


Simulation design and analysis to support decision making in the management of capability programs

D. Jusaitis^a , K. Baddams^b, N. Sardesai^a and R. J. Anderson^b

^a Shoal Group Pty Ltd, ^b Defence Science and Technology Group

Email: russell.anderson@dst.defence.gov.au

Abstract: This paper discusses a study into the initial application of Design of Experiments (DOE)-based sensitivity analysis to low-fidelity simulation of Program-level capability scenarios, in order to provide statistical robustness to evidence, and greater confidence to early-stage decision making. The following excerpt from Hallett et al. (2018) provides the context surrounding scenario modelling work from which this study is motivated:

Following the 2016 First Principles Review, Defence introduced an updated Capability Life Cycle and a Program Management construct to develop and manage Defence Capabilities for the Joint Force. Conventional wisdom holds that to realise the potential of the Joint Force and the capabilities that support it, careful analysis of the capabilities and their integration and interoperability is required. Programs are a step towards more formally managing the complex capabilities and their integration and interoperability demands and are in addition to the [existing] Project/Product level management.

The Joint Capability Narrative, the Joint Capability Needs Statement, the Program Strategy, and the Program Integrating Operational Concept are elements that provide a top-down driven understanding of what a capability must achieve, where it needs to operate, and with whom it needs to operate. The central theme of all these elements can be portrayed in a set of scenarios covering the operation and support of the capability.

An evidence base for decisions is essential. Hallett et al. (2018) explain that “demands on the Defence budget put pressure on Programs to provide robust capability-based business cases” to support decision makers, and that “[t]he business case should include new Project justification”. Foundationally for this study, Jusaitis & Cook (2018) demonstrate how supporting evidence “can be derived from analysis of the current and proposed capabilities in Program scenarios against an agreed set of Program Measures of Effectiveness”.

Hallett et al. (2018) describe and propose an executable model-based approach to Program scenario analysis, and Jusaitis & Cook (2018) explore the practical underpinnings of this methodology and discuss the development of a model framework. The framework supports simulation and analysis of Program capability scenarios, to produce first-order estimations of the effectiveness predictions that can demonstrate the feasibility of the capability design.

The executable model examined in this study addresses a straightforward hypothetical air-defence capability scenario. The goal for the initial application of DOE was to define performance envelope boundaries for a defence fighter platform intercepting a range of attacking opponent platforms. In preparation for DOE, the model was designed to receive fourteen input factors and to provide six output responses. Responses were captured as durations at select events during each simulation. A Nearly Orthogonal Latin Hypercube Design was applied together with factor screening statistical analysis to inform the removal of non-significant factors and support selection of a response-of-interest. This was followed by a full-factorial design focussed on the remaining factors with a regression analysis to inform selection of significant factors for response optimisation. Finally, sensitivity analysis was run to understand which significant factor combinations provided optimal responses; these factors were then used to define the boundaries of fighter performance envelopes against various opponent platforms.

Analysis of the performance boundaries for an existing hostile capability against those of a predicted future hostile capability provided insight into the capability gaps to retain capability overmatch with a future opponent. This result demonstrates that DOE offers a valid and useful approach to providing a level of statistical robustness to the analysis of low-fidelity capability scenarios.

Keywords: Executable architecture, operational architecture, design of experiments, capability design

1. INTRODUCTION

This paper discusses the initial application of Design of Experiments (DOE)-based sensitivity analysis to low-fidelity simulation of Program-level capability scenarios, in order to provide statistical robustness to evidence, providing greater confidence to early-stage decision making. The following excerpt from Hallett *et al.* (2018) provides the context surrounding scenario modelling work from which this study is motivated:

Following the 2016 First Principles Review (Defence, 2015), Defence introduced an updated Capability Life Cycle (CLC) and a Program Management construct to develop and manage Defence Capabilities for the Joint Force. Conventional wisdom holds that to realise the potential of the Joint Force and the capabilities that support it, careful analysis of the capabilities and their integration and interoperability is required. Programs are a step towards more formally assessing and managing the complex capabilities and their integration and interoperability demands and are in addition to the [existing] Project/Product level management. Defence has created around 40 Programs each of which is “a group of related Projects, Products, and activities that are managed in a coordinated way to optimise the capability outcome within allocated resources” (Defence, 2017).

