

시험평가용 M&S V&V 과정에서의 통계 기법 적용 연구

Study on the Application of Statistical Techniques in the M&S V&V Process for Test & Evaluation

이동준^{*1)}, 김용선¹⁾, 이상복²⁾, 노광현³⁾, 윤주일⁴⁾

Dongjoon Lee^{*1)}, Yongseon Kim¹⁾, Sangbok Lee²⁾, Gwanghyun Roh³⁾, Juil Yoon⁴⁾

[초 록]

무기체계 시험평가지 활용하는 모델링 및 시뮬레이션(M&S)의 검증(Verification) 및 확인(Validation) 과정에서 통계적 기법을 적용하는 방법을 제안한다. 시험평가 과정에서 M&S 활용이 확대되고 있지만, 신뢰성 확보가 필수적이며, 본 연구에서는 다양한 무기체계에 대한 실험적 사례를 통해 V&V 과정에서의 통계기법 적용 방안을 제시한다. 도하 장비(기동 체계), 전술 C4I 체계(지휘통제 체계), 감시 레이더(감사·정찰 체계), 대공 유도 무기(정밀 타격 체계) 등 주요 무기체계를 대상으로 회귀 분석, 카이제곱 검정, 분산 분석, 베이저안 추정 등의 기법을 활용하여 M&S 결과와 실제 데이터 간 일관성을 검증사례를 제시하였다. 연구 결과, 통계적 기법을 적용함으로써 M&S의 신뢰성이 향상될 수 있으며, 제한된 실험 데이터를 보완하는 데 효과적임을 확인하였다.

[ABSTRACT]

This paper proposes a method to apply statistical techniques in the verification and validation process of modeling and simulation (M&S) used in weapon system test and evaluation. Although the use of M&S is expanding in the test and evaluation process, reliability is essential. In this study, we present a method to apply statistical techniques in the V&V process through experimental cases of various weapon systems. We presented the consistency between M&S results and actual data using techniques such as regression analysis, chi-square test, analysis of variance, and Bayesian estimation for major weapon systems such as river crossing equipment (maneuver system), tactical C4I system (command and control system), surveillance radar (surveillance and reconnaissance system), and anti-aircraft guided weapon (precision strike system). The results of the study confirmed that the reliability of M&S can be improved by applying statistical techniques, and that it is effective in supplementing limited experimental data.

Key Words : 모델링 시뮬레이션(Modeling&Simulation), 검증 및 확인(Verification&Validation), 시험평가[Test&Evaluation), 통계 기법(Statistical techniques)

1. 서 론

모델링 및 시뮬레이션(Modeling and Simulation, M&S)은 복잡한 시스템의 성능을 평가하고 다양한 시나리오를 실험할 수 있는 강력한 도구로, 특히 무기체계 시험평가에 적용하는 경우에는 매우 효과적이라 할 것이다. 무기체계의 시험평가에 M&S를 활용하면 실제 장비나 무기체계를 운용하는 과정에서 발생하는 비용과 안전 위험을 줄일 수 있으며, 현실적으로 실험이 어려운 조건에서도 신뢰성 있는 평가가 가능하기 때문이다

1) 한성대학교 국방M&S인정기술연구원(The Defense Institute of Accreditation Technology Research, Hansung University, Korea)

2) 한성대학교 산업시스템 공학부 교수

3) 한성대학교 AI응용학과 교수

4) 한성대학교 기계전자공학부 교수

* Corresponding author,

E-mail:kopoleon@hansung.ac.kr

Copyright © The Korean Institute of Defense Technology

Received : November 26, 2024 Revised : December 7, 2024

Accepted : December 30, 2024

다. 이때, M&S의 신뢰성을 확보하기 위한 검증 및 확인(Verification and Validation, V&V) 과정 또한 필수적이며, 선진국에서는 V&V 활동과 이를 검토하는 활동을 통해 사용의도의 충족여부를 공식적으로 결정하는 인정(Accreditation)절차^[1]를 적용하고 있으며, 국내에서도 검증, 확인 및 인정(VV&A) 절차^[2]를 시행하고 있다. 인정결과로 M&S의 신뢰성 확보하기 위해서는 그 근거가 되는 명확한 V&V 활동이 필요하다.

