2017 NDIA GROUND VEHICLE SYSTEMS ENGINEERING AND TECHNOLOGY SYMPOSIUM
SYSTEMS ENGINEERING (SE) TECHNICAL SESSION
AUGUST 8-10, 2017 – NOVI, MICHIGAN

ASSESSING THE TRADEOFF BETWEEN COST AND AVAILABILITY USING SIMULATION

Peter Vergenz
KBRwyle
Jacksonville, FL

Marc Banghart, PhD
KBRwyle
Jacksonville, FL

ABSTRACT

For large populations of vehicles, it is often difficult to estimate how changes to scheduled maintenance plans will impact future operational availability, especially when component failure rates may not be known precisely or the operational environment changes. The primary objective of this contribution is to illustrate a Modeling and Simulation (M&S) approach which determines the minimum amount of maintenance necessary to keep a given threshold of operational availability. The analysis was performed using discrete-event simulation, maintenance data, and anecdotal information from technicians. The information was combined within a model containing over 15 variables including labor and process constraints. The analysis yielded a decision tool that can be utilized to assess several potential long term storage maintenance policies, focused on cost minimization while meeting readiness requirements.

INTRODUCTION

For large populations of vehicles, it is often difficult to estimate how changes to scheduled maintenance plans may impact future operational availability, especially when the underlying component failure rate is unknown or if the operating environment changes. Processes such as Reliability-Centered Maintenance (RCM) exist to iteratively optimize a maintenance plan over time, but RCM will not provide detailed estimates of operational availability, nor does it provide a convenient way to assess all the indirect costs of alternative maintenance plans. To minimize cost, it is often desirable to determine the minimum amount of maintenance necessary to ensure a given threshold of availability. Analytical methods can be utilized to optimize the associated maintenance intervals. However, these techniques are typically complex and may not allow the decision maker to easily conduct extensive what-if analysis. This paper will demonstrate the use of simulation to study the tradeoff between cost and availability.

Case Study: Ground Vehicles in Storage

The U.S. government possesses numerous ground vehicles which are in "storage," yet must be maintained in a state such that they could be issued for operations within a matter of hours to meet contingency requirements. This unique storage situation required a tailored maintenance plan for these vehicles. A simple solution would be to treat these vehicles as if they were in an operational state. However, this would be wasteful because most tasks prescribed in the weekly, monthly, annual, etc. were developed for vehicles which are operated regularly. There is little value in changing the oil for a vehicle which has been in storage for a year.

After analyzing particular failure modes to determine logical maintenance tasks, consulting existing references for storing vehicles such as MIL-STD 3003 [1], and establishing preliminary preventative maintenance (PM) intervals, a discrete event simulation was developed to answer the following questions:

1. What is the maintenance cost of these vehicles in storage?
2. What is the expected operational availability of these vehicles?
3. How much labor and materiel are required? This includes requirements for PM, corrective maintenance (CM), and periodic issue and receipt of vehicles to operational units.
4. What would be the cost to achieve higher operational availability?

In addition to modeling the existing storage location with both current and possible future maintenance policies, the decision-maker also wanted to evaluate the potential costs of two other storage scenarios at an overseas location. This resulted in models for four scenarios. Each scenario had the same basic structure as shown in Figure 1.