The Joint Capability Narrative, the Joint Capability Needs Statement, the Program Strategy, and the Program Integrating Operational Concept are elements that provide top-down driven understanding of what the capability must achieve, where it needs to operate, and with whom it needs to operate. The central theme of all these elements can be portrayed in a set of scenarios covering the operation and support of the capability.

Hallett *et al.* (2018) go on to describe and propose a model-based approach to Program scenario analysis. They claim that one key to implementing this is to adopt “a tool that allows both descriptive and analytic models to be developed and links between separate Program and Project models to be made”. Jusaitis & Cook (2018) explore the practical underpinnings of this methodology and discuss the development of a model framework to support simulation and analysis of Program capability scenarios, in order to produce first-order estimations of the effectiveness predictions that can demonstrate the feasibility of the capability design.

This study builds on the simulation underpinnings described by Jusaitis & Cook (2018) in support of the Hallett *et al.* (2018) methodology by examining the initial application of sensitivity analysis using a Design of Experiments (DOE) approach to the low-fidelity simulation of Program capability scenarios. The objective was to show that this approach could provide a level of statistical robustness to performance predictions in support of early-stage decision making and the ongoing management of capability effectiveness.

2. BACKGROUND

Defence has many descriptive operational architecture models that have been, and are being, developed during the CLC strategy-and-concepts phase. Traditionally this has been for Projects, but now increasingly for Programs. These models have been developed primarily to derive parameters that define the capability; however, many have capability scenarios that are modelled using Enhanced Function Flow Block Diagrams (EFFBDs), an executable graphical language that can be simulated using Discrete Event Simulation (DES).

Hallett *et al.* (2018) and Jusaitis & Cook (2018) have demonstrated that with further development these descriptive models are able to support low-fidelity simulation-based scenario analysis, that can be applied to provide support for capability Program management and decisions. However, when considering the effect of several scenario inputs (factors) on an output (response), there will be an optimal set (or sets) of factors that will achieve a desired response. Montevechi *et al.* (2012) explain that simulation alone does not serve as an optimisation technique. Without employing a tool to support simulation analysis, practitioners are left to simulate limited capability system configurations and simply choose the one that appears to result in the best integrated capability performance.

Seeking to address this gap, the authors explore the application of DOE; a structured statistical analysis and optimisation approach that allows for insight into a model to determine its sensitivity to various factors, and cause-and-effect relationships (Santner *et al.*, 2018). This approach was selected over more exhaustive approaches like Monte Carlo to reduce the time and computing power required. The objective of this study is to increase the statistical rigour behind the outcomes from such scenario analysis and to understand if the optimisation techniques could provide evidence to support Program capability decisions. Section 3 provides

an overview of the executable model, Section 4 outlines the methodology adopted, Section 5 summarises the results, Section 6 offers some discussion, and Section 7 concludes.

3. THE EXECUTABLE MODEL

The executable model addressed a hypothetical dynamic targeting scenario developed around an Airborne Early Warning and Control (AEW&C) aircraft providing Command and Control (C2) and monitoring an Area of Operations (AO) for opponents. Fighter aircraft are operating a Combat Air Patrol (CAP), providing air superiority and protecting a maritime High-Value Unit (HVV) within a Maritime Task Group (MTG). The scenario is illustrated in Figure 1.

The scenario activity flow sees the AEW&C detect an inbound opponent aircraft and task a fighter to intercept. The fighter proceeds to intercept and launches Air-to-Air Missile/s (AAM) attempting to neutralise the opponent before it launches an Air-to-Surface Missile (ASM) at the HVV.

