2019 NDIA GROUND VEHICLE SYSTEMS ENGINEERING AND TECHNOLOGY SYMPOSIUM
SYSTEMS ENGINEERING TECHNICAL SESSION
AUGUST 13-15, 2019 - NOVI, MICHIGAN

TESTING IN A COMPLEX WORLD

Bryan LaRose[1], Melissa Morgan[1], Kristen Stark[3], Dan Kosinski[1], Ken Fischer[1], Mark Brudnak, Ph.D.[1], Greg Schultz, Ph.D.[2], Angelo Christino[3], Martin Wayne[4], Igor Baseski[1]

[1]U.S. Army Combat Capabilities Development Command, Ground Vehicle Systems Center, Warren, MI
[2]U.S. Army Aberdeen Test Center, Aberdeen, MD
[3]U.S. Army Test and Evaluation Center, Aberdeen, MD
[4]U.S. Army Combat Capabilities Development Command, Data & Analysis Center, Aberdeen, MD

ABSTRACT

This paper discusses and outlines ideas regarding changes to how testing is performed in response to new policy regarding rapid integration of technology into Army ground vehicle systems. It specifically presents and discusses the ways that systems can begin testing early using laboratory testing. It discusses how testing is currently performed and then leverages best practices from the Automotive Industry to recommend methods to recommend how the Army can adapt these for its testing function. Specifically it discusses how specific test should be selected, how to define the testing environment and how to use the data generated from the lab test. It concludes with an example case study.

Citation: B. LaRose, M. Morgan, K. Stark, D. Kosinski, K. Fischer, M. Brudnak, G. Schultz, A. Christino, M. Wayne, I. Baseski, "Testing in a Complex World", In *Proceedings of the Ground Vehicle Systems Engineering and Technology Symposium* (GVSETS), NDIA, Novi, MI, Aug. 13-15, 2019.

1. INTRODUCTION

As the U.S. Army transforms under the new Combat Capabilities Development Command Campaign Plan [1], the Army Operating Concept (AOC) [2], the Army Strategy [3] and the Army Modernization Strategy [4] in conjunction with the creation of the Army Futures Command (AFC), demands on the U.S. Army's Ground Vehicle Test and Evaluation (T&E) community will drive the need to maximize program efficacy. Given that ground systems will increasingly rely on rapid integration, customization, and incorporation of Commercial off the Shelf (COTS) parts and components, the Army Test and Evaluation Command (ATEC), the Next Generation Combat Vehicle (NGCV) Cross Functional Team (CFT), and the Combat Capability Development

Command (CCDC) Ground Vehicle Systems Center (GVSC) must develop a T&E strategy to rapidly mature designs as well as validate requirements and/or proposed solutions prior to production and fielding.

Trends in the anticipated systems and technology upgrades of the future indicate that they will be increasingly complex; implying that not only will the number of components increase, but also the integration complexity as defined by the number of component dependencies and interactions. Additionally, as highlighted in the AOC, the future engagements are unknown and unknowable. In order to operate in this environment, the Army must develop and field reliable systems and technology in real-time as the unknowns become known. The ability to do this rapidly and effectively will be the key to winning the Army's future conflicts. To achieve this responsiveness, the Army Ground Vehicle community must adapt and change its processes to shorten the time required between operational need and provision of an effective solution.

As the AFC continues to come together and evolve, it is vital that we align our efforts with those laid out by the Chief of Staff of the Army (CSA), the AFC Commanding General (CG), and the CCDC's priorities. As laid out in the Campaign Strategy, one of the resounding topics is Readiness. And although we are certain the direction of the AFC is to leverage COTS technology solutions, those systems will not be validated to the intended use. Whether it be through proper contracting and requirements for Original Equipment Manufacturer (OEM) led T&E efforts or independent validation through the use of the Army T&E community, it is critical that we all accept and implement a new way of validating.

In this paper we will specifically discuss ways that laboratory validation and testing can and should be integrated into the overall T&E process. We will cover laboratory testing methods and techniques that can reduce test time and cost allowing faster design iteration and reliability growth. Furthermore, this paper outlines methods by which laboratory-based component and subsystem testing can and should be leveraged as part of the overall Test and Evaluation Master Plan (TEMP).

It discusses and presents how approaches which have been widely adopted in industry and other services could be adopted by the Army Ground System community. The Army has specifically addressed its need to reform T&E methodology and on 11 JULY 2013 the Assistant Secretary of the Army for Acquisition, Logistics, and Technology (ASA(ALT)) published their findings [5] from the T&E Efficiency Task Force.

We outline a proposed method to begin testing at the component level as soon as possible. The paper will discuss an approach to leverage resources including Military Standards and Vehicle Responses to begin testing and growing reliability at the component level while a system is still being engineered and developed. This will improve the reliability of a system when it enters system level test, increasing assurance that components and subsystems have individually achieved high reliability, narrowing the number of failure modes which still need to be found and fixed. This approach relies on the following technical capabilities: 1) defining the testing environment and boundary conditions, 2) acceleration/ compression of the test to increase volume and efficiency, 3) attribution/roll up of reliability to the system level. This paper describes strategies to accomplish this from the perspective of vibration related reliability.