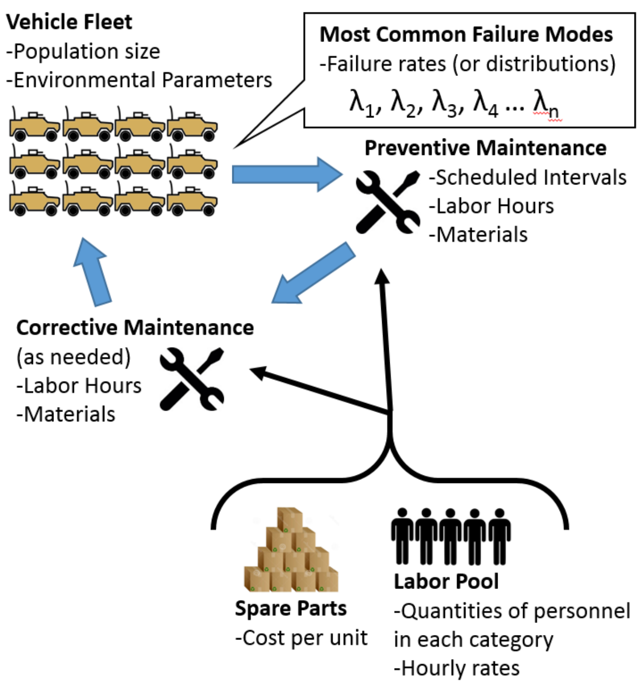

Figure 1: A basic overview of the model structure

MODELING AND SIMULATION APPROACH

Modeling and simulation (M&S) broadly refers to representing a system or process, followed by experimentation on the model in order to solve technical and managerial problems. In some applications M&S allows analysis of systems which are too complex to solve analytically. M&S further provides the capability to easily perform what-if analysis, once the model has been validated. The M&S process includes defining, designing and building the actual model, designing the applicable experiments to be conducted, collecting and analyzing required input data, model construction, validation, experimentation and analysis of model results.

M&S is frequently utilized in order to analyze complex systems, defined as a set of related elements within a stated boundary. Sufficiently complex systems are difficult to analyze without utilizing tools such as M&S due to several behavioral characteristics. In this case, each preventive maintenance event impacted availability in a different way. Additionally, the interactions between preventive maintenance interval length, multiple queues, labor constraints, and availability made analytical computation impractical.

Discrete event simulation is one method in order to analyze such complex systems and is well suited to supply chain, manufacturing and other process-focused problem domains. Discrete event simulation represents a system as a discrete sequence of events in time. Mathematically, the approach utilizes queuing theory and probability distributions to represent random processes.

The steps to conducting a discrete event simulation study could be described as follows:

1. Establish Assumptions and Gather Data
2. Consolidate Model Inputs
3. Develop Model Structure
4. Determine the Number of Replications Necessary
5. Validate Model Structure and Preliminary Outputs
6. Conduct Sensitivity Analysis

7. Compare Final Results

These steps will be explained in detail as applied to our particular case of ground vehicles in storage.

Establish Assumptions and Gather Data

As with any analysis, it is necessary to define some concepts and assumptions to establish a framework. To conduct this simulation, the following assumptions were developed to establish the analysis boundaries. The framework provides the context for identifying materiel, defining required tasks, and quantifying time and cost in order to model and quantifiably measure performance of each model scenario or maintenance strategy. For the sake of brevity, not all the assumptions utilized are discussed here.

No cost-effective maintenance plan can keep a fleet of vehicles 100% ready. Due to the probabilistic nature of material degradation and imperfect abilities to observe some conditions, as soon as PM is completed, there is some possibility that the system degrades below Fully Mission Capable (FMC). With this in mind, preservations and maintenance actions were developed so as to achieve an availability of approximately 90%. This means that if the vehicles in storage were brought out of storage for issue to units, at least 90% would be FMC and up to 10% may require some corrective maintenance.

It was assumed that all materiel resources are readily available to maintenance personnel. Although simulation has the capability to easily assess the impact of materiel resource limitations, labor constraints were the only restrictions considered in this model.

In this case, three years of vehicle maintenance data was used to determine the most common failure modes and develop estimates of failure rates. The team also gained access to the maintenance database in order to drill down into specific maintenance actions by vehicle serial number and location as needed. From those efforts, top degraders were developed and analyzed for applicability across all variants of vehicles analyzed. For many corrective maintenance actions, it was difficult to discern whether the failures occurred while the vehicles were in storage or if they were preexisting upon induction and not corrected at that time. To resolve some of these uncertainties, interviews of approximately eight mechanics and technicians who regularly conduct maintenance on these vehicles were also conducted at the vehicle storage location. These interviews further ensured that the analysis team firmly understood the process to be modeled, and allowed buy-in to be obtained and maintained throughout the analysis.

