{Test is Passed}) \le \beta_c$$ *Bayesian Risk Criteria*

  - **Producer's Risk**

$$P(\text{Test is Failed} \mid MMBF = MMBF_{\text{p}}) \le \alpha_p$$ *Classical Risk Criteria*

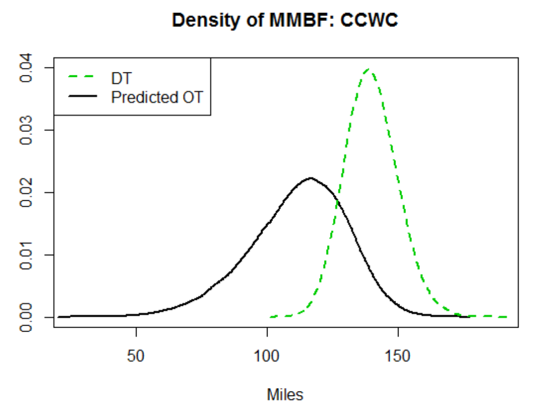
$$P(\text{MMBF} \ge MMBF_{\text{p}} \mid \text{Test is Failed}) \le \alpha_p$$ *Bayesian Risk Criteria*

Assurance Testing: Proof Of Concept

Building a test plan for one vehicle: vehicle 1

$$\eta_{vehicle\ 1} = \frac{\lambda^{*}_{Phase\ 3}}{1-\theta}$$

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.



