

# Establishing Confidence in Combat Simulation Input Data

Denis Shine <sup>a</sup>, Adam Murgatroyd <sup>b</sup>, Richard Dexter <sup>a</sup> and Fred Bowden <sup>a</sup>

<sup>a</sup> *Defence Science and Technology (DST)*

<sup>b</sup> *Defence Science and Technology Laboratory (Dstl)*

Email: [denis.shine@defence.gov.au](mailto:denis.shine@defence.gov.au)

**Abstract:** Combat simulation software routinely requires extensive input data files describing the performance of military systems. Typically these input files are built as lookup tables and can require hundreds of thousands of individual data points describing information such as weapon accuracy and lethality, sensor performance and mobility.

To combat the increasing complexity created by the manual generation of these data, input data generation processes are becoming increasingly automated, both in terms of technical performance models and algorithms to construct simulation databases. This has greatly reduced the time and manpower requirements of generating useful input data, such that databases that once took weeks to construct can be built in days or even hours.

However, such automation has introduced the danger of erroneous data or processes propagating through such an automated system. Depending on the nature of an error, it could potentially affect large portions of the output data. Worse still, automating processes tends to hide intermediate steps where an error may be caught early, with error in the final result being obfuscated.

To combat this we have developed a comprehensive verification and validation approach to our data generation practices. This approach uses a variety of methods to inspect various elements of the system, including source input data, intermediate data, physics-based algorithms and output data.

This paper will briefly present our automated data generation processes, which have also been described in a number of previous publications. It will then highlight, through examples, some of the dangers we have observed through the use of our automated system. Finally it will describe our verification and validation approach to ensure data quality.

**Keywords:** *Validation, verification, simulation, data*

## 1. INTRODUCTION

Within the Defence Science and Technology (DST) group the Land Capability Analysis (LCA) branch relies on the use of combat simulation as part of its analytical capability to support Army Modernisation decision making. Combat simulation allows for the assessment of the operational effectiveness achieved through acquiring new capability, implementing tactical changes or changing force structures. Key elements of a combat simulation include the entity representations (such as a soldier or combat vehicle), their individual and collective behaviours, the terrain and scenarios Scheme of Manoeuvre (SoM). The combat simulation also provides adjudications based on entity interactions, in particular detection and combat effects.

All of these elements require enormous quantities of input data, which is difficult to generate and organise. A decade ago, this process required heavy human intervention, with a simulation analyst collecting data, generating the interaction data for a limited set of entities, building the required databases and verifying the final configuration worked as intended. To alleviate this burden, LCA has developed the Simulation Repository (SimR [Holden et al, 2016]) application, which automates much of the data collation, generation and organization processes. SimR stores basic characteristic and performance data in a simulation-agnostic manner, and uses external tools to generate more detailed performance data without human intervention. It then transforms this data into a specific combat simulation database, one of the key inputs for tools such as COMBAT XXI<sup>1</sup>[US ARMY, 2017], which is LCA's primary simulation tool. SimR has greatly improved the agility of LCA's combat simulation data generation capability, allowing combat simulation databases to be built within hours, rather than days or weeks.

Increasing automation has not been without its problems. By nature, automation tends to hide intermediate steps in a process. Therefore, errors can potentially cascade through the chain of models. Furthermore, the result is often a simulation database, which is difficult to manually inspect efficiently. As a simple example, if one introduces an error in the initial velocity of a ballistic munition, it will impact the trajectory of the munition in the combat simulation. However, this error will also impact on the munition's dispersion, as accuracy has a relation to velocity and will impact on lethality, as penetration is also a function of velocity. If the initial error is not identified at the source, it can be difficult to detect through inspection of the resulting combat simulation database. Given all intermediate processing is currently automated the inspection of intermediate results is not easily achievable.