Consolidate Model Inputs

The model utilized numerous inputs. Some of these were provided by process owners, while others were developed by analyzing failure data, maintenance manuals, or estimations provided by subject matter experts. These model inputs included the following:

- Vehicle quantities
- Preliminary PM intervals
- Existing quantities of personnel for each of eight labor categories
- Hourly rates for each labor category
- Additional semi-fixed costs per person for overseas locations (housing, etc.)
- Process times or distributions for
    - PM actions
    - CM actions
    - Data entry
    - Inventory management
    - Supervisory actions
- Itemized materiel costs for PM and CM
- Failure rates in storage (for both indoor and outdoor storage)
- Work schedules for each location
- Frequency, vehicle quantity, and duration of periodic operations

Develop Model Structure

Simulation software packages implement queuing theory utilizing varying nomenclature or structural elements that form the model. However, the underlying mathematical theory holds regardless of the specific software package utilized. Within Arena discrete event simulation software, there are several types of model structure elements which were utilized in the simulation. First, there are entities. Entities are the elements which flow through the model, may change status, may affect other entities, and affect the output metrics of the model. Kelton describes entities as the "players" of the simulation [2]. Another element within Arena is attributes, the properties of entities which define states. For instance, whether an entity is in an "up" or "down" status would typically be captured by use of an attribute. Another model element is variables, which are global values which may be used by the model in various ways. Another model element is resources, which are elements which are consumed or utilized when entities undergo various processes. Resources are typically used to model labor or materiel used during processes. Lastly, queues are areas where entities may build up if they are unable to move on to the next process for some reason. This may be due to lack of resources or limited capacities in the next process. Queues are automatically generated within Arena for most constrained processes. For more information describing modeling and simulation within Arena, see reference [2].

The model was structured with the vehicles as entities. Personnel and materiel were modeled using resources. Multiple distributions were used to model each process time to accommodate the different times required for different variants, statuses, and environments of vehicle.

At time zero of each replication, the appropriate amount of vehicles for each variant enters the model. Each of these vehicles is assigned attributes which define the variant and storage status. In the next step, each vehicle is assigned a unique serial number, a future date at which each of its PM actions is due (based on staggering all the vehicles throughout each interval) and a future date on which a failure will occur for that vehicle based on the failure rate for that variant, status, and environment using the exponential distribution for time between failures. Each vehicle then remains in a hold until a PM comes due. When a failure occurs, the vehicle is pulled from the hold, assigned a "failed" attribute, and then put back into the hold. Each vehicle can incur multiple failures at a time. The failure remains on the vehicle until it is identified during a PM and corrected during CM.

When a PM action comes due for that vehicle, it goes through several modules which determine exactly which PM actions are due. If a particular PM is due, it goes through the appropriate processes which accrue labor hours for the appropriate resources. Each process within the model utilized the specific process time defined for that variant, status, and environment. The model also includes logic which diverts vehicles from certain processes which are not applicable in certain scenarios (i.e. particular locations which lacked certain resources).

After PM processes are completed for a given vehicle, the vehicle progresses to the CM processes. If no failures were identified, the vehicle exits the CM processes and goes back to the hold. If failures were identified, it continues through the CM processes, sometimes multiple times if multiple failures are identified on a vehicle. Once all failures on a vehicle have been corrected, the availability variable for the model is updated.

After CM, the model checks whether vehicles are needed for an operation. If vehicles are needed for an operation, they are routed to the operations area. The vehicles receive a Limited Technical Inspection (LTI) and associated processes then go into the operations queue until the necessary amount of vehicles are collected. Once the appropriate amount of vehicles are collected, the group of vehicles goes on the operations process for the duration specified for that instance based on the input distribution. After vehicles complete the operation, they again receive an LTI and associated processes, then go back into the hold for PM or failures.

There are also "logical entities" and associated modules defined within the model to screen for PMs, screen for failures, collect statistics, and record model outputs.

Determine the Number of Replications Necessary

To obtain results with the appropriate resolution for the application, numerous replications are typically necessary when conducting simulation. These replications could be viewed as observations in a sample being taken from the population. For each metric of interest, such as cost or availability, the results reported are typically the mean values from all the replications conducted. However, each replication results in a slightly different value. If the variation across these replications is large, there is uncertainty in the metric of interest, hence the necessity to determine the number of replications (sample size) necessary to achieve sufficiently precise results. "Sufficiently precise," specifies the level of detail necessary for the given metric of interest. For instance, there is probably no reason to calculate the mean availability down to 6 decimal points. If we know availability is 90.9% +/- 0.25%, this is probably sufficiently precise for this metric. This variation is termed the acceptable error. So if results have been obtained for 20 replications of a model have been performed, how does one determine whether 20 is enough?