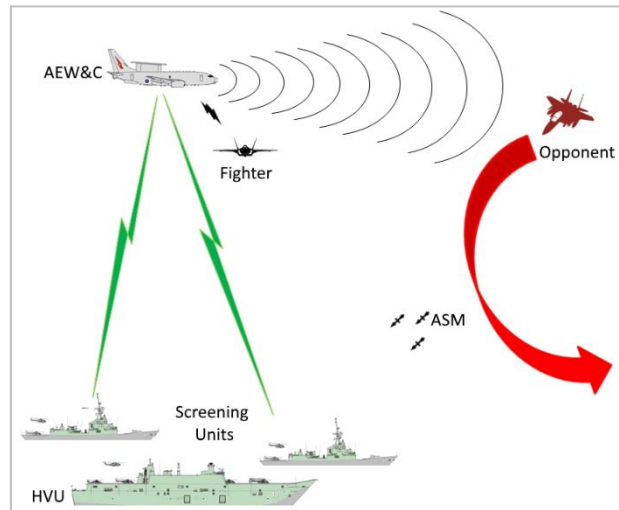


Figure 1. Scenario illustration

The activity flow of the executable model top-level scenario is shown in Figure 2. Table 1 summarises each operational activity (OA). There is further complexity below this high-level summary of the scenario, where decisions made by each performer are based on calculations of the locations and speeds of other performers¹.