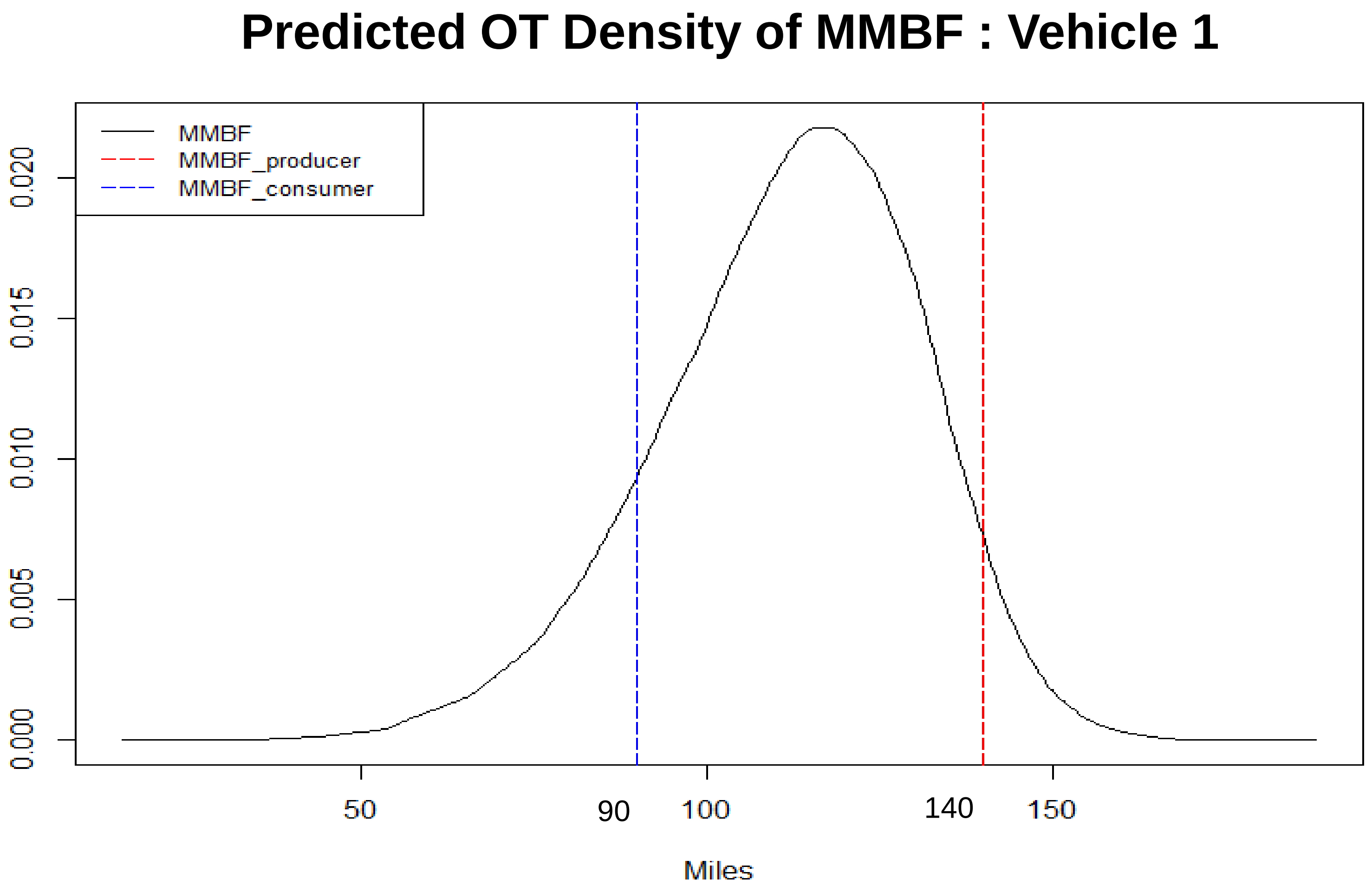
Posterior Consumer Risk

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

Posterior Producer Risk

$$P(\mathrm{MMBF} \ge MMBF_{\mathrm{p}} | \text{Test is Failed}) \approx \frac{\sum_{j=1}^{N}\left[1-\sum_{y=0}^{c}\left(t_0\eta^{(j)}\right)^{y} e^{-t_0\eta^{(j)}}\right] I\left(\eta^{(j)} \le \frac{1}{MMBF_p}\right)}{\sum_{j=1}^{N}\left[1-\sum_{y=0}^{c}(t_0\eta^{(j)})^{y} e^{-t_0\eta^{(j)}}\right]} \le \alpha_p$$