The confidence interval for an estimation of a metric is often given with an equation using the standard normal distribution. However, the use of the normal distribution for determining the confidence interval from a sample is only valid if the population standard deviation is known or the sample size is large, which is often not the case. Sometimes the population standard deviation is assumed to be equal to the sample standard deviation, but this assumption may be avoided by using the Student's t distribution, which does not require a known population standard deviation. Compared to the normal distribution, the Student's t distribution is also shaped like a bell curve, but it has fatter tails than the normal distribution. As the degrees of freedom (dof) or sample size increases, the tails get thinner to the point that once the sample size is over 100, the Student's t distribution is approximately equal to the normal distribution, so use of the normal distribution could be valid for these large samples. Some sources say that a sample size of 30 is sufficient for this purpose, but it depends on the level of precision required. Because the normal distribution could induce miscalculations when the sample size is small, it is recommended to always use the t distribution, especially if conducting the calculations in Excel or some other automated tool. See reference [3] for more about the t distribution. Using the Student's t distribution, the confidence interval will be

$$Confidence\ Interval = \bar{x} \pm s_M t_{dof,\propto/2}$$

Where

$\bar{x}$ = mean of the sample (for all replications)
s_M= standard error of the mean
$t_{dof,\propto/2}$= value from t distribution, given $\propto$ and dof
$\propto$ = 1 – (confidence level)
dof = n – 1
n = number of replications (sample size)

While some simulation software suites automatically calculate s_M, the standard error of the mean, not all programs do this. If not, it can be calculated using:

$$s_M = s/\sqrt{n}$$

Where

s = sample standard deviation, and
n = number of replications (sample size)

In the model developed, availability is the metric of interest which was determined to have the most variation so it was analyzed to determine the number of replications necessary. Additionally, multiple scenarios were being modeled and the particular scenario with the smallest population size was selected for this step. A smaller population size will result in larger variation due to the higher probability of extreme mean values. To conceptualize this in plain terms, if you are repeatedly flipping a coin, a smaller number of flips is more likely to result in an extreme value (for instance all heads) than a larger number of flips. The probability of getting all heads in five flips is much more likely to occur than getting all heads on 20 flips. For this reason, the model scenario with the smaller population was selected to determine the number of replications needed.

Results from several iterations of simulations are presented in the tables below. A 95% confidence level that the mean availability is within 0.25% was chosen as the level of precision desired. In other words, if the model predicts 90.5% availability, continued replications of the model may result in a mean availability as low as 90.25% or as high as 90.75%. Values in the following descriptions are shown as decimals rather than percentages to avoid possible misinterpretations.

Number of Replications	Mean Availability	Standard Error of the Mean	Confidence Interval Half Width
3	0.9090	0.0031	0.01474
5	0.9084	0.0034	0.01027
10	0.9052	0.0033	0.01012
20	0.9056	0.0012	0.00283
25	0.9055	0.0010	0.00250
30	0.9049	0.0009	0.00210
50	0.9050	0.0007	0.00160
70	0.9057	0.0006	0.00140
150	0.9056	0.0004	0.00095

Table 1: Means, standard errors, and confidence interval half widths for various quantities of replications

Suppose we had initially conducted 20 replications. At that point, the mean was 0.9056 and the standard error of the mean was 0.0012. At a 95% confidence level ($\propto = 0.05$ or $\propto/2 = 0.025$), with 19 degrees of freedom (n-1), we can look up in a table (or calculate in an Excel formula) that $t_{dof,\propto/2} = 1.729$. Substituting the values for t and $\bar{x} = 0.9056$ into the equation, we calculate the confidence interval. Using this calculation, we can say at a 95% confidence level that the mean is between 0.9031 and 0.9082, or a confidence interval half-width of 0.0101. This range is outside our desired precision of 0.0025, so more replications would be necessary to achieve the desired precision in our results. At this point, we would run the model as many more times as necessary to achieve the confidence interval that is sufficiently precise. The figure below shows the means and confidence intervals for a range of replication quantities on the model described.