## 2. LCA SIMULATION VALIDATION AND VERIFICATION (V&V) FRAMEWORK

The LCA framework is based on many elements of traditional V&V [SISO Vol 2(2012)], but expands them for our particular use case. Most traditional simulation V&V revolves around model V&V; ensuring a model performs as intended, and is fit for purpose [Girardot, D et al (2004)]. While this approach has these elements, we are faced with additional challenges, such as, the SimR application, which stores empirical data also requires a process to ensure the validity of its data. In addition, the models within SimR tend to operate as chains – for example, the ballistics model feeds the penetration model, which feeds the vulnerability model, which feeds an exporter, which produces combat simulation data. Not only does each model need to undergo a V&V process, but the intermediate data passed in between each chained model also requires assessment.

To achieve this, LCA has developed a comprehensive framework that is specific to our combat simulation needs. This is based on a mixed methods approach, where multiple independent views are applied to a sub-element of the system. This both allows for a fit-for-purpose method to be applied, and also improves the confidence of the overall V&V process. Broadly, the methods adopted fall into six categories [SISO Vol 2 (2012)], which are described below.

*Comparisons with empirical data.* All performance models built for combat simulations are ultimately built from characteristic data describing military platforms. Therefore, empirical data can be used to validate this data; indeed the empirical data is often the source. In addition, the intermediate models in the data generation process (for example, a ballistics model) also produce their own data, which, if available, can also be validated through empirical data.

*Internal consistency review.* Internal consistency reviews are used when empirical data is not available, either due to a lack of access to sources or due to a lack of characteristic/performance data in the first place. In the absence of such data, the best that can be achieved is to ensure that the data being reviewed makes sense within the confines of the data generation system. An excellent example is fragmentation patterns. While

---

<sup>1</sup> Combined Arms Analysis System Tool for the 21<sup>st</sup> Century

empirical data exists for a wide variety of bursting munitions, some munitions have no available information. Therefore, fragmentation models predict the pattern based on other physical data, and an internal consistency review is used to ensure the resulting pattern aligns with expectations.

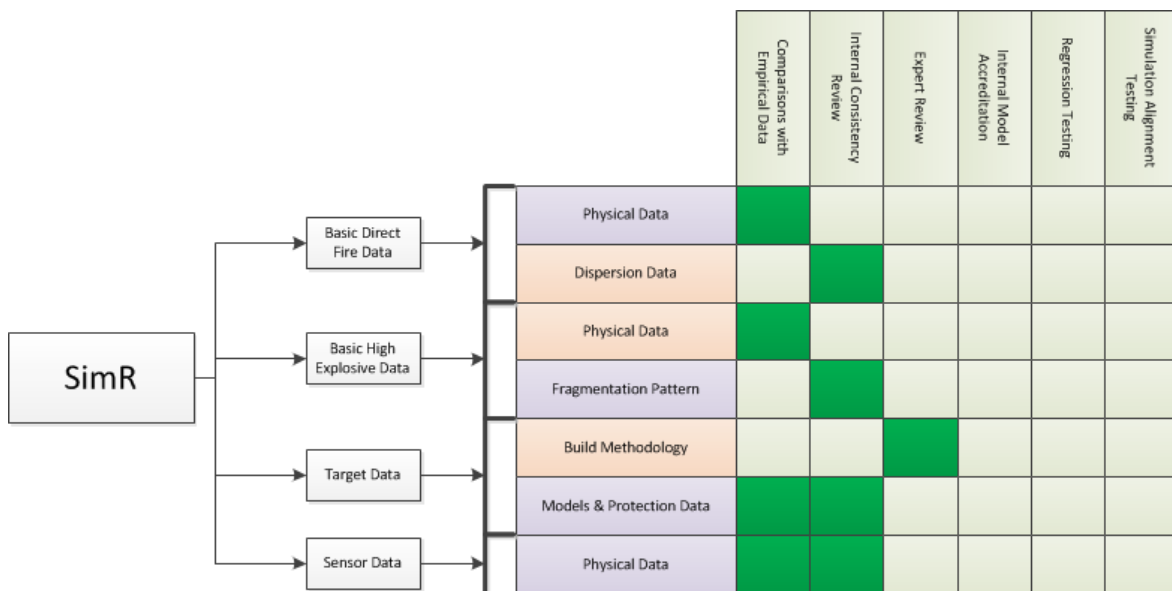
*Expert review.* As LCA is not a recognised authority at technical modelling of physical systems and interactions, expert review is sought in two key areas. First, models developed by LCA need review to ensure they align with best practice and are fit for purpose [Mazonka et al, 2016 (a and b), 2013(a and b), 2012]. Second, the resulting data from each model needs review against higher fidelity models of the same type. As an example, LCA uses a number of simple algorithms to represent ballistic penetration. Results from these can be compared with more complex models, such as MAVKILL and the Unified Weapons Model- Weapon Target Interaction (UWM-WTI) models, [to determine if they align and can be corrected if necessary.

