

# A Capability Maturity Model for Benchmarking in Wargames

Mehwish Nasim<sup>1,2,4,\*</sup>, Adam J. Wilden<sup>1</sup>, Peter Williams<sup>3</sup>, Timothy Legrand<sup>4</sup> and Patricia A. H. Williams<sup>1</sup>

<sup>1</sup>Flinders University, Adelaide, Australia

<sup>2</sup>The University of Western Australia, Perth, Australia

<sup>3</sup>Defence Science and Technology Group, Adelaide, Australia

<sup>4</sup>The University of Adelaide, Australia

[mehwish.nasim@uwa.edu.au](mailto:mehwish.nasim@uwa.edu.au)

[adam.wilden@flinders.edu.au](mailto:adam.wilden@flinders.edu.au)

[peter.williams2@defence.gov.au](mailto:peter.williams2@defence.gov.au)

[tim.legrand@adelaide.edu.au](mailto:tim.legrand@adelaide.edu.au)

[trish.williams@flinders.edu.au](mailto:trish.williams@flinders.edu.au)

**Abstract:** This research provides an analysis of maturity models, and insights from specific game studies such as unclassified non-kinetic games, supported by contributions from the wargaming community. By proposing a design framework inspired by capability maturity models used in software development, cyber security, and people management, this research introduces a new benchmark for evaluating wargames, in a reproducible and standardised fashion. This model facilitates the identification of strengths and areas for improvement, offering a structured path to higher maturity levels. It aims to enable wargame designers to assess and compare wargame components systematically, enhancing the ability to validate outcomes, predict gameplay effects, and support decision-making with greater confidence. Such advancements could significantly impact policies and improve disaster resilience, particularly within Defence strategy and capabilities, marking a significant advancement in the academic and practical enhancement of the wargaming field.

**Keywords:** Validation; Benchmarking; Reproducibility; Wargaming; Maturity Mode

---

## 1. Introduction

In our previous research (Wilden et al. 2023), a scoping review on wargaming literature identified a notable lack of standardised methodologies for analysis and benchmarking within the field. The review identified various areas of wargaming research including general and conflict modelling, the use of artificial intelligence in game design, and evaluation practices that highlighted the challenge of replicating studies and validating outcomes due to the lack of uniform standards. Additionally, there is even a distinct lack of standardization in the presentation of information in wargaming Wojtowicz (2019). This level of fragmentation across the wargaming research landscape complicates efforts towards achieving standardisation and reproducibility. While it can be suggested that wargames are more art than science, and developed to be bespoke artifacts for single-use scenarios, research does suggest that there is a methodological definition for wargaming, that can be standardised without losing the artistry, and customisation involved (Banks 2024).

This is a work-in-progress paper. We propose a design framework that draws inspiration from Capability Maturity Models - CMMs (Curtis, Hefley & Miller 2009; Paulk 2009; Rea-Guaman et al. 2017; Silva et al. 2015) that have successfully been applied to domains such as software development, cyber security, and people management. CMM are structured frameworks used to assess and improve the processes of an organization. The proposed model can be used as a benchmark for comparison between wargames. It will enable wargame designers to identify strengths and improvement points and provide a pathway to reach higher maturity levels. Our framework can assess the qualitative positioning of the wargame components in various stages or phases. A standard framework would be a step toward allowing the designers to validate, compare, and predict the effects on gameplay and for decision-makers to draw conclusions with more confidence that could impact policies and disaster resilience for the government's strategies and capabilities.

While there has been a considerable focus and output of academic research in the Software Engineering domain, professional wargaming has had less academic focus and a more fragmented and practical focus concerning design and benchmarking. When considering the history of Capability Maturity Models about software, Paulk (2009) suggests that two key themes emerge. First, there is a definite need for best practice frameworks, and second that said frameworks need to be flexible to accommodate a wide variety of different applications. This need arises from organisational requirements of being able to judge cost-effectiveness, and overall efficiency of software projects, while also being able to adapt to changing software requirements and allow for innovation.