2. SYSTEM DEVELOPMENT PROCESS

Requirements development is one of the earliest steps in the current Army systems development process. Initial requirements are based on operational capability needs and include system performance and Reliability, Availability, Maintainability (RAM) requirements, as well as their numeric values. The RAM requirements are based on the forecasted usage defined in the Operations Mode Summary/ Mission Profile

(OMS/MP). Anticipated failures are defined in the accompanying Failure Definition Scoring Criteria (FDSC). System performance and RAM operational requirements are then translated into performance specifications that are placed into the contract and tested for compliance. In general, the majority of testing occurs after full-system test articles are delivered to the Army's proving grounds, starting with Developmental Testing (DT) and then progressing to Operational Testing (OT).

DT-level performance specification testing is generally straightforward and relatively expedient (e.g., measuring 0-60 mph vehicle acceleration times and determining pass/fail). However, system reliability testing is very time consuming, particularly if immature systems are delivered to the proving grounds.

Historically, reliability testing is based on demonstrating the aforementioned requirements operating the system in accordance with the OMS/MP. The process to determine requirements compliance uses reliability scoring conferences where the severity of failures observed in test are cataloged and discussed using the FDSC. The process of testing systems using average use case is effective at determining the system's compliance to the reliability requirement, typically defined as a mean time between failure. The failures that are observed during the testing and generally fixed, depending on the severity and cost of correcting the issue, and results in a test-fix-test paradigm.

3. THE NEED TO CHANGE

To understand the need for change, we must first discuss recent Army reforms and initiatives geared towards speeding up the Acquisition process. These acquisition reforms and initiatives include the establishment of AFC; streamlining and improving ongoing acquisition activities such as contracting, sustainment and testing; creating CFTs focused on rapidly defining requirements; refocusing science and technology (S&T) priorities and investment; and changing oversight and decision making related to major acquisition programs [6]. The goal of the CFTs is to bring together different experts in contracting, requirements, logistics, T&E, and S&T to facilitate collaboration to provide immediate input as opposed to the more traditional requirements development process, in which input has typically been provided separately [7].

Traditionally, the test and evaluation period is a large portion of the acquisition schedule, therefore, streamlining is a critical component to expediently field new equipment and technology to the warfighter safely and effectively. The Middle Tier Acquisition (MTA) for Rapid Prototyping and Rapid Fielding (Section 804 of 2016 National Defense Authorization Act) provides a pathway to rapidly prototype innovative technologies or field proven technologies within a relatively short time. In conjunction with the MTA, the Program Management Offices (PMOs), ATEC and CFTs must work as partners to build a test program that is focused and provides essential data early on to characterize risk and inform decision makers. And this process must be continued post fielding to ensure feedback from operational environments are considered in engineering changes processes (Figure 1).

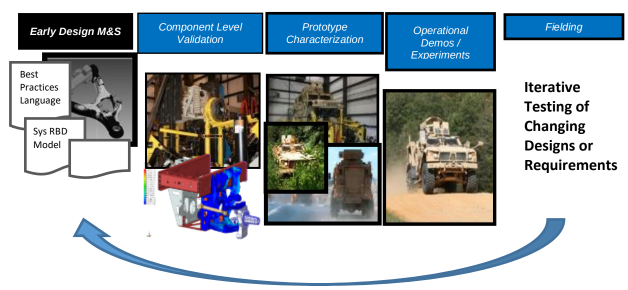

Figure 1: Iterative Validation Post Fielding

There are inherent risks and challenges to a shorter, schedule-driven test program that are already being realized with current ground system programs. Recent history has provided many examples of poor design for reliability or insufficient validation of technology during the design stage, which has led to contractors delivering systems to ATEC for Developmental Testing (DT) that are not mature enough for demonstration of requirements. This has led to programmatic delays due to the test-fix-test cycle in what should be a confirmation period and not a discovery period. In addition to technology readiness shortcomings, manufacturing quality issues or long lead items have also led to test assets being delivered to ATEC behind schedule. This results in lack of test data available in critical decision making events. These risks and challenges should be considered during the planning process to ensure the right resources are dedicated to early system design and reliability.

The dramatic overhaul in the Army acquisition process demands an equal transformation in future testing. The next section of this paper discusses the approaches that have been considered "industry standards" for several decades. These T&E strategies include increased use of Modeling and Simulation (M&S); delivery of contractor test data to be used in formal evaluations (Figure 2); and reliance on contractor expertise and M&S for MTA system reliability. Key to ensuring that these strategies are adapted in the Army Acquisition process is inclusion of contract language that mandates delivery of (1) test plans and associated results, (2) M&S inputs and outputs, and (3) Contract Deliverables (CDRLs) that are submitted on a regular basis and contain the necessary data to perform a comprehensive evaluation.

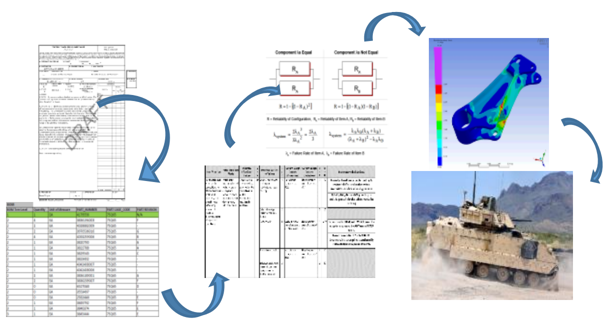

Figure 2: Assessing Contractor Data

The Army has specifically addressed its need to reform T&E methodology and on 11 JULY 2013 the ASA(ALT) published their findings [5] from the T&E Efficiency Task Force. The Task Force identified one of its key aspect as utilization of M&S, stating: "Given the scarcity of available resources, the Army T&E community should determine how to leverage existing M&S tools (Hardware-In-The-Loop (HWIL) as well as purely computer-based) to ensure that high value acquisition programs with substantial T&E requirements are performed in the most efficient manner possible."