*Internal model accreditation.* The internal models developed by LCA undergo an accreditation process to ensure they behave as intended. In the past this process was relatively ad-hoc, but more recently this process has evolved into a more formal review and documentation approach to unmask potential weaknesses and error sources systematically. As an example, the penetration model used [Mazonka et al (2013b)] is currently being upgraded to include best practice behind amour debris effects. This upgrade has been managed by an accreditation process which involved the production of an extensive set of test data, along with two internal review sessions. This model has now been accepted and will be moved into the production environment.

*Regression testing.* Regression testing forms a key part of the model V&V process and ensures that software updates over time do not unintentionally impact the results generated by the model. This is managed by using reference data as an input and identifying that the output from the model matches an expected set of values and there has been no changes from corrections/updates made within the software.

*Simulation alignment testing.* The purpose of SimR and its sub-models is ultimately to produce input data for combat simulations. Although the key models are all generic, they are connected to exporters which convert this data into the exact format required by the target simulation, taking into account any assumptions. Testing is required to ensure that generic results, as much as possible, align with simulation-specific results.

**Figure 1** illustrate how the framework is applied. It shows how it is applied to the basic data in SimR, such as direct fire, high explosive and sensor data. Given this part of the system contains no actual models, only the first three of the six noted methods are used.



**Figure 1.** V&V Framework - Basic SimR Data

**Figure 2** shows how the framework is applied to SimR’s direct fire sub-model. In this case, data is the result of a performance model, such as a ballistics model. In some cases, this data is comparable to empirical data, but as the model complexity increases or we are investigating future weapon concepts, availability of empirical data decreases. Therefore, for data such as probability of kill, the primary V&V methods lean on consistency checking and expert review.

At this stage, the models used also require scrutiny. This includes internal scrutiny, to ensure the model was implemented correctly and is fit-for-purpose, along with expert review from external partners.