Assurance Testing: Proof of Concept



Results: Comparison to Traditional Test Plan

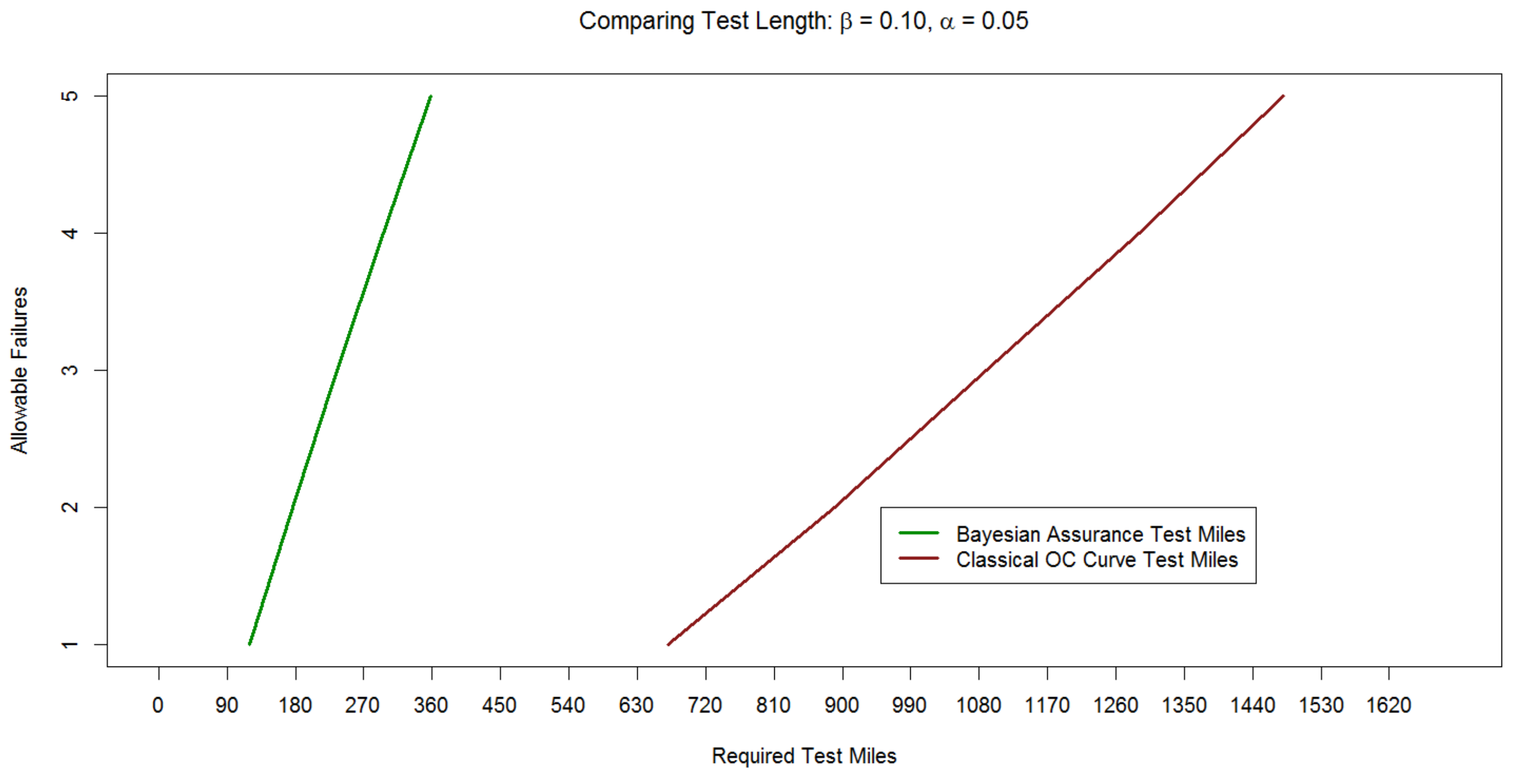
- **A traditional test plan approach in the DoD fixes consumer risk at the requirement (e.g., $\beta_c = 0.10$ for a MMBF = 90)**
  - Plans the minimum test around a fixed number of failures:
  - $Test\ Duration = Req * \left(\chi^2{}_{1-\alpha,2*N_f+2}\right)^{-1}/2$
  - Ignores producer risk