4. FUTURE OF TESTING IN THE ARMY

In order to be effective and responsive, the Army must change the way it does system acquisition, specifically with respect to the way it conducts testing. It must embrace and utilize subsystem and component testing early to accelerate system reliability growth. Fortunately, the automotive industry, driven by pressures to decrease time-to-market, reduce cost and improve quality has forged a path of best practices regarding system design and testing which can be leveraged by the Army. It is important to note the differences between automotive and defense, but from a technology and method point of view, there exists significant and beneficial overlap. The location of the Army ground vehicle community (Program Executive Offices (PEOs) and GVSC) in the automotive capital of the world was a strategic decision, and it is vital that we leverage industry best practices to ensure we maximize the effectiveness of that decision. Using the automotive industry as a guide, the Army must begin to adopt and adapt the processes, tools and methods of laboratory testing to fulfil its mission of delivering timely and effective equipment to our soldiers.

According to Madden, the automotive industry first used full vehicle simulators in 1959 [8]. Since that time, Leese et al [9] reported that simulator technology became widespread in the automotive industry by the 1980's. Today they are commonplace with an entire industry focused on developing and providing laboratory simulator technologies. Just one single test equipment provider (MTS Systems) reported $778M in gross revenue in 2018 [10]. Steve Haeg of MTS Systems Corporation states that "Complete elimination of expensive prototypes and proving ground testing is the goal of most vehicle development managers. Almost every company has reduced its number of full-vehicle prototypes over the last decade, and are opting to rely more heavily on Computer-Aided Engineering (CAE) and lab testing to verify product performance." [11]. Hence the evolution of laboratory testing in the automotive sector has been driven by the need to innovate and deliver reliable vehicles to their customers. Given that a typical vehicle may be in service for 100,000+ miles over a lifetime of 10-15 years.

The automotive industry's adoption of laboratory simulation has been a continual and adaptive process. With respect to hardware, it began with four-post test rigs [8] and evolved to multi-axial rigs, first with 3 Degrees of Freedom (DOF) and eventually 6 DOF per wheel. Corresponding to the development of the hardware, the control evolved from the playback of road profiles [12], to the development of Remote Parameter Control (RPC) [13], to the development of non-square control and use of Singular Value Decomposition (SVD) [14] to operational RPC. Likewise, methods of test editing were developed to compress or shorten the amount of time required while maintaining a substantial amount of test severity. Methods in this regard have been reported to shorten test time to anywhere from 10% to 25% of the original length [9], [15], [16], [17]. Similarly, the incorporation of and integration of analytical modeling and simulation with testing has likewise grown and increased [18], [19]. Finally, automotive has, over the years, refined their processes regarding how M&S, lab testing and field testing should be brought together to improve reliability of their vehicle designs.

As discussed in previous sections, the Army ground system community is significantly lacking in capability and utilization compared to our peers in industry. The automotive sector has been utilizing modeling and simulation to validate their designs almost exclusively for decades.

Understanding that the Army and automotive industry are different in significant ways, we believe that these differences are not prohibitive. The primary differences are enumerated in the following table:

Aspect	**Automotive**	**Army**
Volume	Millions of units/yr	1,000's of units/yr
Duty cycle	12,000 miles/yr	Varies
Model update	Yearly	5-10 years
Buy decision	After production	At milestone
Buy unit	Single vehicle	Fleet
Environment	On-road	On/off Road
Incentives	Many OEMs + Market	Single-few OEM + Requirements
Testing	Test to confirm	Test to discover

In order to adopt these tools, techniques and procedures, the Army must commit to a transformation process which will include changes to both acquisition and testing policies. It is certain that such a transformation will be a learning process, but to attain these efficiencies, the Army must begin the journey, understanding that the right solution may take a couple of iterations to achieve the full benefit. Given that MTA programs are specifically directed to seek innovative methods to accelerate design and testing, these programs provide a more risk tolerant approach to validation, out of which new and innovative processes, policies and techniques can grow to maturity and be applied to traditional acquisition programs.

Current capabilities unique to the US Army Ground System community utilizing M&S for designs ready for integration (Technology Readiness Level (TRL 6+)) exist at both the Army

ATEC and the CCDC GVSC. Other DoD agencies and industry partners can be utilized; however, sensitive data, limited payloads, and equipment capacity often limits the availability of those options.

To date, much of the capability available has been utilized to make specific engineering decisions for components integrated on mature, fielded systems. This application has proven to be effective for generating rapid data to make decisions on engineering changes where vehicle response data has been generated and on track loads monitored for the history of the vehicle. This is a great start, but in order to truly maximize the useful data generated from M&S while minimizing cost and schedule impact to the Army, it must adopt a more consistent plan to introduce these capabilities early in development prior to Full Up System Level (FUSL) testing.

This paper and the communities supporting it will provide an approach that leverages the lessons learned from private industry, best practices from engineering communities, and the capabilities currently available to the Army Ground System community to propose a crawl, walk, run (see Nolan et al [20]) approach to validating the development of technologies prior to release for Soldier integration in OT.

Crawl being the utilization of digital (Finite Element Analysis (FEA), Physics of Failure) and/or theoretical (Failure Modes and Effect Analysis (FMEA), Fault Tree) modeling to identify high risk components and subsystems for focus. This should also be accompanied by proper contract language to dictate the quality of these models and explanation of how these identified subsystems will be monitored during the contracting phase along with developmental phases prior to Low Rate Initial Production (LRIP).