In addition to an overall capability measured by the CMM, there is also the need for a focus on how data is dealt with. To this end, there are data governance maturity models; tools and methodologies used to measure the capability of an organization relating to the acquisition, holding, sharing, use, and exploitation of data (Caballero et al. 2023). Like that of the CMM, categorizing an organization as having mature levels of data governance policies involves all processes that are required to manage, access, and innovate using organizational data. The assessment process for data governance involves the analysis of four core domains (Javaid & Iqbal 2017); People, Technology, Data Management and Process aspects. These four domains are also relatable to the wargaming field.

## 2. Methodology

To better develop a CMM for wargaming, that would not only be fit for purpose but also consistent with other CMM frameworks, it was decided to use the Australian Energy Sector Cyber Security Framework (AESCSF) (Australian Energy Market Operator (AEMO) 2024) as an inspiration for structuring the wargaming framework (Figure 1).

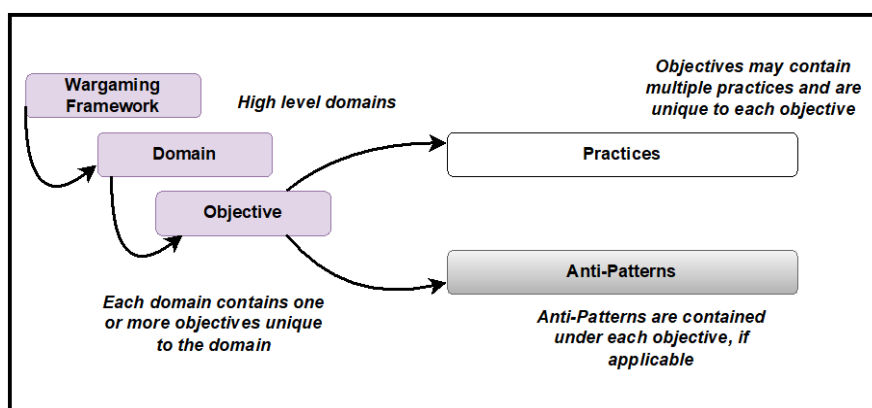


Figure 1 The wargaming framework consisting of domains and objectives

At a lower level, the AESCSF outlines how the practices are benchmarked to judge the maturity of that objective and domain (Figure 2).

As described by the AEMO (2024) in their description of the AESCSF, each of the shown (Figure 2) Practice and Anti-Patterns are assigned a maturity indicator level - MIL (in this case MIL-1, MIL-2 or MIL-3), that describes the level of maturity, and overall implementation of the practice. The AEMO (2024) describe each of these maturity levels and their characteristics that impact the assessment of Practices as follows:

- MIL 1 (Initiated): Initial Practices are performed but may be ad-hoc.
- MIL 2 (Performed): Practices are more complete or advanced than at MIL 1 with the introduction of management characteristics that drive consistency and repeatability.
- MIL 3 (Managed): Practices are more complete or advanced than at MIL 2 with the addition of further management characteristics that drive governance and continuous improvement.

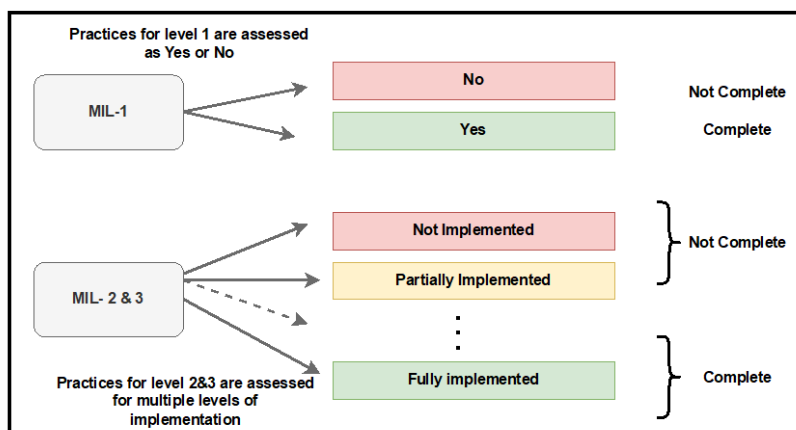


Figure 2 “AESCSF Maturity Levels” AEMO (2024)

While it is not necessary to stick to a three-stage maturity level benchmarking for wargaming (further refinement may necessitate adding levels), this approach can be adopted initially and can be changed without issue later.