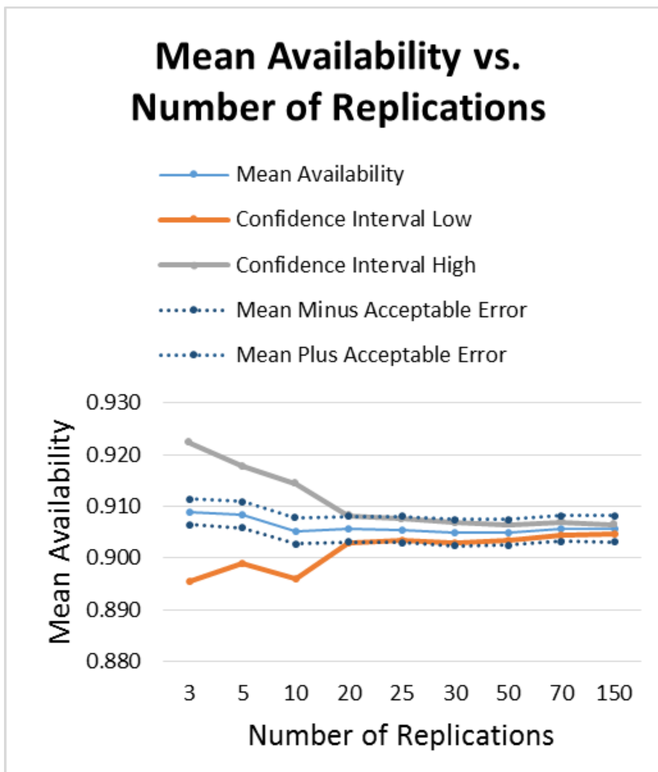

Figure 2: Means, confidence interval half widths, and acceptable error for various quantities of replications

In this particular case, it was determined that 25 replications will yield a confidence interval that is within the desired precision. For more information on determining the number of replications needed, see references [4] and [5].

It is important to keep in mind that this does not take into account possible errors in the input parameters or the structure of the model. In the case of modeling and simulation, the confidence interval merely means that if more instances of the model were replicated, the mean from all replications would not be likely to shift outside that interval. This is an important distinction that is often misinterpreted when model projections are presented with a confidence interval. Additionally, projections from the model assume the parameters are accurate. If for instance, the failure rates double due to increased humidity in the environment, the real world performance would not be likely to fall within the predicted confidence interval.

The value of models is the insight into relationships in the performance of the system with regard to what is important and how it influences performance of the processes being modeled. In that sense, modeling is best understood as a decision support tool providing a quantifiable understanding of how processes and systems operate. Model predictive value can be improved by refining the model using comparisons of the model output to actual data collected through a validation and verification process. In short, a model based on past performance does not guarantee future results.

Validate Model Structure and Preliminary Results

To validate the structure of the model, several In-Process-Reviews were conducted to ensure the logic was developed accurately. Additionally, at the beginning of the project, the team developing the model had visited all maintenance sites being modeled to gain an understanding of the processes in place. Additionally, further interviews of maintenance supervisors were conducted after the model was built. Upon development of preliminary results, additional In-Process-Reviews were conducted to ensure the resulting behavior of the model matched what would be expected by the process owners and stakeholders.

After maintenance policy decisions are made and new procedures are implemented, there will likely be opportunities to further assess the validity of the model. It could also likely be improved based on revised information. Data should be collected to validate and further refine the input parameters, particularly in areas where robust data may not have been available at the time of model development. This data could be collected by the personnel performing maintenance, then used in the model to further improve its validity for future use when conditions change, other scenarios are necessary, or other maintenance policies are considered.

Conduct Sensitivity Analysis

Sensitivity analysis is the process used to determine how changes to input values would likely affect output values of a model. This could be useful to know for a variety of reasons. For instance, some of the input values may change over time. Predicting the expected system performance due to these changes would have obvious benefits. Additionally, in cases where input parameters are estimated due to incomplete data, it would be useful to know how imprecise estimations of different input parameters would affect the model results.