미 국방부의 V&V 기법 실무가이드북^[3]에서는 75개의 V&V 기법과 특히, 18개의 통계기술을 제시하고 있으며, 국내에서도 이를 참조로 적용하고 있으나 복잡하고 많은 노력이 필요한 통계 기술의 적용은 미흡한 실정이다. 실물시험을 하는 시험평가시 통계적 기법을 적용하는 방법에 대한 연구는 다수 연구되어왔으나 M&S를 활용한 결과값 분석이나 V&V의 기술로 적용하는 구체적인 방법에 대한 연구는 많지 않은 편이다.

본 연구의 목적은 시험 평가 과정에서 실물 시험이 어려운 항목에 활용하는 M&S 체계와 시뮬레이션 결과값의 신뢰성을 확보하기 위한 V&V 활동에서 통계기법을 적용하는 방법론을 제시하는 데 있다.

본 연구의 실효성을 위하여 기동 체계(도하 장비), 지휘통제 체계(전술 C4I 체계), 감시 및 정찰 체계(감시 레이다), 정밀 타격 체계(대공 유도 무기) 등을 예시로 제시하였다. 다만, 통계적 기법은 해당 무기체계의 특성보다는 검증하고자 하는 통계량의 특성이 중요하므로 해당 무기체계로 제한하지 않고 시험항목의 특성에 따라 통계적 기술은 적용되어야 할 것이다.

2 국내 M&S V&V 관련연구

2.1 실제 시험평가에서의 통계적 방법 적용

비용 및 기간 등의 제한에 의해 발생하는 실제시험의 횟수 제한을 극복하기 위해 통계적 방법을 적용해 시험횟수를 결정하고 있다. 시험횟수에 대한 주요 연구로는, 실제시험 수량을 결정할 때 통계적 관점에서 적정 신뢰수준으로 고려한 이항확률분포를 적용하는 것이 합리적이라는 연구^[4], 시험수량을 결정할 때 신뢰수준을 결정하고 요구명중률과 표본오차를 고려한 해야 한다는 연구^[5], 실제시험 횟수 산정에 적용할 적정 신뢰수준을 산정하기 위한 미군 및 한국군 등의 기준을 비교한 결과 60%의 신뢰수준 적용이 적절하다는 판단을 제시한 연구^[6], 통계적인 방법을 고려해 합격률을 결정하고 적정 시험횟수를 결정할 때 유의수준(허용오차), 검정력, 표본오차, 효과크기(유효크기)를 고려해야 한다는 연구^[7] 등이 있다. 위 연구들의 공통점은 시험평가 횟수를 결정하기 전 명중률 또는 합격률을 설정해야 한다는 것으로 시험평가의 결과를 판정할 때에 중요한 요소로 작용한다.

2.2 M&S 시험평가에서의 통계적 방법 적용

설계와 제작을 위한 과학적 분석 수단으로 활용되는 M&S는 확보 가능한 실제 또는 유사 데이터의 활용을 통한 검증 및 확인으로 신뢰성을 확보하며, 다양한 분석과 추정을 하고 있다. 그러나 무기체계의 경우 실제 시험이 불가능한 경우에

M&S의 활용을 권장하지만 이를 검증 및 확인하는 실제 데이터를 충분히 확보하기가 어렵다. 비용 및 기간 등의 제한에서 실제 시험보다는 자유로운 방법인 M&S 시험평가에 대한 통계적 방법의 주요 연구는 신뢰성 보장에 중점을 두고 있다. AWAM(Army Weapon effectiveness Analysis Model, 지상 무기효과분석모델)의 최소 반복 수행 횟수는 신뢰성 있는 결과를 획득하기 위해서 1500회 이상의 시뮬레이션을 반복수행해야 한다고 실험을 통해 제시했다^[8].