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	120	350 ($\alpha_p = 0.71$)
2	176	480 ($\alpha_p = 0.67$)
3	235	601 ($\alpha_p = 0.62$)
4	296	719 ($\alpha_p = 0.58$)
5	358	834 ($\alpha_p = 0.55$)

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

Results: Matching Risk Criteria

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



Future Directions

- **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 could be discovered in a future test phase
- **Exponential distribution assumption**
  - Assess the fit of our distributional assumptions
- **Incorporate Covariates**
  - Vehicle Variant
  - Test Site
    - » Difficulty of Terrain
    - » Weather Conditions

BACKUP

References

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.

A Traditional Analysis

- **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):**

$$\text{M}\widehat{\text{MBO}}\text{MF} = \frac{\text{Total Miles Driven}}{\#\text{ of }\ \textbf{OMF}\text{ Failures}}$$

An Overly Simplistic Analysis!

Summary of Results

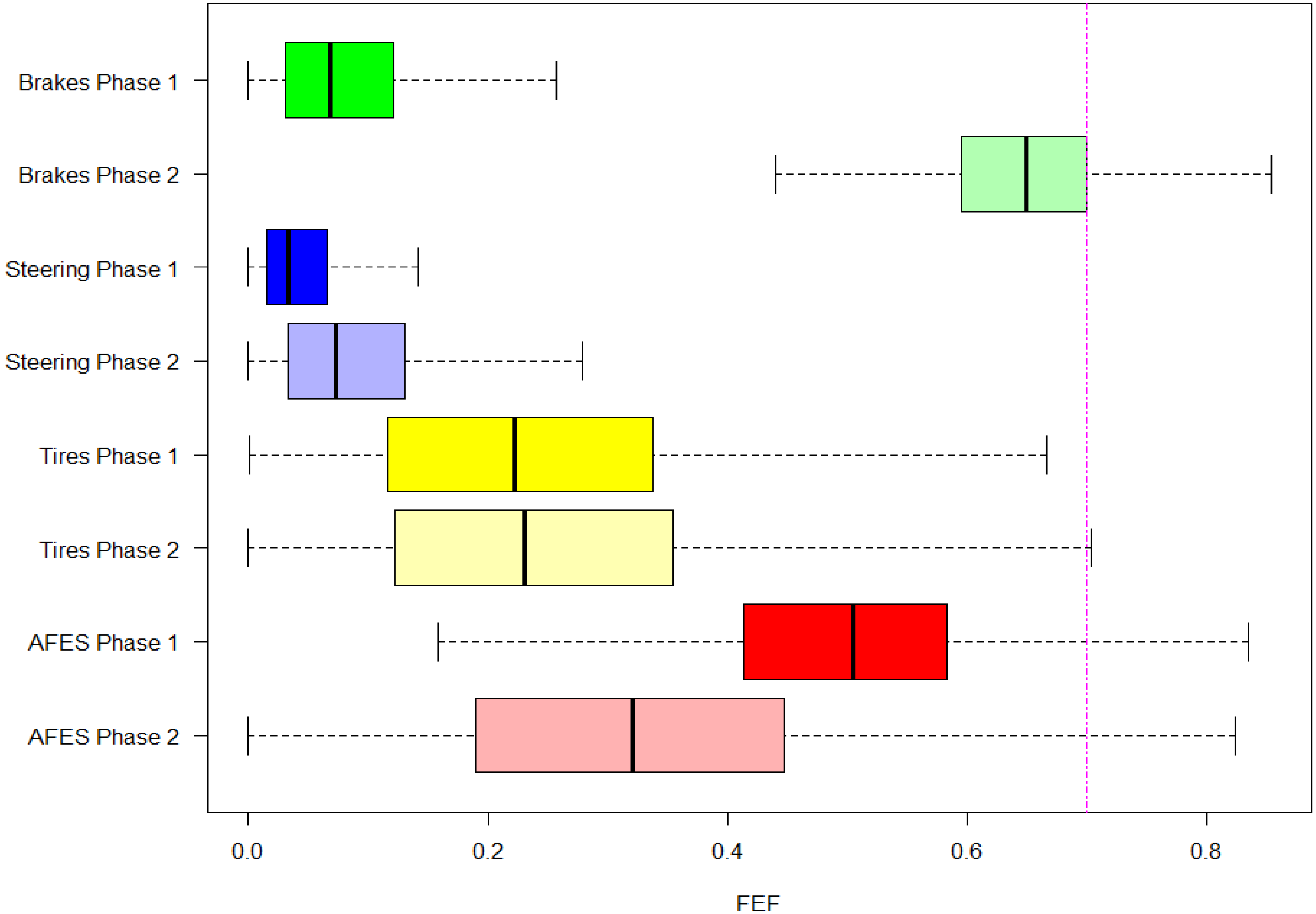
- **Estimate of a vehicle's failure rate**
  - Phase 1: $\sum_j \boldsymbol{\lambda_{ij}}$
  - Phase 2: $\sum_j \boldsymbol{\lambda_{ij}\rho_{1j}}$
  - Phase 3: $\sum_j \boldsymbol{\lambda_{ij}\rho_{1j}\rho_{2j}}$

- **Estimates of a failure mode rates**
  - Common failure modes: λ_j
  - Not Common failure modes: λ_{ij}

- **Estimate of the fixed effectiveness between phases 1 and 2.**
  - If the $\boldsymbol{\rho_j}$ estimates are between (0,1), the failure rates have decreased.
  - If the $\boldsymbol{\rho_j}$ estimates are greater than one, the failure mode has an increased failure rate.

Results: Fix Effectiveness Factor

Commonly Assumed Average FEF is 0.7 !!!



Unknown Failure Modes: Proof of Concept

- The Bayesian hierarchical model assumes that the number of observed failure modes is fixed at 26 and that there are no other, unseen failure modes.
- Assuming the number of failure modes is a random variable is not trivial in practice!!
  - » It is simple to complete a sensitivity analysis on the effect of adding unobserved failure modes.