Walk being the utilization of digital and physical simulation in order to validate the failure rate of those components and subsystems when integrated into their intended systems and subjected to loads that are expected in a relevant environment. This can be done by either leveraging time history data from prototype systems, Military Standards for exact or similar systems, or vehicle responses generated through physics based models. Any deficiencies identified in this stage would be subject to redesign in order to maximize reliability prior to integration.

Run being the integration of all components and subsystems into a FUSL test and introducing realistic loads and environments in order to validate suitability and safety prior to OT.

The utilization of this crawl, walk, run approach offers countless advantages over the current method of validating development at the FUSL DT only; most of this involves more comprehensive data along with a cheaper and faster cost to the government. In this section we will merely address the pros of the "walk" phase of utilizing simulators to test vehicle components and subsystems prior to integration.

By leveraging physical simulation to validate technologies prior to FUSL testing, it is possible to drastically reduce the schedule and financial cost to the government. Cost savings can be realized in many ways, but most of those are realized during developmental testing and sustainment.

Developmental testing cost is impacted by simulation by eliminating the need to stress mature systems that have already been validated simply to test a handful of new technologies. Where specific components and subsystems have been identified as high risk, simulation of those technologies can be focused and accelerated to generate data faster and with greater confidence then the FUSL level approach. FUSL testing stresses an entire vehicle, shuts down all testing for one failed component, and is limited to the available vehicle resources in order to generate sample sizes that create statistical confidence. The following discussion outlines the key elements regarding choices that must be made in the application of lab testing to an acquisition program.

4.1. Deciding what to test

Determining which components to test must be driven to directly address risk. Unknown aspects of the system need to be identified early through the development of a Failure Modes, Effects and Criticality Analysis (FMECA). Components of unknown durability must be identified early and have targeted testing programs designed. Tests should be highly accelerated and designed to run to failure to establish a picture of the reliability of the component. Additionally, the component must be tested in sufficient quantities to establish statistical variation in the reliability. Then as components come together as subsystems and eventually a full system, laboratory tests must likewise be designed to surface failure modes due to integration.

Formal tools have been and are being developed to add rigor to this process. These tools include the Maintenance Aware Design Environment (MADE) tool and the Army Lifecycle Test Optimization (ALTO) tool [21]. These tools provide very specific recommendations regarding testing given overall system-level objectives.

This approach is incredibly important for new development systems that are going through rapid prototyping. The schedule and cost constraints introduced in these cases make the extensive FUSL reliability growth testing completely infeasible. Despite the limited time availability for executing a reliability growth program, the soldiers' expectations of being delivered a reliable system remain. This can only be achieved for these systems through aggressive subsystem and component level testing.

4.2. Deciding on test conditions

Nolan et al [20] outline a process for determining test conditions. This process is roughly as follows:

Testing begins at the component level as early as possible using a combination of historical tests to define the test parameters. These tests have limited confidence because loads or materials may have changed since data were recorded, but still provide useful information in that they can surface failure modes very early and quantify initial reliability. Second, testing is conducted on a mule or hybrid vehicle (i.e. combination of production and prototype parts). It is used to approximate loads experienced in service for use in FEA modeling and/or to identify major design defects early. The third stage is pre-prototype which can be used to gather more accurate information for component tests and can be used for early systems level testing in the laboratory. Test conditions in this case can be acquired by operating the vehicle in a relevant environment. Note that the emphasis for developing a physical system is to feed data into the laboratory tests (i.e. having one physical system serves as a multiplier for lab testing). The fourth stage consists of vehicles made with prototype tooling which are used for durability validation as opposed to evaluation. (Note that in the Army process this is the equivalent of the end of Engineering & Manufacturing Development (EMD)). The three facets of testing are illustrated in Figure 3.

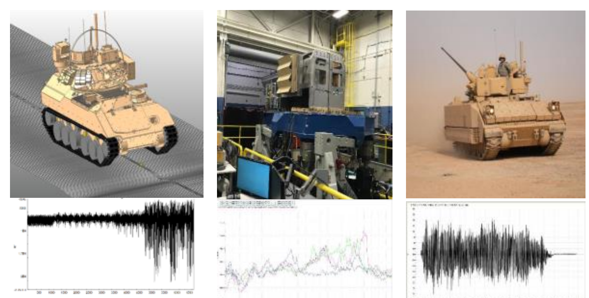

Figure 3: Multi-Faceted Testing.

This is when the Army has a set of vehicles delivered from the OEM to begin evaluating reliability against a requirement. It is at this time that the Army begins officially surfacing failure modes. In automotive, reliability has already gone through several levels of testing, and at this stage there is a reasonable level of confidence in the reliability of the system. In automotive testing, at this stage, is a confirmation rather than a discovery process. They conclude with a fifth stage of testing in both simulation and proving ground using early production vehicles. (This would be the Army equivalent to LRIP).

Early identification of test conditions is extremely important. Nolan suggests using historical data. To be more precise we recommend the following steps to identify environment data for an early component test.

1. Find an existing vehicle which is similar to the target vehicle.
2. Adjust the level of the data using one or more of the following methods.
   a. Transform to a new location using known kinematic or other physical relationships.
   b. Scale the data based on an understanding of "level" between the existing vehicle and the target vehicle. (i.e. Use M&S to determine that vehicle X's vibration is 87% of vehicle Y's vibration under similar circumstances). This scaling may be single value or a function of frequency.
3. Use early mule to validate a physics model which can then be used to generate the required test condition data.
4. Use a pre-production prototype as a data acquisition rig to acquire a set of data in a relevant environment to be used as test conditions.