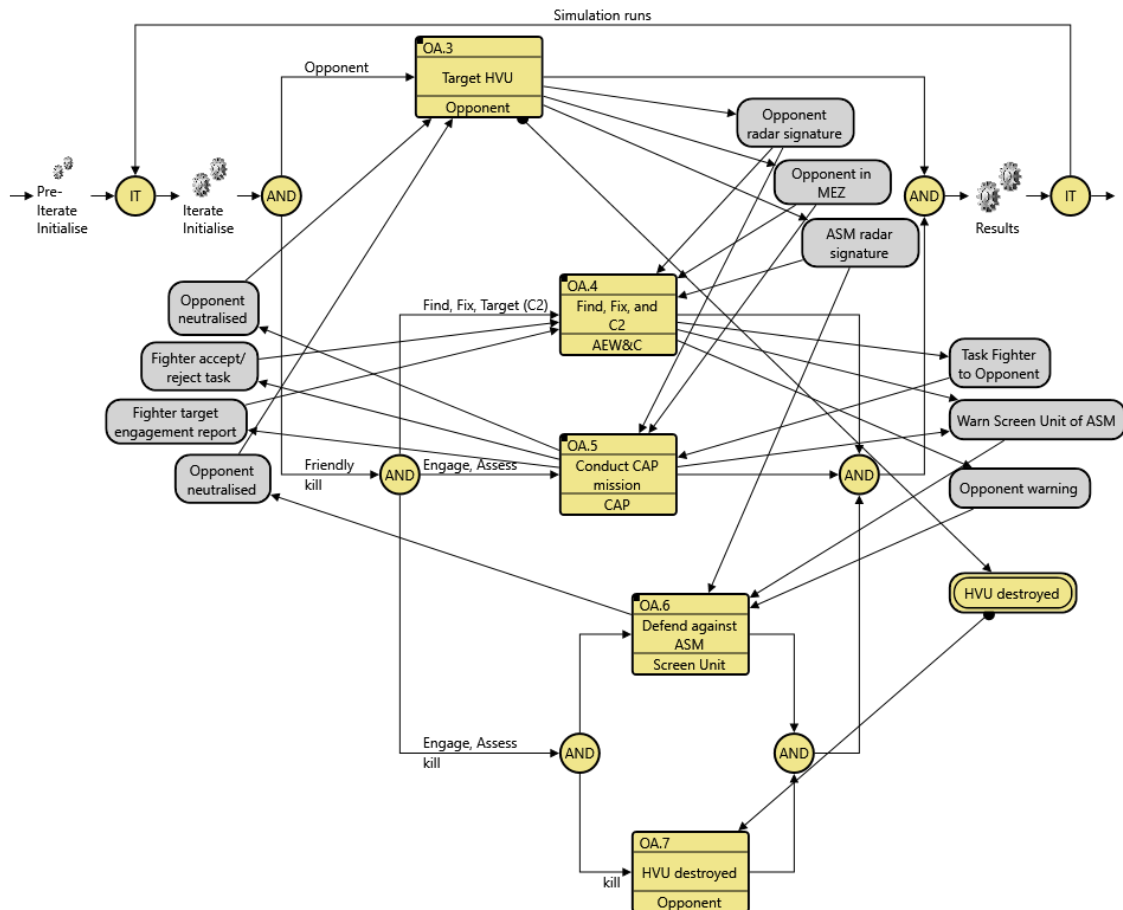


Figure 2. Scenario activity flow²

¹ See Jusaitis & Cook (2018) for a detailed overview of the executable model development approach.

² MEZ is the acronym for Missile Engagement Zone.

Visible in Figure 2 are also the simulation management activities (signified by gear icons). These activities do not represent the behaviour of the dynamic targeting scenario but are used to manage the multiple simulation runs required to perform the DOE.

Table 1. Scenario operational activity descriptions

Activity reference	Activity description
OA.3: Target HVU	Models the activities carried out by the opponent to engage and destroy the HVU.
OA.4: Find, Fix, and C2	Models the activities, including C2 of the AEW&C. For hostile contacts, the AEW&C will assess effector (fighter) availability and task accordingly.
OA.5: Conduct CAP mission	Models the activities of the Combat Air Patrol (CAP) mission. Focussed on fighter activities in response to being tasked to neutralise the opponent.
OA.6: Defend against ASM	Models the activities of the MTG, focussed on the Screening Unit activities to defend the HVU from the incoming ASMs.
OA.7: HVU destroyed	Only run if the ASM is successful in avoiding Screening Unit counterattacks, and consequentially successful in hitting the HVU. This activity ends the scenario.

Given the exploratory nature of this study, several general assumptions were made during model development. These assumptions were largely based on sensible simplifications; for example, fighter radar range being equivalent to its targeting range, and restricting the AEW&C, HVU, and screening units within the same vicinity. A full list of assumptions is not included here as the focus of this paper is on the applicability of DOE for low-fidelity executable models to provide a level of statistical rigour in support of decision making.

4. SIMULATION AND ANALYSIS METHODOLOGY

The approach to applying DOE was informed by Schmidt & Launsby (1994) and the DOE sensitivity analysis approach documented by Montevechi *et al.* (2012). The following sub-sections summarise each step.

4.1. Definition of model factors and responses

Factors (inputs) were selected based on values that were likely to vary between different scenarios and solution systems or have an impact on the operational scenario. Fourteen factors were initially selected covering both the friendly and opposition performers and their activities in the scenario. Default opposition and fighter speeds, missile speeds, and launch ranges were selected, along with radar range, information timeliness factors, and distances between the performers. These values could then be changed to represent different levels of capability and different relative locations for the performers and thus allow for insight to be derived from measuring the effectiveness of a large range of scenarios. Similarly, six model responses (outputs) were defined and captured as the time at which selected events occurred. The two selected for examination through the process were: the time of AAM arrival at the opponent and the time of ASM launch by the opponent.

To correctly define the factors required for the DOE, suitable upper and lower bounds are required. This is to inform the design of the simulation runs, such that the factors can be simulated at a few levels between and including the limits to determine the effect on the response. Upper and lower bounds for the factors were guided by the authors' objectives and informed by Subject Matter Experts (SMEs). They were representative of the present and future capabilities.

4.2. Model Verification and Validation

Model Verification and Validation (V&V) is essential if the model is to be used to accurately inform decisions. Montevechi *et al.* (2012) warned that using a non-validated model can lead to erroneous conclusions and undesirable decision making. For this study, an SME was engaged throughout model development to conduct a moderate level of V&V by observation and examination of the simulation behaviour and responses based on several sets of factors, covering low, mid, and high values, and values above and below the defined ranges. This level of V&V was considered appropriate for the purposes of the study.