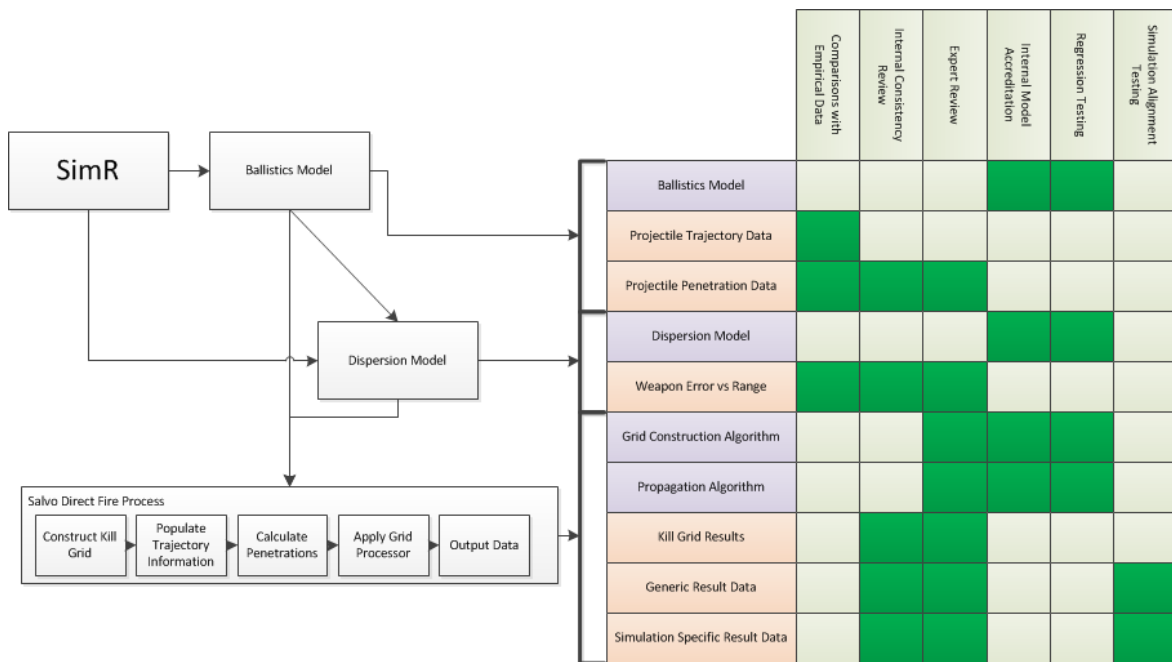


Figure 2. V&V Framework - Ballistics & Dispersion Model

### 3. IMPLEMENTATION & LESSONS LEARNT

In early 2017, the initial implementation of this framework was commenced. The V&V process was implemented in two different ways, depending on where the data has originated.

Initially the focus was on the:

- Munition Characteristic Data
- Ballistic Data
- K Values (used in the development of accuracy errors)
- Target Model Data
- Propagation Data( Automated Process that identifies the penetration performance of a munition)

The first four areas had a focus on manually validating the data within SimR to ensure that the automated follow on processes would have the correct inputs.

For the 5<sup>th</sup> area the V&V of these automated processes to ensure that, whatever inputs are entered, the model in question created the expected outputs.

#### 3.1. Characteristic Validation

The characteristic data within SimR (which included parameters such as munition calibre) were compared against a variety of documents (both internal and open source) in order to ensure that there was an accord between the two, regardless of the level of confidence in the data source. If the data was found to be different between the two, the data within SimR would be updated, with a description within the entity in the database explaining the changes made. It was important during this part of the process to also identify the provenance source and the consequent level of confidence that could be attached to that data.

For example, if the muzzle velocity and ballistic coefficient of a munition and weapon pairing has been delivered from a range trial conducted by DST Group (or a similar trusted organisation), then a high weight of confidence can be attached to this data.

If the data is from open source documentation, then this data can be used within simulations if there are no alternate data sources available, but it should be expressed clearly to the stakeholder that any conclusions

pertaining to this munition have a lower level of confidence attached and are subject to change. It should also be noted that the lower level of confidence in the data may require a sensitivity analysis within a test scenario to identify what level of impact the generated data has on the outcome. If the analysis finds that the data has a significant impact then the sourcing of the data will need to be further addressed to ensure accuracy of results.

This is especially pertinent when there is no data available (this could occur for a number of reasons for example when the weapon or munition in question is still conceptual so the empirical data doesn't yet exist) and has been generated either by surrogation or through subject matter expert opinion, or when the data is within the repository but has no traceable provenance attached to the data field. The latest configuration within SimR allows the user to attach the provenance from various documents to the data fields within each entity contained in the repository, which eases this process for future data audits.

For this type of data it is imperative that a comprehensive list of assumptions are captured and stored in place of a provenance document, so that the data for similar munitions generated in this manner can use the same assumptions and processes. Also, this data is compared to similar entities within the pre-existing validated dataset to ensure that it fitted as expected into the trends existing within the dataset. (As an example if a HMG is more accurate at 5km than a stabilised cannon, then perhaps there is something wrong with the initial data assumptions?)