4.3. Why Lab Testing?

The way that the Army currently tests vehicle systems is partially driven by the way that we acquire systems. The Army acquires a system as an integrated platform which is developed and manufactured by private industry OEM. The current approach emphasizes that the Army specify the "what" not the "how" when developing requirements. This applies to reliability as well. The Army will specify a requirement at the system level in terms of a metric such as mean time between system abort (MTBSA). It is then the OEM's responsibility to deliver a vehicle meeting that requirement. If the delivered vehicle does not meet this requirement, the program undergoes a process called reliability growth in which failure modes are discovered and fixed.

The OEM's, in an effort to demonstrate reliability prior to proposal, can perform their own system, subsystem and component, but often the data obtained in these tests cannot be properly qualified due to insufficient quantification of loads to the OMS/MP.

Of course for new systems, this introduces a problem of the testing environment being derived from systems-level responses that cannot be measured, and components cannot be realistically tested without a valid testing environment.

Fortunately we have tools at our disposal such as military standards, digital models and environmental data from other vehicles. Laboratory testing in the Department of Defense (DoD) has grown and matured over the past 50+ years. This is reflected in the publication of the first version of MIL-STD-810 back in 1962, which was originally applied to Air Force systems. The standard evolved to incorporate inputs for ground systems as well. The ground system community in the Army first began laboratory durability testing in the 1970's with the building of the Ground Vehicle Simulation Lab (GVSL) at the Detroit Arsenal in Warren, MI [22]. Since then the GVSL has built a robust laboratory testing capability by leveraging tools and techniques from the automotive sector, only higher capacity for the Army's heavier vehicles. These include N-post simulators, vibrations tables, environmental chambers, characterization fixtures for mass properties, suspension properties, and fixtures for single suspension unit testing. Additionally the propulsion test group at GVSC has capabilities such as chassis dynamometers, engine test cells, battery test cells and environmental chambers. Aberdeen Test Center has likewise grown their laboratory testing capability with the addition of the Roadway Simulator (RWS) and the Vehicle Durability Simulator (VDS). We here outline some of the advantages and disadvantages of lab testing.

The main advantages of laboratory testing over field testing are flexibility, repeatability, efficiency, and affordability. Xu et al enumerate six advantages [23]. Laboratories allow for testing under an unlimited combination of environmental conditions and loading scenarios. The ability to rapidly change test profiles provides significant flexibility and test efficiency that cannot be found in field experiments. Laboratory testing is repeatable and reproducible over time. This allows for accurate performance comparison of competing technologies, or in support of reliability growth. Laboratory testing allows for several test compression strategies from removal of non-damaging events, to Step-Stress testing, to statistical DOEs. Each of these advantages of laboratory testing result in lower test cost and less time required to test. All simulations are by definition approximations of reality. This is true whether they are analytical or physical, and whether they are conducted in a laboratory or in the field. In all cases, the level of fidelity in the simulation increases both the cost and complexity of the test system. While there are many pros to physical simulation, it is important to also understand the limitations and risks that come with this approach. Laboratory testing is capable of matching most loads a component or subsystem will see, but not all. Simulation also gives little insight into user operation that would be seen by even a test driver. The varying environment such as weather (rain), terrain (mud/sand), or temperature are not simultaneously available. Due to the lack of investment in simulation capability, the current infrastructure maintained by the Army would not allow for even one system to go through this approach currently, let alone multiple competing systems at once. Current infrastructure also limits the creativity the contractor can provide. Any design can be validated on course as that is it's intended environment. Unique designs require the design and acquisition of simulators not currently available, which is very costly and time consuming. Most of the disadvantages with laboratory testing center around cost. While the cost of test projects may be lower than a comparable field experiment, there are several funding issues for the Army that make laboratories a business challenge.

Investment challenges

- Upfront challenge with Program Objective Memorandum (POM) funding process.
- inability of the government to profit
- Color of money challenges

Sustainment challenges

- Challenge with quantifying cost on a per project basis
- Color of money challenges
- Multiple commands (Installation Management Command (IMCOM), CCDC) need to coordinate budget plans

Test system cost of complexity

- Reduces number of independent control variables
- Limits fidelity in surfacing failure modes that are multi-dimensional (e.g. fails under certain loads at temp)

4.4. Policy and procedural changes

As discussed earlier in the paper, in order to adopt a radical change in the way the Army validates Ground System Reliability it is vital that policies and procedures be modified to ensure the process is effective. To address this, the Army must focus on three critical areas: Requirements that support subsystem characterization, contract language and deliverables to ensure industry delivers acceptable data packages, and policy to support the acquisition and test communities as they plan, execute, and evaluate new test methodology.

Current Army ground system reliability requirements focus on the system as a whole [24]. In order to adopt and implement a new approach to reliability validation, a change in how reliability requirements are written is needed. This is difficult under current DoD Acquisition policies, as the technology solution to a need is not prescribed. This can be overcome by ensuring proper deliverables, their format, and their fidelity are spelled out in specifications to the industrial base.

Leveraging tools currently available, including FMEAs, Reliability Block Diagrams, and WBS levels it is possible for the Acquisition community to dissect a FUSL requirement and add engineering rigor to the technologies being provided. A starting example can be seen in the JPO JLTV RFP from 2012 that requested contractors deliver an FMEA of the components integrated into the vehicle and the risk of those components failure during operation. This model can be built upon through CDRLs that require specific WBS levels (as noted in DoD Reliability, Availability, Maintainability-Cost (RAM-C) policy).