2.3 시험평가 신뢰도 향상을 위한 유형별 적용

기존의 통계적으로 신뢰할 수 있는 적정 시험 횟수 판단에 대하여 최근 연구에서 유형별로 횟수를 제한한 연구가 있는데, 시험평가 유형을 세 가지(평균과 기준 비교, 성공률 검증, 3개 이상 집단의 평균 비교)로 구분하여 평균과 기준을 비교시에는 Wilcoxon 부호순위 검정의 횟수, 성공률 검증시에는 정확 방법(Exact method)의 횟수, 3개 이상 집단의 평균 비교시에는 분산분석의 가정을 만족할 시에는 ANOVA(Analysis of variance, 분산분석)를 적용하고 그 이외의 경우에는 Kruskal-Wallis 방법이 적절하다고 제시했다^[9].

2.4 시험평가 결과 판단에서의 통계적 방법 적용

시험평가 결과 판정시 정량적인 기준과 정성적인 기준으로 인한 일부 불협화음이 존재한다. 예를 들어, 정량적인 기준은 '최대사거리 40km 이상인 요구사항에 대해 39.97km를 달성하여 불합격'인 상황을 들 수 있고, 정성적인 기준은 '군 운용 적합성의 용이성과 적절성'이 사람에 따라 다른 기준일 수 있다는 것이다. 이러한 상황을 개선하기 위해 사업 초기에 '개발 목표에 대한 허용오차'와 '정성적 평가 항목에 대한 수치화 설정'을 제시했다^[10].

3. 무기 체계 시험 평가에서 M&S 활용

3.1 M&S 활용의 장점 및 기대효과

M&S는 무기체계의 동작과 성능을 가상 환경에서 모의 실험함으로써 실제 시험의 복잡성을 줄이고, 다양한 환경과 시나리오에서 시스템을 평가할 수 있는 기회를 제공한다. 예를 들어, 다양한 전장 환경에서의 성능을 평가하거나, 극한의 온도, 압력, 기후 조건에서의 반응을 측정하는 등 실물 시험으로는 어려운 상황을 시뮬레이션을 통해 재현함으로써, 무기 체계의 설계 및 제작, 운용의 검증을 가능하게 하는 것이다.

또한, M&S는 실물 시험평가 전에 무기 체계의 설계 단계를 평가할 수 있도록 하여, 초기 오류 발견 및 수정에 기여한다. 이는 개발 과정에서의 리스크를 줄이고, 개발 비용과 시간을 크게 절감할 수 있는 장점을 제공한다. 특히, 이러한 설계 및 제작과정에서 최근에는 디지털트윈의 적용이 확대되면서 재할용 관점을 포함하여 궁극적으로 M&S는 무기 체계의 효율적인 시험과 평가를 통해 체계의 신뢰성을 높이고, 더 나아가 국방 예산 절감에도 기여할 수 있다.

3.2 시험평가 환경과 M&S체계 연동요소

시험평가용 M&S 체계는 실제 시험이 제한되는 환경에서 사

용하게 되는데 시험평가 대상 무기체계와 상호작용하는 요소는 자연환경(지형 및 기상), 무기체계의 표적 및 대응체계, 인접체계라 할 수 있다. M&S를 활용하여 시험평가하는 것은 이러한 상호작용 체계 중 M&S의 범위와 수준을 어디까지 하는냐하는데서 출발한다.