The AESCFS scoring model judges the completion of maturity levels using the following guidelines.

- A Practice is “Complete” if it is assessed as “Largely Implemented” or “Fully Implemented”.
- A MIL is “Achieved” if all Practices within it are “Complete”.
- Scored based on a combination of “Practice implementation” and “Management Characteristics”.
- The MILs apply independently to each domain and are cumulative.
- For a participant to gain a MIL in each domain, they must Complete all practices, and not exhibit any Anti-Patterns, at that MIL in that Domain.
- For example, to achieve a MIL-3 the participant would have to perform all Practices and not exhibit any of the Anti-Patterns, in MIL-1, MIL-2, and MIL-3.

These criteria may or may not be appropriate for all wargaming practices, however, it can be initially used as a guide for developing the wargaming CMM. Wargaming maturity levels may end up being considered complete with partially implemented practices, and/or combinations of other complete practices and lack of anti-patterns.

We have modified the AESCFS’s definition of anti-pattern in the context of wargaming, as follows:

- Anti-Patterns are included in the wargame to enable the identification of behaviours/practices that hinder a game from achieving a higher maturity.
- Anti-Patterns are developed in consultation with wargamers and stakeholders.
- In essence, they are ‘bad’ activities that undermine the effectiveness of an element in a wargame. Therefore, additional focus is given to them to encourage wargamers to fix these behaviours.

For wargaming maturity benchmarking, an anti-pattern could be identified through collaboration with wargaming subject matter experts, and from observations of wargaming activities.

### **3. Evaluation**

The evaluation was conducted on three wargames. Eight wargaming designers conducted the evaluation.

Wargame 1: Notes were taken during the execution of this wargame. This wargame involved three teams (Blue, Orange, Green) with unique units and capabilities, spanning motorised platoons, cyber operations, intelligence networks, electronic warfare, commandos, and local security. Each team's units were designed with specific roles, such as reconnaissance, cyber surveillance, human intelligence, and special operations, reflecting a comprehensive simulation of modern warfare and intelligence-gathering activities. Only one analyst evaluated this wargame.

For the other two wargames, the designers provided the feedback post-wargaming exercise.

Wargame 2: It was a kinetic wargame focused on peer enemies engaged in a conflict with long-range fires and a significant air-sea battle. It used a structured, manual wargaming system and was attended by around 20 people (8 players and 12 analysts/support staff).

Wargame 3: It used a bespoke tabletop system to play an influence scenario where 9 different actors (teams of players) were trying to influence a local population and otherwise positively or negatively affect the other players.

Each wargamer filled out an Excel sheet. A snippet of the Excel sheet is shown in Figure 3. The blank template is available on GitHub (Nasim 2024).

In response to the question, “Can this model be useful for standardizing the wargaming-design process for comparison of wargames and their validation?”, all eight wargamers responded Yes.

### **4. Results**

We evaluated the agreement between the wargamers for each wargame. The wargamers/analysts independently filled the last column of the Excel sheet (Figure 3). For brevity, we have summarised the results as follows.

Wargame 1: N/A

Wargame 2: There was 90% agreement on the assessment.

Wargame 3: There was 92% agreement on the assessment.