Solely asking for these sort of deliverables is not enough however to allow for these and other data sources to be used in evaluating system reliability. To trust that the provided data is truly representative of the system in its envisioned usage environment and is thus suitable for supporting the evaluation of the system, these Design for Reliability activities should all be linked together in a reliability case report type format. Data developed in a vacuum is unsuitable for evaluating system reliability. It must instead be demonstrably developed as an integral part of the overall design process for the system and the acquisition community needs to ensure contract language requires this evidence.

The Army has already begun passing some policy directives that give overarching guidance to what sort of new data sources are acceptable for augmenting the reliability analysis of MTA systems. As schedule constraints no longer allow for full system reliability testing, the new guidance specifies that greater reliance on contractor expertise and M&S will be necessary for evaluation of the reliability of MTA systems and that the T&E strategy should incorporate the reliability design activities as a primary source of data (Policy Directive for Test and Evaluation of Middle-Tier Acquisition Programs). The following section goes into what sort of considerations need to be made when developing reliability contract language for these MTA systems.

4.5. How to interpret the results

As the what and how of Army testing changes from FUSL to more focused subsystem and component level testing, the ways in which these newly accepted datasets are evaluated by the Army to assess system reliability will need to change as well. The end goal of the reliability analysis for a system remains to provide confidence that the system will achieve its system level reliability requirements and in turn that it will be operationally suitable for the soldier. This goal will now need to be achieved by piecing together data from numerous different data sources instead of the standard analysis of full system level data from OT.

The fact that numerous data sources will need to be combined together to assess system reliability necessitates that contracts must mandate the development and maintenance of a full system level reliability model. This overall model needs to be detailed enough such that different data can feed into reliability estimates of its different subsystems and components and that these all roll up together to provide an overall system reliability measurement. With a dependence on numerous data sources that all will become available and evolve over different timeframes, contracts must further mandate that the modeling solution must be readily updateable while also maintaining a sync with the system architecture model such that the reliability model is continuously representative of the current configuration of the system.

Different tools exist that can be utilized to develop and maintain these types of models. Some reliability modeling tools support rapid development of reliability block diagrams and fault trees for systems to define how failure of individual components and subsystems can lead to system failures. Another model possible is the implementation of Bayesian Statistics in order to correlate non-linear data sets in the overall evaluation of reliability (Figure 4).

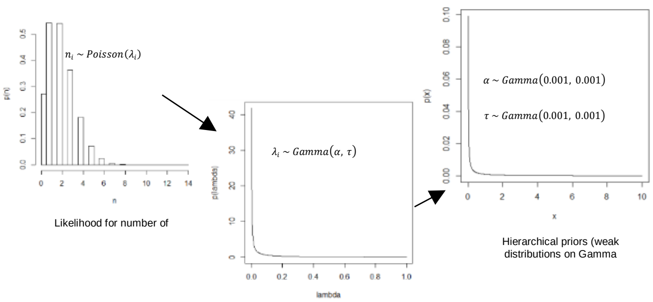

Figure 4: Bayesian Model Example

These tools support utilizing different data sources to provide estimates of reliability for different parts of the model and then easily roll up all of the information to estimate system level reliability. If these models are developed separately from the architecture models for a system however, they can quickly become out of sync and less representative of the system over time unless diligent coordination is carried out between the reliability and architecture teams as the system design evolves.

One solution to this is to carry out modeling of the system's architecture and reliability all within a single toolset designed to accommodate both. This is the ideal scenario and tools exist to allow this but they are not typically the tool of choice for modeling system architecture. For systems where other architecture modeling tools are utilized that do not support robust reliability modeling, a solution to this is to maintain a subset of smaller reliability models built within reliability modeling tools, with each representing a subsystem or component of the overall vehicle, down to an agreed upon level of detail. Then very basic automated reliability rollups can be built within the architecture tool that regularly queries these smaller models and rolls them up together within the architecture tool to provide an overall system level reliability rollup that stays in sync with the architecture model of the system over the development timeline. Regardless of which of these two approaches is taken, contracts must ensure that it is communicated that for data to be used in the evaluation of a system's reliability, it must be part of an integrated model that is representative of the system.

As to how the subsystem and component test data can inform the system's reliability model, there are numerous ways depending on the subsystem or component being tested, the type of test being carried out on it and the prior data for the reliability of the subsystem or component undergoing test. The most straightforward way the test data can inform the reliability model is when the testing profile can be directly correlated to the intended usage of the system. Then based on the quantity of testing and observed failures, a reliability distribution can be defined into the model for the subsystem or component. In other cases, if methods such as Highly Accelerated Life Test (HALT) are carried out to rapidly surface failures so that they can be fixed, this data could be meshed with historical reliability data for similar components along with assumptions about the distribution of fix effectiveness factors for the addressed failure modes to project out a likely range for the final reliability for that subsystem or component. There are many other ways to utilize the data from component and subsystem testing to populate the overall system reliability model as well. The contract needs to specify that the key to maximizing the amount of other data that can be utilized is to thoroughly document what is done in a logical and defensible way such that the context behind the overall system reliability model is fully understood.

5. EXAMPLE/CASE STUDY

In 2014, the ATEC and Tank Automotive Research, Development and Engineering Center (TARDEC) teamed together to create the Design of Automotive Reliability T&E (DART) effort. This effort aims to balance traditional proving ground testing with M&S in order to expedite an understanding of vehicle suitability characteristics in support of shorter decision cycle time.