4.3. Screening design, simulation, and analysis

'Levels' are the steps between the upper and lower bounds of a factor that are varied for each run defined in an experiment design. Although a minimum of two levels is required for DOE, further levels may increase the fidelity of results, especially if non-linear relationships are suspected. However, more levels also increase the

number of required simulations, which takes more time. The simulation duration was not significant, and five levels were used (inclusive of the upper and lower bounds) as a trade-off between granularity and time.

A full-factorial experimental design defines n^k simulations, where n represents the number of levels of the variable and k is the number of factors that will be varied. Varying all fourteen factors over the five levels would require many simulations ($n^k = 5^{14} = 6,103,515,625$). To reduce the time required, a DOE screening approach called Nearly Orthogonal Latin Hypercube Design (NOLHD) was adopted. NOLHD is a statistical method for generating sets of factor values to provide a ‘space-filling’ representation of the experimental region with minimal unsampled regions (Cioppa & Lucas, 2007).

Using this approach, the required number of runs was reduced significantly down to 70 (the number of factors multiplied by the number of levels (Air Academy Associates, 2019)). NOLHD is optimised for main effects. It is possible that a coupling effect could exist between factors, that is the impact of a given factor is dependent on the level of another factor. This can also be handled by the NOLHD, with the design allowing for this when generating factor sets for each run. However, the decision was made not to examine coupling effects in the analysis until later full-factorial runs.

The model was simulated based on the NOLHD and the responses were recorded. Regression analysis was applied to the responses and the non-significant factors were removed using backwards elimination. The DOE tool automatically applied Analysis of Variance (ANOVA). The next stage focussed on the remaining six significant factors and the two responses-of-interest: time of AAM arrival and time of ASM launch.

4.4. Full-factorial design, simulation and analysis

With only six significant factors, a 3-level full-factorial design was then conducted. This design required 729 runs to simulate all possible permutations of the high, medium, and low values, and allowed for cross-coupling effects to be identified. A fractional-factorial design with two levels could have reduced the number of runs (and thus resources) required, but with relatively short runs a full-factorial design was examined, and three levels provided better fidelity. Regression analysis was performed on the responses to quantify the effect of each of the factors and to determine a best-fit predictive model. This identified *Opponent Speed* as the only common significant factor. Table 2 and Table 3 show the best-fit models and the interactions between the significant factors and responses-of-interest. The sixth factor, *AAM speed (E)*, was no longer identified as significant.

An issue occurred where AAM arrival time was undefined when the fighter could not successfully engage the opponent. To produce the model these times were set as the respective simulation end time. However, whilst the adjusted R^2 value was reasonable, there were extreme standardised residuals and further investigation is needed to understand this. Regardless, the model still provided a strong indicator of the significant factors.

Table 2. Best-fit model AAM arrival time, Adjusted $R^2 = 0.8966$

Factor	Coefficient	Factor	Coefficient
Const	1149.88	DF	227.93
Dist. to Opponent (A)	1327.06	ABF	-284.81
Opponent speed (B)	-602.42	ADF	-255.23
Fighter speed (D)	-1375.38	BDF	-203.86
AAM range (F)	-217.86	ABDF	433.81
AB	-146.19	AA	-223.74
AD	-870.21	BB	121.89
AF	119.35	DD	920.57
BD	691.15	FF	138.94
BF	145.79		

Table 3. Best-fit model ASM launch time, Adjusted $R^2 = 0.9982$

Factor	Coefficient
Const	961.00
Opponent speed (B)	-356.25
ASM range (C)	-487.50
BC	168.75
BB	118.75

4.5. Optimisation for definition of performance boundaries

The fighter scenario objective was to deliver an AAM to the opponent before the opponent could launch an ASM. It follows that the fighter could only achieve this objective when:

$$t_{AAM\ arrival} \leq t_{ASM\ launch}$$

All fighter factors were held constant to represent a static level of capability. Three levels of *Opponent speed* were chosen to represent current, near future, and advanced opponents. The quadratic equation, provided by the best-fit models, provided the greatest value of fighter *Distance to Opponent* that would meet the objective for each *Opponent speed*. These results supported boundary definition. Further analysis was done through

deterministic sensitivity analysis (DSA). The DSA results were compared with the equation results and provided data points to support definition of two further performance boundaries for each *Opponent Speed*.