Domain	Objective #	Objective	Practice #	MIL	Identifier	Practice Text	Wargame
DATA	9	Data capture	1	1	DATA-9.1.1	During wargame planning create detailed data capture plan.	Partially Implemented
DATA	9	Data capture	2	1	DATA-9.2.1	Ensure data capture personnel understand responsibilities and process.	Partially Implemented
DATA	9	Data capture	3	1	DATA-9.3.1	Store backups of wargame data for future analysis.	Partially Implemented
DATA	9	Data capture	1	2	DATA-9.1.2	Data capture plan must explicitly incorporate adjudication rulings and associated decision making process.	Partially Implemented
DATA	9	Data capture	3	2	DATA-9.3.2	Include advance backup plans in case of failure.	Partially Implemented
DATA	9	Data capture	3	3	DATA-9.3.3	Off-site storage and/or multiple levels of backup for redundancy and disaster recovery.	Not Implemented

Figure 3 Snippet of the model for Objective # 9 which lies in the data domain

## 5. Discussion and Future Work

### 5.1 Conclusions

We presented a capability maturity model for benchmarking wargames. We found the designers have a high level of agreement in rating the objectives and practices in two wargames. In future, we will test the framework on different types of wargames and will expand on objectives relevant to different kinds of scenarios.

### Acknowledgment

\*This research was supported by the South Australian Defence Innovation Partnership – Collaborative Research Funds, awarded to Dr Mehwish Nasim (2021-2023).

### References

- Australian Energy Market Operator (AEMO), ACSCA, Cyber and Infrastructure Security Centre (CISC), 2024, *Aescsf Framework and Resources*, Australian Energy Market Operator (AEMO), viewed 25 Mar 2024, <<https://aemo.com.au/en/initiatives/major-programs/cyber-security/aescsf-framework-and-resources>>.
- Banks, DE 2024, 'The Methodological Machinery of Wargaming: A Path toward Discovering Wargaming's Epistemological Foundations', *International Studies Review*, vol. 26, no. 1, p. viae002.
- Caballero, I, Gualo, F, Rodríguez, M & Piattini, M 2023, 'A Maturity Model for Data Governance, Data Quality Management, and Data Management', in *Conference on Cloud Computing, Big Data & Emerging Topics*, pp. 157-70.
- Curtis, B, Hefley, B & Miller, S 2009, 'People Capability Maturity Model (P-Cmm) Version 2.0', *Software Engineering Institute*, p. 18.
- Javaid, MI & Iqbal, MMW 2017, 'A Comprehensive People, Process and Technology (Ppt) Application Model for Information Systems (Is) Risk Management in Small/Medium Enterprises (Sme)', in *2017 International Conference on Communication Technologies (ComTech)*, pp. 78-90.
- Nasim, M 2024, *CMM-wargaming*, GitHub, viewed 03 Apr 2024, <https://github.com/mehwishnasim/CMM-wargaming>.
- Paulk, MC 2009, 'A History of the Capability Maturity Model for Software', *ASQ Software Quality Professional*, vol. 12, no. 1, pp. 5-19.
- Rea-Guaman, AM, San Feliu, T, Calvo-Manzano, JA & Sanchez-Garcia, ID 2017, 'Comparative Study of Cybersecurity Capability Maturity Models', in *Software Process Improvement and Capability Determination: 17th International Conference, SPICE 2017, Palma de Mallorca, Spain, October 4–5, 2017, Proceedings*, pp. 100-13.
- Silva, FS, Soares, FSF, Peres, AL, De Azevedo, IM, Vasconcelos, APL, Kamei, FK & de Lemos Meira, SR 2015, 'Using Cmmi Together with Agile Software Development: A Systematic Review', *Information and Software Technology*, vol. 58, pp. 20-43.
- Wilden, AJ, Nasim, M, Williams, P, Legrand, T, Turnbull, BP & Williams, PA 2023, 'On Benchmarking and Validation in Wargames', in *22nd European Conference on Cyber Warfare and Security, ECCWS 2023*, pp. 533-43.
- Wojtowicz, N 2019, 'From Sandboxes to Laboratories: Evolution of Wargaming into a Method for Experimental Studies', *International Journal of Scientific and Research Publications*, vol. 9, no. 12, p. 399.