The MaxxPro Long Wheel Base (LWB) Ambulance (Figure 5) was created through a vehicle modification from the original MaxxPro LWB vehicle. The upgrades from the original variant included an upgraded suspension to support increased weight, an ambulance capsule with medical equipment, and a medic seat/gunner platform conversion. Jointly led by TARDEC, the Army Evaluation Center (AEC), and Program Management Office (PMO) Mine-Resistant Ambush Protected (MRAP) the MaxxPro LWB Ambulance reliability test program utilized physical simulation to supplement traditional on-track reliability testing to increase program efficacy. The test program is the first to assess ground system reliability utilizing the data generated from laboratory simulation. Historically, vehicle suitability T&E has been reliant on test vehicles accumulating thousands of miles on a test course at a proving ground to surface system failures and "prove out" mitigations. Such testing optimally subjects the vehicles to the operational loads, stresses, and environments defined in requirements documents. However, traditional testing requires significant schedule, funding, and other valuable resources such as test and support personnel. By supplementing previous similar system data and truncated proving ground test with laboratory simulation, the team was able to realize a cost reduction exceeding $200,000 and a decrease of six months in schedule.

In order to evaluate the reliability of the MaxxPro LWB Ambulance, the team utilized data from three separate sources. It is important to note that without on-track testing, the laboratory simulation of the vehicle would not have been possible. The first data source was generated from the original MaxxPro LWB program, which accumulated 12,000 miles of on-track reliability DT from January 2011 through February 2012. This test generated 34 separate Operational Mission Failures (OMFs) (25 non-vibration) and resulted in a demonstrated reliability of 302 Mean Miles Between Operational Mission Failure (MMBOMF) at 80% confidence. During this test, the vibration induced failure modes included door hardware misalignment (4), various mounts/fasteners loosening (4), and battery terminal disconnection (1).

The second data source was generated from the MaxxPro LWB Ambulance program, which accumulated 3,000 miles of on-track reliability DT from September 2014 through February 2015 and 700 miles of on-track reliability operational testing from April 2015 – June 2015. This series of test events generated 6 separate OMFs (3 non-vibration) and resulted in a demonstrated reliability of 408 MMBOMF at 80% confidence. It is important to note that not only were the ambulance variant modifications added, but many of the failure modes listed in the first test were addressed with corrective actions. During the second tests, the vibration induced failure modes included shock mount bolt shear (1), tow bar mount loosening (1), and a battery terminal disconnection (1).

The third and final data source was also generated from the MaxxPro LWB Ambulance program; however, this test consisted of 9,000 miles of reliability simulation on a vertical actuator vehicle simulator. The simulator miles were validated and verified (V&V'd) by instrumenting the MaxxPro LWB Ambulance during the 3,000 miles of on-track reliability DT and then utilizing the accelerations and displacements seen during that test to build the damage profile and subsequent drive files on the simulator. This test event generated 4 separate OMFs and resulted in a demonstrated reliability of 529 MMBOMF at 80% confidence. This number was generated by combining the vibration induced failures with the extrapolated non-vibration based failures (3 OMFs x 3 sets of 3,000 miles = 9 total non-vibration OMFs). Of the most significance from this test was the failure modes themselves, which included a shock mount bolt shear (1), tow bar mount loosening (1), battery terminal disconnection (1), and a sub-frame crack (1).

Figure 5: MaxxPro LWB Ambulance on N-Post Simulator

The introduction of laboratory simulation on the formal evaluation of system reliability proved to be a successful venture. The team was able to utilize previous test data and engineering best practices to target test specific upgrades and engineering changes through hardware in the loop (HWIL) simulation. Upon completion of the test, the exact vibration induced failures surfaced during traditional on-track testing were also surfaced during vertical actuated simulator testing. By utilizing the instrumentation data gathered during the 3,000 mile DT, TARDEC was able to generate a V&V'd drive file which was absent of non-damaging segments and significantly compress the amount of time needed to replicate the damage on the targeted subsystems. The result of this test directly led to PMO MRAP's ability to achieve Full Material Release (FMR) quicker than traditional testing would have allowed.

6. Conclusion

This paper discussed how the changing environment in which the U.S. Army will need to operate will drive faster fielding of systems and system modifications. The paper describes how the automotive industry uses laboratory testing to accelerate vehicle testing. It proposes way that the U.S. Army can leverage and adapt these methods to its T&E mission. It concludes with a case study comparing field testing and lab testing for a MaxxPro LWB Ambulance.

7. References

[1] Combat Capabilities Development Command, "Campaign Plan 2019+," U.S. Army, 2019.

[2] US Army TRADOC, "Army Operating Concept: Win in a Complex World," US Army, 2014.

[3] US Army, "The Army Strategy," The US Army, 2018. [Online]. Available: https://www.army.mil/e2/downloads/rv7/the_army_strategy_2018.pdf. [Accessed 22 May 2019].

[4] US Army, "The Army Modernization Strategy," US Army, 2018.

[5] Assistant Secretary of the Army for Acquisition, Logistics, and Technology, "Army Test and Evaluation Efficiencies-Task Force Findings," US Army, 2013.

[6] D. Goure, "Will Proposed Reforms Really Make Army Acquisition Faster?," The National Interest, Dec 2017. [Online]. Available: https://nationalinterest.org/blog/the-buzz/will-proposed-reforms-really-make-army-acquisition-faster-23725. [Accessed 6 May 2019].