5. RESULTS

The equation prediction was within 2% of two of the three DSA runs but was almost 10% off for the other. This highlighted the need to explore alternative designs and optimisation methods.

The DSA results were extrapolated to define three sets of envelope boundaries, each with three probability-of-success levels. The boundaries define a maximum distance to opponent within which the friendly fighter must be positioned for an opportunity of one, two, or three engagements. Figure 3 shows an example of the resulting performance envelope boundaries for defence against a range of opponent speeds; Low (L), Medium (M), and High (H). It is evident that to allow two or three engagements (the smaller boundaries), the fighter must be closer to the opponent. If the opponent appears inside AAM launch range, the fighter may detect and target the opponent without AEW&C assistance. In this case it may still be effective if found outside the AEW&C radar range. The visible differences provide an example of the insight that could be used to inform system upgrade, refresh, or replacement decisions.

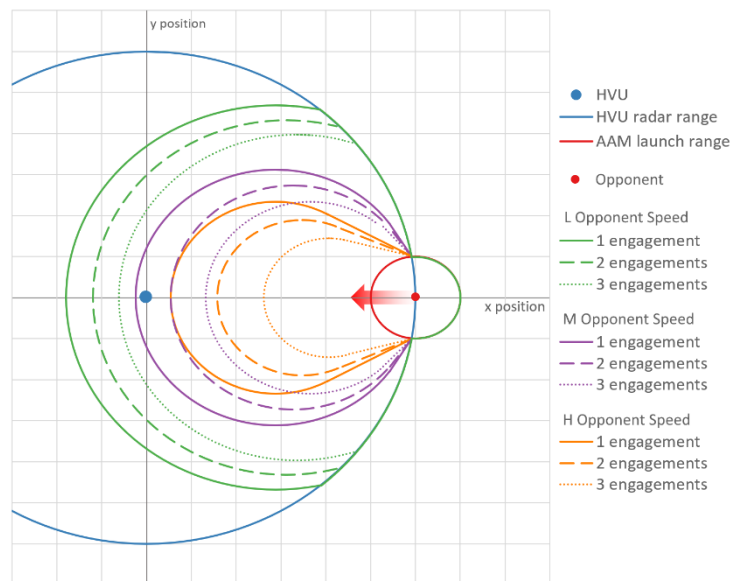


Figure 3. Example fighter performance boundaries against opponents with Low (L), Medium (M), and High (H) speed

6. DISCUSSION

This study investigated the application of DOE to scenario activity flows present in an executable descriptive model, to demonstrate that the concept had merit. Significant factors identified through DOE analysis largely aligned with pre-DOE predictions. Based on these observations, the authors see no obvious constraints in applying similar methods to model scenarios of larger scale and complexity.

The scenario activity flows used for this study were developed similarly to those typically found in Defence capability design descriptive models. The process of updating the model infrastructure to support DOE was reasonably straightforward, and mainly involved adding parameter variables to define performer behaviour and scripts to manage parameter influence on activity duration and flow. The simulation analysis objective will determine the extent of model changes and development required. This study concentrated on event timing and input factor optimisation to produce a desired response. This required focus on a somewhat linear scenario thread through the activity model. Other analysis objectives will require different and tailored approaches.

Developing the analytical side of scenario-based descriptive models to support simulation presents an efficiency opportunity realised by removing the need to duplicate the scenario flow in a separate analysis tool such as MATLAB. This inherently removes the risk of duplication or translation errors. Further, keeping the simulation functionality within a single tool affords the opportunity for feedback loops to be captured and traced explicitly to the scenario and to requirements.

Large-scale full-factorial or Monte Carlo-style simulations were not required to conduct a rigorous statistical analysis of the simulation. DOE delivers a method that achieves satisfactory results with a substantially smaller number of simulations; using a NOLHD provides confidence that the full scope of the experiment space is represented, allowing for a smaller scale full- or fractional-factorial simulation using only significant factors.