Ideally, as part of this process, the level of confidence that the provenance source provides should be above some agreed upon level within a rating system. This means that once the data within the repository has been validated against the provenance documentation within the repository, effort should be made to try to source additional information which comes from a more reputable or up to date source, if the weight of confidence in this provenance documentation falls below that required for a study.

As part of the work we developed a validation record for each entity as they were verified so that it is easy for the analyst/SME to identify when this entity data was last checked and validated. At this time SimR allows the user to identify a particular study the data was last used in, which provides the user a limited understanding of when a data audit was last conducted and the confidence level they can have about the data used to support that study.

Finally, if the data in the selected provenance documentation matches that which is in the repository, then no action is required, otherwise the entities within SimR must be updated, including descriptions on why the data is being changed.

### **3.2. Automated Data Validation and Verification**

As identified earlier, some of the data is used within an exported database is generated through the application of a number of automated processes. As part of the generation of data for simulation activities there is a requirement that these undergo the V&V process.

This was implemented in a limited form through the use of a series of unit tests where code responsible for the translation/generation of variables used in the generation of performance data were tested in an automated manner ensuring that only the range of possible values (according to empirical data) are accepted into the model.

These tests must be created for data that is impossible (such as a ballistic coefficient of -1) as well as that which falls on the edge of possibility (like a ballistic coefficient of 0) to identify that these errors are captured and treated correctly during the generation process.

A pass within a unit test indicates that the variables being passed are being handled correctly and that no change to the model should be required whereas a fail results in an investigation into the code for that process (assuming the unit tests have been set up correctly). This is especially important as it will identify if changes to other areas within the code has impacted on the data generation process.

### **3.3. Documenting Validation and Verification**

As part of the V&V process reporting, a number of documents were generated that provided a formal reporting channel for the user. At an early stage it was decided that the reports should be structured in an identical manner to allow for a rapid understanding of the results of the V&V process. To do this the format of the V&V reports are structured in the following way:

**Overview.** Defines the data which is being generated, the challenges around collecting this data, and how it'll be compared against provenance data (when available).

**Process.** Details the actual V&V process undertaken, so that it can be recreated in the future, along with what constitutes a pass or a failure, and the reason for this.

For munition characteristics comparisons, a document will exist for each type of entity available detailing whether or not it passed the validation, and if not why not.

For unit tests, then this will be a separate document detailing each of the tests including; what each test is trying to replicate, the output that is expected, the output that was found, whether or not that constitutes a pass or a fail, and the date on which this test was completed. (Where appropriate this will contain the results, however it is often simpler to do so in separate documents which are weapon/munition specific).

**Outstanding Issues.** Areas found within the data, which couldn't be validated and require additional explanation or confirmation from another analyst/source, before a solution can be identified.

**Recommendations** – Suggests a series of improvements to the current process, which are not essential but would be nice to have, this has included things like adding another step to a part of the V&V process, or altering the GUI of a model slightly to aid the ease of data comparison.

While this process allows us to ensure that the data being used within the simulation environment is traceable and valid for its required purpose, the V&V process itself was found to be very resource intensive. Validation of a single entity could take up to 1 day if the provenance supplied was either lacking or was in great detail. It is anticipated that once we have established a foundation of data that has been through the V&V process this will have a minor impact on the require resourcing but establishing the foundation of data using the existing repository will need to be managed closely.

#### 4. FUTURE

The data generation team interacts heavily with the Close Combat Studies team, who manage and run the COMBAT XXI simulation. Although our data generation capability is designed with multiple target simulations<sup>2</sup> in mind, COMBAT XXI remains the primary customer for the data generation capability. At present, the focus of our V&V capability stops at the point of delivery when the performance database is released to the Close Combat Studies team. In the relative near term, this testing regime will be expanded to include tactical level test scenarios within COMBAT XXI which will be used to benchmark any changes in the overall data generation system, including those to input data or underlying models. The result will be a fully integrated end-to-end V&V capability, from raw input data through to operational results.