There are several possible methodologies commonly utilized to conduct a sensitivity analysis. The simplest form is a One-At-a-Time (OAT) approach. In this case, one input parameter is varied at a time while the others are held constant. The changes to results are then observed to determine the sensitivity of the model to each of the input parameters. Several Design of Experiment (DoE) approaches to include full factorial and Taguchi designs can be used. Full factorial designs vary all possible combinations of all factors at many levels, while Taguchi designs fraction the levels of each factor required to reduce the required number of trials or experiments.

The advantage of the OAT approach is that is does not require many trials and results are fairly straightforward to interpret. The disadvantage of the OAT approach is that it does not capture interaction affects. For example, a given output may not be affected by either of two input parameters changing by themselves, but the output may be affected when multiple factors change at the same time.

The advantage of DoE approaches are that they will capture interaction affects if conducted properly. However, the disadvantages are that they require significantly more trials and the results require more sophisticated statistical methods to interpret (Analysis of Variance). For a full factorial DoE approach, the number of trials required is given by k^n, where k is the number factors examined and n is the number of levels examined for each factor. A sensitivity analysis examining three factors at five levels would require 3^5, or 243 trials for each output variable examined. Each trial requires multiple replications as described previously, so the time required to perform a full factorial DoE approach is often significant.

In the sensitivity analysis for this model, the OAT approach was utilized. Three factors were examined: the PM interval lengths, materiel costs of all maintenance actions, and the process times. The interval lengths were examined at seven levels and the other parameters were examined at three levels. The effect of these changes were observed on both availability and total cost (although not shown in this report for brevity).

COMPARE RESULTS

After performing all the steps previously described, the results from the models of each scenario were analyzed and compared. These results included total costs for PM, CM, and operational issue; costs per vehicle for each of those metrics (because different scenarios had different vehicle quantities); labor hours for each category (total and per vehicle); and the recommended personnel staffing levels for each scenario based on the labor hour requirements and avoidance of bottlenecks negatively affecting availability.

For the location with existing policies already in place, the model showed that significant savings per vehicle could be achieved by extending all the existing PM intervals. Additionally, the simulation gives stakeholders confidence that extending the PM intervals will not degrade operational availability below desired levels. The intervals were extended only to the point that the availability target of 90% could still be achieved. Some of the results from these simulations are shown in the figures below.

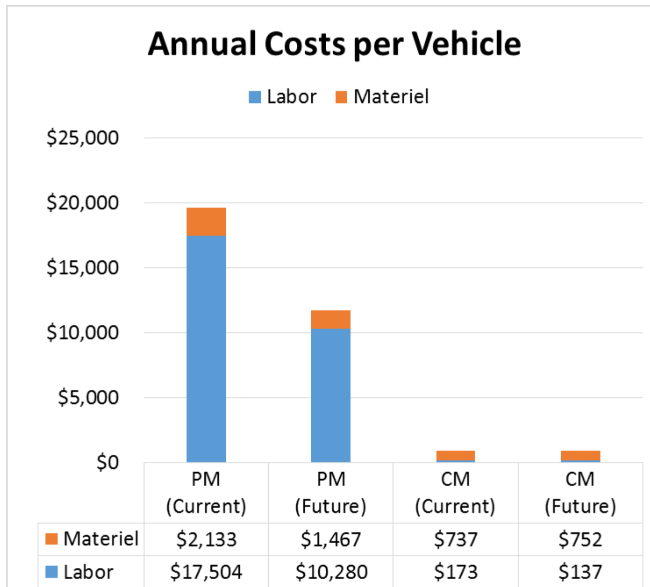

Figure 3: Annual costs per vehicle modeled for the current state and future state

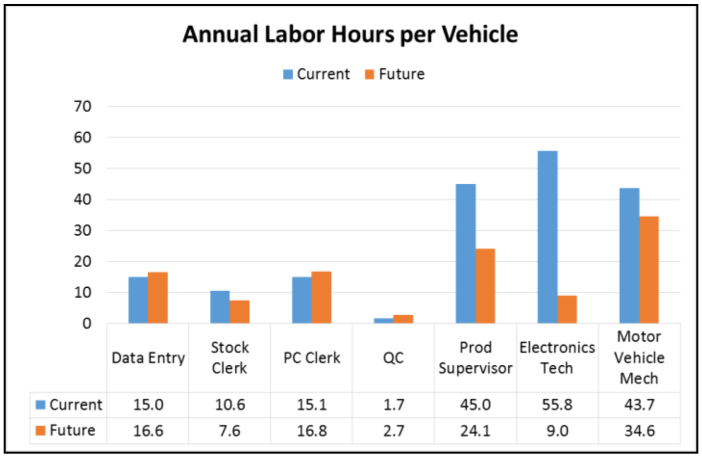

Figure 4: Annual labor requirements per vehicle modeled for the current state and future state