Quantification of the extent to which each significant factor affects the simulation responses can provide statistical rigour for capability upgrade and acquisition decisions. For example, the clear indication of reduction of friendly fighter envelope in proportion to increasing opponent speed in the performance boundary results. Significant factors identified through DOE highlight where capability improvements can have the largest impact on scenario outcomes. For example, to have a chance of successfully defending against a future high-

speed opponent, the current fighter platform must be positioned within the defined boundaries. With limitations statistically defined and understood, significant factors can be optimised against a future opponent profile, providing a statistical foundation to inform requirement definition for capability upgrades. Further, solution concepts and options can be explored and simulated in early project stages to examine how well they perform against Program scenarios and Measures of Effectiveness, providing evidence to support final decisions.

6.1. Future development

Future development directions are summarised below:

- **Further examination of DOE design and optimisation methods.** This study was only an initial application of DOE to this kind of problem. There are several other designs and optimisation methods that could be applied. Further work should seek to understand which is best for this kind of problem.
- **Integration with system architecture models.** This work may be extended to test the scalability of this DOE application by increasing the model fidelity through the integration of system architecture models with the operational architecture model. These lower-level models would add further detail and parameters for the simulation and could support test and evaluation of desired configurations.
- **Information flow fidelity.** Information flow timeliness factors did not have a comparatively noticeable effect on the simulation responses. There is a desire to apply DOE to an example that involves greater reliance on complex information flow (larger numbers of factors) with potential for delays and/or congestion in support of analysis and insight of information flow timeliness issues.
- **Extending the scenario.** There is also the opportunity to extend the scenario beyond the bounds of targeting, and to examine a larger picture that includes elements like force composition and structure, or logistics. This would also serve to test the scalability of the presented DOE application across a growing and increasingly complex scenario scope.

7. CONCLUSION

The application of DOE to the analysis of capability scenario simulation in executable operational architecture models is a promising method to increase statistical confidence in the results. This foundation of statistical rigour has potential to improve confidence in the resulting evidence provided to decision makers in the early stages of capability design in Defence Programs and Projects.

ACKNOWLEDGEMENTS

The authors acknowledge DOE support and guidance from Mr Matthew Dragovic and input from reviewers.

REFERENCES

- Air Academy Associates, 2019. *Accelerating the Analysis of Test Data Using Effective and Efficient Experimentation*. [Accessed Online 05 Oct 2019]
Available at: <https://airacad.com/wp-content/uploads/2019/07/TIW-Las-Vegas-Tutorial-2019.pdf>
- Cioppa, T. M. & Lucas, T. W., 2007. Efficient Nearly Orthogonal and Space-Filling Latin Hypercubes. *Technometrics*, 1(49), pp. 45-55.
- Defence, 2015. *First Principles Review - Creating One Defence*, Canberra: Commonwealth of Australia.
- Defence, 2017. *Interim Capability Life Cycle Manual*, Canberra: Department of Defence.
- Hallett, J., Psalios, M., Jusaitis, D. & Cook, S., 2018. *The role of scenarios in model-based management of capability Programs*. Melbourne, Engineers Australia, pp. 183-197.
- Jusaitis, D. & Cook, S., 2018. *A framework to support model-based management of capability Programs*. Melbourne, Engineers Australia, pp. 266-277.
- Montevechi, J. A. B., Miranda, R. d. C. & Friend, J. D., 2012. Sensitivity Analysis in Discrete-Event Simulation Using Design of Experiments. In: E. W. C. Lim, ed. *Discrete Event Simulations - Development and Applications*. Itajubá(MG): IntechOpen, pp. 63-102.
- Santner, T. J., Williams, B. & Notz, W. I., 2018. *The Design and Analysis of Computer Experiments*. 2nd ed. New York(NY): Springer.
- Schmidt, S. R. & Launsby, R. G., 1994. *Understanding Industrial Designed Experiments*. Colorado Springs: Air Academy Press.