All mileage is censored: 0 failures

$$t_{DT_1} \sim exp(\lambda_{ij}), \quad t_{DT_2} \sim exp\left(\lambda_{ij}\rho_{1_j}\right), \quad t_{DT_3} \sim exp\left(\lambda_{ij}\rho_{1_j}\rho_{2_j}\right)$$

Fix K at some number > 26

$i = 1,2,\ldots,8$ *(vehicles)* $j = 1,2,\ldots,26,\ldots,K$ *(failure modes)*

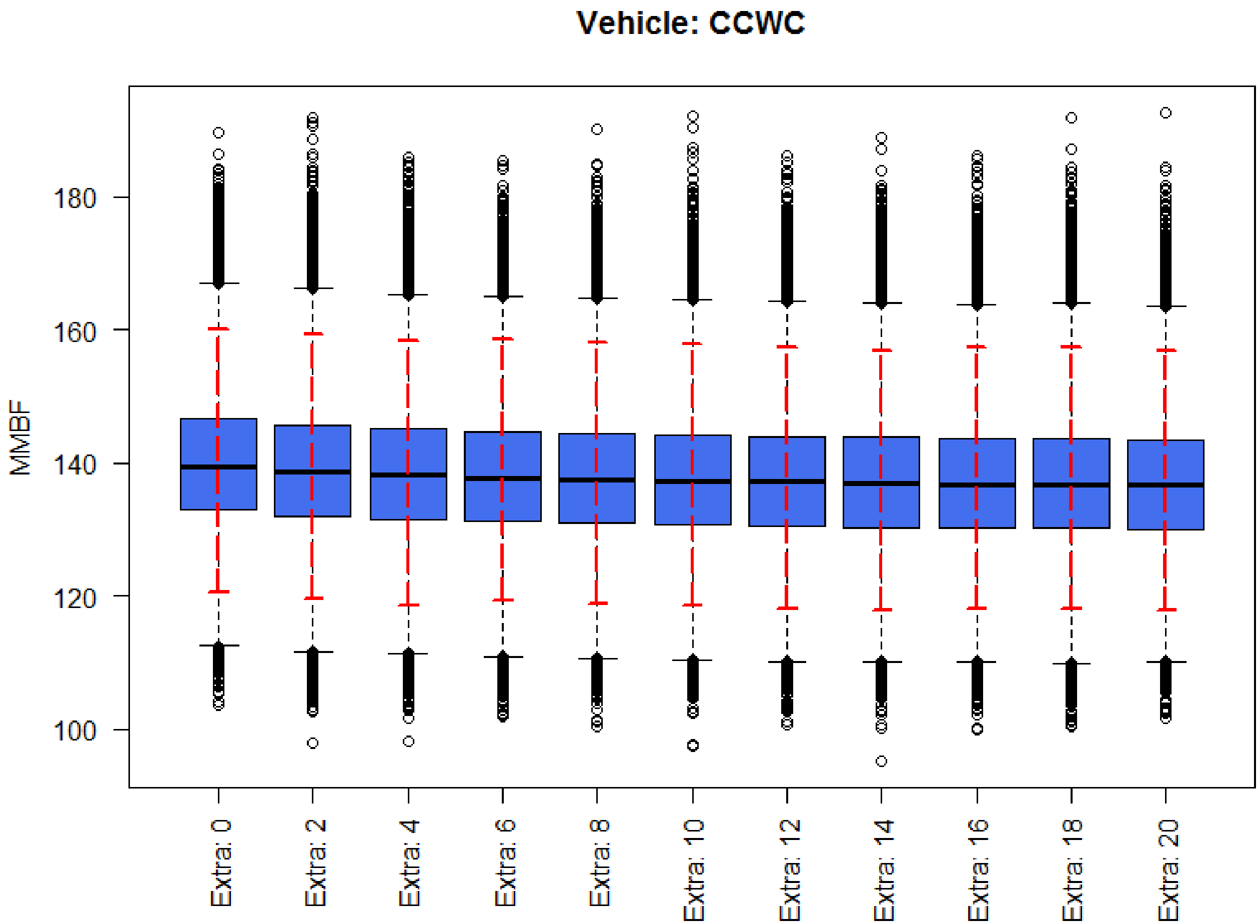
$$\lambda_{ij} \sim gamma(a,b)$$

Failure rates are not common across vehicle

$$\rho_1 = 1; \quad \rho_2 = 1$$

Given that the failure modes are not observed, the failure rates are constant across all phases

Unknown Failure Modes: Proof of Concept



Estimates for the MMBF based on our model are robust to the number of unobserved failure modes.