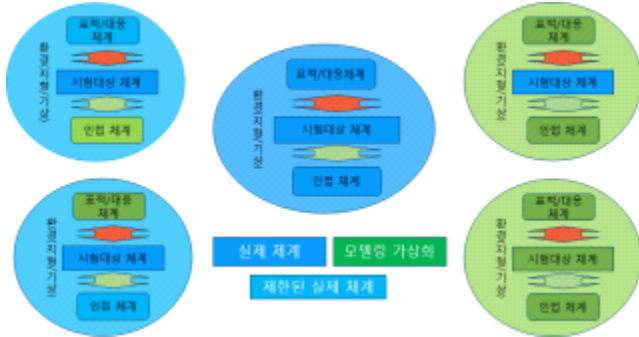


그림 1. 시험평가용 M&S 체계의 활용

Fig 1. Utilization of M&S system for Test & Evaluation

시험대상체계를 포함하여 그림 1과 같이 관련 체계를 모두 실제체계와 환경에서 하는 것이 이상적이거나 이는 불가능하므로 실물시험이라 하더라도 가용하는 범위의 지형 및 기상 환경, 제한되는 표적 및 대응체계, 제한되는 인접체계를 사용하고 있다. 따라서 M&S로 모의하는 지형 및 기상환경, 표적 및 대응체계의 모의기, 인접체계 모의기로 대체하며 시험평가하는 것이며, 시뮬레이션된 극한 환경에서의 해당 무기체계 성능 평가를 위해서는 해당무기체계도 모델링을 해야하는 경우도 있다. 이러한 관련체계의 특성은 실제 데이터를 확보하기 어려우므로 시험평가용 M&S의 활용에 제한이 발생한다. 무기체계의 설계 또는 개발 및 제작 과정에서 해당 데이터를 수집하거나 해당 데이터를 기반으로 M&S 환경, 또는 디지털트윈 환경을 구축한다. 그러나 시험대상체계의 시험환경 모의는 실제체계와 실제 환경에서의 획득가능한 데이터가 충분하지 못하다. 왜냐하면 실제에서 실행할 수 없어서 M&S로 하기 때문이다.

3.3 시험평가용 M&S의 검증 및 확인 대상

시험평가용 M&S의 경우 대부분은 동일한 종류의 체계 개발 및 시험에 활용한 M&S를 재활용하는 경우이다. 다양한 유도탄, 다양한 레이더, 다양한 기동장비 등은 무기체계가 신규개발이라 하더라도 M&S의 근본 특성은 동일하기 때문이다. 또한 부체계, 구성품 단위가 M&S 결과값을 입력값으로 활용해야하는데 이런 경우 이들 하부체계는 상용화된 M&S 제품을 사용하는 것도 일반적이다. 신기술에 의한 첨단무기의 경우 참조모델이 없다면 신규 개발을 할 것이다.

따라서 시험평가 M&S의 신뢰성 확보대상에 대한 명확한 식별이 필요하다. 보다 단순화하여 신규개발, 재활용, 상용도구 활용의 3개의 유형으로 고려할 수 있다.

M&S 체계의 신뢰성 확보를 위한 인정은 실물 시험이 가능한 상황에서 획득한 실제 데이터로 M&S 체계가 동일 시나리오로 했을 때 수락가능한 수준임을 인정하는 것과 M&S체계로 시험평가 항목에 대한 결과를 추정하는 것에 대한 인정이라고

할 수 있다. 즉, 시험평가용 M&S의 신뢰성 확보의 핵심 요소는 사용할 M&S 체계의 검증과 확인, 신뢰성이 확보된 M&S 체계를 활용하여 시뮬레이션한 결과를 시험평가결과로 사용하기 위한 검증과 확인이 핵심이다.

3.3.1 시험평가용 M&S 신규 개발

신규개발이란 새로운 무기체계의 개발로 재활용할 수 있는 M&S가 없을 때, 시험평가항목이 기존의 모델로는 평가가 불가능한 경우 또는 상용도구로 활용이 어려운 경우 등이다. 신규개발하는 M&S의 경우 M&S 개발 절차와 V&V 절차를 기본 절차에 따라 진행한다. 실물시험, 실물시험과 동일한 M&S, 시험평가항목과 유사한 참조데이터, 추정하는 실제작전 데이터가 필요하다.



그림 2 신규개발시 M&S 활용(예)

Figure 2. Using M&S in new development(example)

일반적으로 국방 M&S의 VV&A 절차에서 검증 활동은 요구사항, 설계, 구현, 데이터에 대하여 순차적으로 진행한다^[2] 이때 구현검증은 SW/HW의 단위시험과 통합시험으로 진행되고 있다. 그러나 시험평가용 M&S의 경우 실물시험이 제한되는 시험평가 환경으로 사용되므로 그림 2에서와 같이 구현 및 데이터 검증은 실물 시험평가와 같은 시나리오로 신규 개발된 M&S를 적용하여 그 결과값을 활용하여 비교함으로써 최종 검증을 해야 한다.

