Statistical Methods for Combining Information: Stryker Family of Vehicles Reliability Case Study

Rebecca Dickinson, Laura Freeman, Alyson Wilson, Bruce Simpson

April 13, 2016
Outline

• **Motivation for Using All Information**

• **Background Information**
  – The Stryker Family of Vehicles (FoV)
  – Operational Testing (OT) and Developmental Testing (DT)

• **Methods**
  – The current DoD reliability analysis using OT data only
  – Frequentist and Bayesian models for combining DT and OT data

• **Results / Conclusions**
  – Tighter confidence intervals and better reliability estimates
  – Benefits are greatest for vehicles with only 0-2 reported failures in OT

• **Continuing Work**
  – Joint Light Tactical Vehicle (JLTV)
Motivation For Using All Information

• **What is the Current Practice?**
  – DOT&E in most cases uses only operational test data for reliability analyses
    » Stryker Beyond Low Rate Initial Production (BLRIP) Report
    » Benefit: ensures data is representative of operational test conditions
    » Drawback: discards expensive information from previous testing that provides information on system reliability

• **National Research Council Studies**
  – *Statistics, Testing and Defense Acquisition, 1998*
    » Emphasizes that all relevant information be examined for possible use in both the design and evaluation of operational tests …
    » State-of-the-art statistical methods for combining information should be used, when appropriate, to make tests and their associated evaluations as cost-efficient as possible

  – *Improved Operational Testing and Evaluation, 2004*
    » Focuses specifically on methods of combining information for the Stryker family of vehicles
The Stryker Family of Vehicles

Infantry Carrier Vehicle (ICV)

• Antitank Guided Missile Vehicle (ATGMV)
• Reconnaissance Vehicle (RV)
• Fire Support Vehicle (FSV)
• Commander’s Vehicle (CV)
• Medical Evacuation Vehicle (MEV)
• NBC Reconnaissance Vehicle (NBCRV)*

Engineer Squad Vehicle (ESV)

Mortar Carrier Vehicle (MCV)

* The NBC RV was excluded from the study because of its different acquisition timeline.
There are four essential functions
- Move
- Shoot
- Command and Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance (C4ISR)
- Survive

A failure is an event in which the vehicle or a subsystem of the vehicle does not perform as it should

The Army failure definition scoring criteria (FDSC) categorizes the severity of failures
- System Abort Failure
  » The vehicle is unable to complete the mission
- Essential Function Failure
- Non-essential Function Failure

Army Reliability requirement:
- Mean Miles Between System Aborts (MMBSA) = 1,000 miles
Developmental and Operational Testing

**Developmental Testing**
- **Primary Purpose:** Verification
- **Controlled Conditions**
- **Experienced Technicians operating the vehicles.**
  - They have done this for years and they know the courses really well
- **Courses**
  - Use courses that are designed to replicate the primary roads, secondary roads, and trail like conditions

**Operational Testing**
- **Primary Purpose:** Validation
- **Operational Conditions**
- **An army unit comes in to do this testing**
- **Courses**
  - OT data set comes from testing that was done at Fort Knox
  - Most of the testing was done using secondary road type conditions
- **Limited amount of Time**
  - Due to operator availability and range availability
  - Operational testing may be too short to discover many reliability deficiencies

DT and OT are Different!
- Environments
- Operators
- Test Durations
The Stryker FoV Data Set

Developmental Testing

- DT Estimate was 2197 MMBSA

Operational Testing

- OT Estimate was 8494 MMBSA, because of limited miles on each vehicle and only one failure

Very limited information available for the MEV in both DT and OT
A Traditional Analysis - Using OT Data Only

- The table below is similar to that which was included in the report written for DOT&E when considering this data set.

- These results serve as the reference when comparing the new methods that look at combining information across the developmental and operational test phases.

### Stryker Reliability by Variant using Operational Test Data

<table>
<thead>
<tr>
<th>Vehicle Variant</th>
<th>Total Miles Driven</th>
<th>System Aborts</th>
<th>MMBSA</th>
<th>MMBSA 95% LCL</th>
<th>MMBSA 95% UCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATGMV</td>
<td>10334</td>
<td>12</td>
<td>861</td>
<td>492.9971</td>
<td>1666.62</td>
</tr>
<tr>
<td>CV</td>
<td>8494</td>
<td>1</td>
<td>8494</td>
<td>1524.505</td>
<td>335495.1</td>
</tr>
<tr>
<td>ESV</td>
<td>3771</td>
<td>13</td>
<td>290</td>
<td>169.6326</td>
<td>544.7885</td>
</tr>
<tr>
<td>FSV</td>
<td>2306</td>
<td>1</td>
<td>2306</td>
<td>413.8815</td>
<td>91082.13</td>
</tr>
<tr>
<td>ICV</td>
<td>29982</td>
<td>35</td>
<td>857</td>
<td>615.9437</td>
<td>1229.84</td>
</tr>
<tr>
<td>MCV</td>
<td>4521</td>
<td>4</td>
<td>1130</td>
<td>441.4354</td>
<td>4148.219</td>
</tr>
<tr>
<td>MEV</td>
<td>1967</td>
<td>0</td>
<td>-</td>
<td>656.6007</td>
<td>-</td>
</tr>
<tr>
<td>RV</td>
<td>5374</td>
<td>2</td>
<td>2687</td>
<td>743.8384</td>
<td>22187.42</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>66749</strong></td>
<td><strong>68</strong></td>
<td><strong>982</strong></td>
<td>774.2946</td>
<td>1264.074</td>
</tr>
</tbody>
</table>