A benefit of conducting sensitivity analysis for both cost and availability is that these outputs may be plotted to observe their relationship to each other. It is commonly understood that maintaining a higher level of availability costs more, but the quantification of this trend can be difficult. The simulations conducted for the sensitivity analysis enable development of this cost curve. Note that the costs increase much more rapidly as availability goes beyond 90%. The availability curve will continued to approach 100%, but will theoretically never reach it. This is because no matter how often you inspect or PM a vehicle there is always some possibility of a failure occurring shortly after that.

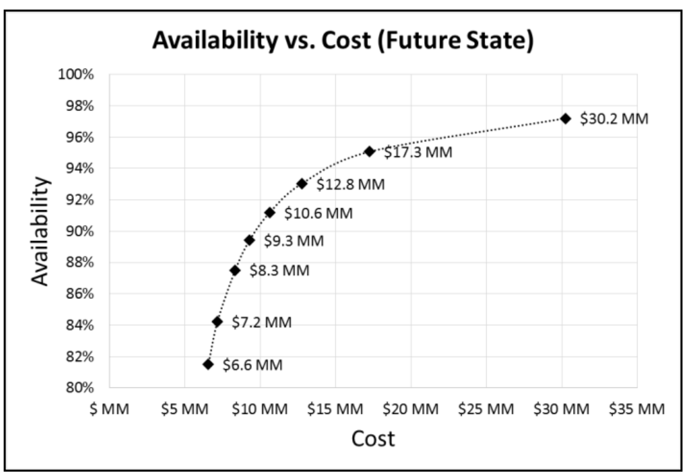

Figure 5: The tradeoff between cost and availability derived from simulation

MODEL LIMITATIONS

All models are an abstraction of the underlying system, thus an error (although small) will always exist between the true system and the model. The model structure should be periodically reviewed, and any additional information that becomes available can be integrated. Additionally, if sufficient information is available, model assumptions can be further relaxed.

Military field data poses significant challenges due to likely error within the data sets, as well as the limited amount of information that is typically captured as part of a work order. The vehicles in storage should be closely monitored as new processes derived from the M&S analysis are implemented. Monitoring will allow data to be collected measuring the effectiveness and efficiency of these new processes as well as validate the availability predictions of the future state. Additionally, the model could be improved by further refining some of the input data. In particular, some of the failure rates were estimated using anecdotal data due to lack of this information in the maintenance database.

CONCLUSION

Ultimately, every military vehicle program strives for additional availability at a lower cost. Managers need to make decisions about budgets, manpower, and maintenance policies which affect the availability of their systems. While analytical solutions can often yield reasonable estimates for certain manpower or cost problems, modeling and simulation provides a way to solve problems which are too complex to solve analytically. It provides a way to assess possible courses of action without spending resources on pilot studies to physically test alternative policies.

REFERENCES

[1]Department Of Defense, “MIL-STD 3003B Preparation For Shipment And Storage of Wheeled Vehicles”, United States, 2007.

[2] D. Kelton, R. Sadowski, and N. Zupick, "Simulation with Arena, 6th Edition", New York, NY, 2015.

[3] "NIST/SEMATECH e-Handbook of Statistical Methods", http://www.itl.nist.gov/div898/handbook/ accessed April 2017

[4] K. Murphy, "How Long Should I Simulate, and for How Many Trials? A Practical Guide to Reliability Simulations", Institute of Electrical and Electronics Engineers (IEEE) Reliability and Maintainability Symposium (RAMS), United States, 2001

[5] M. Driels and Y. Shin, "Determining the Number of Iterations for Monte Carlo Simulations of Weapon Effectiveness", Naval Postgraduate School, United States, 2004