[7] United States Government Accountability Office, "Army Should Take Steps to Reduce Risk," 2019. [Online]. Available: https://www.gao.gov/assets/700/698858.pdf. [Accessed 6 May 2019].

[8] M. B. Madden, "Road Simulator Testing of Large Vehicles," in *SAE Mid-Year Meeting*, Detroit, 1970.

[9] G. E. Leese and R. L. Mullin, "The Role of Fatigue Analysis in the Vehicle Test

Simulation Laboratory," SAE , Warrendale, PA, 1991.

[10] MTS, "Investor Relations," MTS, [Online]. Available: http://investor.mts.com/. [Accessed 7 May 2019].

[11] S. Haeg, "Accelerated Vehicle Testing," April 2010. [Online]. Available: https://www.mts.com/en/forceandmotion/groundvehicletesting/MTS_003479?article=1. [Accessed 26 April 2019].

[12] R. A. Cripe, "Making a Road Simulator Simulate," SAE, Warrendale, PA, 1972.

[13] B. W. Cryer, P. E. Nawrocki and R. A. Lund, "A Road Simulation System for Heavy Duty Vehicles," SAE, Warrendale, 1976.

[14] J. W. Fash, J. G. Goode and R. G. Brown, "Advanced Simulation Testing Capabilities," SAE, Warrendale, PA, 1992.

[15] M. Brudnak and M. Pozolo, "Using Modeling and Simulation for Failure Analysis," SAE, Warrendale, PA, 1999.

[16] M. J. Brudnak, J. S. Walsh, I. Baseski and B. LaRose, "Durability Test Time Reduction Methods," SAE, Warrendale, PA, 2017.

[17] H. J. Lee, M. Brudnak, J. Walsh and B. LaRose, "Diminishment of Cuts in Durability Test Time Reduction Methods," SAE, Warrendale, PA, 2018.

[18] J. W. Fash, R. A. Lund, R. Pountney and K. White, "Enhancing the Design Development Cycle Through Computer Integrated Engineering or Durability," SAE, Warrendale, PA, 1987.

[19] K. Dressler, M. Speckert and G. Bitsch, "Virtual durability test rigs for automotive engineering," *Vehicle System Dynamics,* vol. 47, no. 4, pp. 387-401, 2009.

[20] S. A. Nolan and N. A. Linden, "Integrating Simulation Technology into Automotive Design Evaluation and Validation Processes," SAE, Warrendale, PA, 1987.

[21] J. Luna, S. Snider, M. J. Brudnak, B. LaRose, M. Morgan, D. Kosinski and M. Scott, "Getting Ahead of the Game: Quantifying the Value of Laboratory Testing Prior to Full-up System Level," in *Proceeding of the 2018 NDIA Ground Vehicle Systems Engineering and Technology Symposium (GVSETS),* Novi, MI, 2018.

[22] J. M. Dasch and D. J. Gorsich, The TARDEC Story: Sixty-five Years of Innovation, 1946-2010, Warren, MI: U.S. Army Tank Automotive Research Development and Engineering Center, 2012.

[23] P. Xu, D. Wong, P. LeBlanc and G. Peticca, "Road Test Simulation in Light Vehicle Development and Durability Evaluation," in *SAE World Congress*, Detroit, 2005.

[24] US Army, "Army Regulation 702-19: Product Assurance: Reliability, Availability, and Maintainability," Department of the Army, Washington, DC, 2015.

[25] J. D. Babbitt and D. Schlomer , "The Need for Speed," *Army ALT Magazine,* 22 April 2019.

[26] A. Karbassian, D. P. Bonathan and T. Katakami, "Nissan Speeds Up Truck Bed Durability Testing," *Automotive Engineering,* 20 April 2009.

[27] S. Anderson, "Tire-Coupled Road Simulation, Then and Now," January 2007. [Online]. Available: https://www.mts.com/en/forceandmotion/groundvehicletesting/MTS_003444?article=1. [Accessed 26 April 2019].

[28] R. A. Lund and K. H. Donaldson, Jr., "Approaches to Vehicle Dynamics and

Durability Testing," SAE, Warrendale, PA, 1982.

[29] S. S. You, C. Leser and E. J. Young, "Tools for Integration of Analysis and Testing," SAE, Warrendale, PA, 2003.

[30] A. Hurd, "Combining Accelerated Laboratory Durability With Squeak and Rattle Evaluation," SAE, Warrendale, PA, 1991.

[31] P. Xu, D. Wong, P. Leblanc, G. T. Peticca and G. T. Andrews, "Road Test Simulation Technology in Light Vehicle Development and Durability Evaluation," SAE, Warrendale, PA, 2005.

[32] R. A. Antonelli, M. D. Garcia, E. Rodrigues, L. C. Ferraro, H. Onusic, M. H. Tamashiro and C. L. Sanches, "New Policy of Durability Tests for Commercial Vehicles Through Development of Low Cost Road Simulators," SAE, Warrendale, PA, 2002.

[33] B. A. Suresh, D. J. Klinikowski and B. J. Gilmore, "Comparison of Vehicle Durability Testing Methods," SAE, Warrendale, PA, 1993.

[34] S. Haq, M. Temkin, L. Black and P. Bammel, "Vehicle Road Simulation Testing, Correlation and Variability," SAE, Warrendale, PA, 2005.

[35] C. J. Dodds and A. R. Plummer, "Laboratory Road Simulation for Full Vehicle Testing: A Review," in *SIAT 2001*, 2001.