Mean Miles Before a System Abort (MMBSA) = \( \frac{\text{Total Miles Driven}}{\# \text{System Aborts}} \)
Failure-Time Regression Models

We can use the exponential distribution to model the miles before a system abort

\[ t_{ijk} \sim \text{exponential}(\lambda_{ij}) \]

\( i = 1,2 \) (test phase)
\( j = 1,2,...,7 \) (vehicle variant)
\( k = 1,2,...,n_{ij} \) (miles)

We can express rate parameter, \( \lambda \), as a function of explanatory variables to find estimates for the MMBSA

Model 1: Average over vehicle type (assumes vehicle type does not matter)

\[ \lambda_i = \gamma_0 + \gamma_1 \text{Test Phase} \]

Model 2: Average over test phase (assumes vehicle type does not matter)

\[ \lambda_j = \gamma_0 + \gamma_1 \text{ATGMV} + ... + \gamma_6 \text{MCV} \]

Yes, we combine information – but we completely ignore the test phase!

Model 3: Look at differences based on Test Phase & Vehicle Type

\[ \lambda_{ij} = \gamma_0 + \gamma_1 \text{Test Phase} + \gamma_2 \text{ATGMV} + ... + \gamma_7 \text{MCV} \]
Exponential Regression Results

MMBSA Estimates from Model 3 for DT and OT
(95% Confidence Intervals)

This model estimates a 37% degradation in the MMBSA moving from DT To OT
Comparing Confidence Intervals

Operational Test MMBSA Estimates
(95% Confidence Intervals)

- Traditional Analysis
- Exponential Regression

Tighter confidence intervals & better estimates for MMBSA
A Bayesian Analysis: Incorporating More Information

- Bayesian model is specified by:
  - Parametric statistical model (just as before) and a prior distribution

- Informative Priors
  - Based on subject matter expertise

- A Hierarchical Model
  - Assumes the parameters are related, the data tells us how closely related
  - Allow us to estimate the MEV reliability based on other data

\[
t_{DT} \sim \exp(\lambda_i) \quad t_{OT} \sim \exp(\lambda_i/\eta)
\]

\[
i = 1, 2, \ldots, 8 \text{ (vehicle variants including the MEV)}
\]

\[
\lambda_i \sim \text{gamma}(a, b) \\
\eta \sim \text{beta}(1, 1) \\
a \sim \text{gamma}(0.001, 0.001) \\
b \sim \text{gamma}(0.001, 0.001)
\]
Comparing the Results of the Three Analyses

Operational Test MMBSA Estimates
(95% Confidence Intervals)

- Traditional Analysis
- Exponential Regression
- Bayesian Analysis

MMBSA

ATGMV  CV  ESV  FSV  ICV  MCV  RV  MEV
Stryker Case Study Conclusions

• We can use basic statistical models to incorporate information from multiple testing phases into OT assessments

• The results are:
  – Tighter confidence intervals (an average of a **60%** reduction in the interval width)
  – Better estimates for MMBSA
    » Commander’s Vehicle estimates were optimistically high before incorporating information from DT
  – Benefits are greatest for vehicles with only 0-2 reported failures in OT

• Bayesian techniques provide:
  – Ability to incorporate more information than is contained in the data
    » Subject matter expertise
    » Historical information not directly contained in data

• Analysis requires more statistical knowledge than the Traditional OT analyses
  – Information gained is worth the effort
Continuing Work: The JLTV Family of Vehicles

• Family of Vehicles designed to replace the Legacy Humvee Fleet.

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.

*Analysis based on notional data*
Continuing Work: Leveraging Additional Failure Information

• 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 (brakes, steering,…).
    » There are 26 observed failure modes across the three phases of testing.

• Incorporating all failures in the analysis (i.e. System Aborts and Essential Function Failures)
  – Across all eight vehicles tested and the three test phases, there are 91 SAs and 1,321 EFFs.
  – EFFs include a large portion of the failure modes that drive maintenance costs and reduce system availability

\[
t_{DT1} \sim \exp(\lambda_{ij}), \quad t_{DT2} \sim \exp(\lambda_{ij}\rho_{1j}), \quad t_{DT3} \sim \exp(\lambda_{ij}\rho_{1j}\rho_{2j})
\]

\[
i = 1,2,\ldots,8 \text{ (vehicles)} \quad j = 1,2,\ldots,26 \text{ (failure modes)}
\]

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

\[
\rho_{1} \sim \text{gamma}(c, d) \quad \rho_{2} \sim \text{gamma}(c, d)
\]

\[
a \sim \text{gamma}(0.001, 0.001) \quad b \sim \text{gamma}(0.001, 0.001)
\]

\[
c \sim \text{gamma}(0.001, 0.001) \quad d \sim \text{gamma}(0.001, 0.001)
\]

*Analysis based on notional data
Continuing Work: A Few Results

Operational Test MMBF Estimate
(95% Credible Intervals)

Mean Miles Between Failure

CCWC  GP2  HGC  UV2

*Analysis based on notional data
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?

Demonstration Test (Operational Characteristic Curve)

– 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.
References