This expansion of the V&V architecture will serve a number of purposes. As the data generation system evolves, it will allow identification of any combat simulation results that change as a result, which may warrant further investigation. It will also improve confidence in combat simulation results through the provision of a fully traceable data generation chain. Finally, if a simulation produces erroneous results, it will reduce the number of variables that would need to be considered when attempting to identify the source of the simulation error.

LCA is also exploring the option of automating some aspects of V&V using data analytics techniques. [Potter 2017] showed that machine learning can be used to learn what a performance database should look like and then detect anomalies. For example, if one detects a catastrophic kill (K-Kill) probability in a dataset for a particular weapon/munition combination, one should also expect to see other types of kill, like mobility or firepower, and that if they were not present then an anomaly might exist. This work showed that not only could anomalies like this be detected by such systems, but that the systems could teach themselves the rules of the database composition.

#### 5. CONCLUSION

Through the use of the V&V process LCA has been able to establish a comprehensive V&V framework for its unique combat simulation environment primarily focused on the process for generating input data provided to the simulation user. This framework uses a multi-methods approach, incorporating both data and model V&V throughout the data generation process ensuring that the data produced is both fit for purpose and is as accurate as possible through the use of automated methods employed within SimR. As there is an

---

<sup>2</sup> Other simulations currently being explored are Dstl's WISE simulation as well as Lawrence Livermore National Laboratories Joint Conflict and Tactical Simulation

increased reliance on the automation of performance metric generation which has introduced new challenges this framework seeks to address these issues at an early stage of development. The framework has, so far, been partially implemented with a number of lessons being learned. Through the use of the V&V process we have established a robust reporting method for use within the data generation area and provide a traceable record for any data used or generated during the process. This framework also increases the confidence in the outputs of the target simulation as well as the results of any study that uses the simulation as an analytical tool. Future developments will focus on integrating the V&V framework with the combat simulation process itself and incorporating data analytics as a part of the data generation process

## REFERENCES

- Girardot, D. and Jacquart, R., (2004). *A Proposed Evolution of Validation Definition*. Proc. Foundations of VV&A, Tempe, AZ, USA.
- Hartley III, Dean S. (1997). *Verification and Validation in Military Simulations*, Proceedings of the 1997 Winter Simulation Conference
- Holden L, Dexter R., Shine. D.R (2016). *SimR: Automating Combat Simulation Database Generation*, Presentation at the 24<sup>th</sup> Australian Society for Operations Research (ASOR) National Conference, Canberra
- Mazonka, O. & Shine, D.(2012) *Simple Physical Models in Support of Vulnerability and Lethality Data for Wargaming and Simulation Environments*; Paper presented at the SimTecT Conference 2012, Adelaide
- Mazonka, Oleg (2013a). *Cumulative Probability of Blast Fragmentation Effect*; <http://arxiv.org/abs/1304.2285>
- Mazonka, O. & Shine, D, (2013b). *Methods and Models in Preparing Weapon-Target Interaction Data for Combat Simulations*, Presentation at the 20<sup>th</sup> International Congress on Modelling and Simulation (MODSIM2013), Adelaide;
- Mazonka, Oleg (2016a). *Ballistic Model: Empirical Data to Determine Transonic Drag Coefficient*, The Computational and Mathematics Conference (CTAC), Adelaide
- Mazonka, Oleg and Shine, Denis (2016b). *Blast Lethality Factor of Penetrating Projectiles against Vehicles*; Presentation at the 24<sup>th</sup> Australian Society for Operations Research (ASOR) National Conference, Canberra;
- Potter, T.S., (2017). *An Introduction to Data Science and Analytics with Application Examples from AMSAA*, Presentation at the 85<sup>th</sup> Military Operations Research Symposium, New York
- Simulation Interoperability Standards Organization (2012). *SISO-Guide-001.2-2012 Guide for Generic Methodology for Verification and Validation (GM-VV) to Support Acceptance of Models, Simulations, and Data, GM-VV Volume 2: Implimentation*
- United States Army. (2017). *COMBATXXI Users Guide [Online]*.