또한 확인 활동은 개념모델 확인, 데이터확인, 결과확인이다. 시험평가 대상 무기체계로는 실제값 확보가 불가능한 시점이므로 실제와 일치여부를 비교하는 확인은 불가능하다. 따라서 그림 2에서와 같이 확인은 현재 실제 운용 중인 유사 데이터와의 일치 정도로 결과 확인을 해야 한다.

이러한 비교활동이 V&V 활동이며 일치성, 유사성을 추론하는 통계기법 적용이 필요하다.

3.3.2 시험평가용 M&S 기존 모델 재활용

기존 모델의 재활용은 가장 많이 적용되는 경우이다. 국내 무기체계 개발에 대한 경험이 오랜 기간 축적됨에 따라 실험실 및 시험장에서 획득한 개발 및 시험 데이터를 확보하고 있다. 예를 들어 유도탄 유사 체계인 경우 대체로 환경은 유사하면서 탐색기, 제어, 발사대, 표적의 성질 등에 따라 재활용이 가능하다. 또한 레이더 유사 체계인 경우 역시 표적의 성질, 지상 고정 이나 탑재 이동여부, 탐지거리 등의 차이에 따라 재활용이 가능하다.

따라서 그림 2와 같이 신규 개발 M&S와 같은 V&V 활동이 이루어지지만 그림 3과 같이 기존 모델 재활용시의 검증활동



그림 3 기존 모델 재활용(예)

Figure 3 Reusing legacy models (example)

은 기사용 M&S 체계 운용시와 다른 점에 대한 요구사항이 가장 중요한 부분으로 이후 설계 및 입력데이터 검증은 구현검증과 통합하여 시뮬레이션 작동을 하면서 검증한다. 확인활동은 요구사항 재정의에서 식별된 부분에 대한 모델의 변경에 대한 실제와의 확인이 이루어지면 결과와 출력 데이터의 확인은 유사 체계의 유사 시험평가 항목과의 일치여부로 진행된다. 통계 기법은 이러한 측면에서 일치여부를 추정하는 기법으로 활용되어야 한다.

3.3.3 시험평가용 M&S 상용도구 활용

상용도구의 활용은 대체로 부체계 및 구성품 단위의 해석용 시뮬레이션 도구 들이며 해외에서 운용수준의 시험이 가능한 도구들도 개발 활용되고 있어 이를 구매하여 활용하는 경우가 빈번하다.

다만, 상용도구를 활용하는 경우에는 그림 2의 신규개발이나 그림 3의 재활용과 비교해 볼 때 검증과 확인활동의 중점과 방향은 동일하지만 몇가지 차이점이 있다. 첫째는 도구의 안정성이다. 신규 개발시 가장 어려운 점인 도구의 안정성면에서 상대적으로 우수하다. 많은 참조 사례가 있고 제작사의 운영유지 차원에서 신뢰성은 확보하였다고 간주되는 것이다. 두 번째로 개발과정의 V&V가 단순하다는 점이다. 요구사항에 대한 검증은 모두 동일하지만 개념모델 확인은 해당 도구의 내부 모델이 요구사항에 부합하는 모델인지를 확인하는 것다. 세 번째는 상용도구가 시뮬레이션 결과로 산출하는 결과값이 실물 시험과 동일한 시나리오로 출력되는지 확인이 필요하고 부족한 경우 지속적인 민감도 분석을 통해 과적합 없이 조정되어야 한다.



그림 4 상용도구 활용(예)

Figure 4. Using commercial tools (example)

따라서 그림 4에서의 검증 활동은 요구사항 및 설계 검증은 요구사항대로 입력 가능한지, 출력값은 실물시험 데이터와 일치하는지까지 검증하는 것이며, 확인은 개념모델이 원하는 내용과 일치하는지, 결과값은 상용도구를 활용한 유사 체계에 대

해 축적된 유사 결과값을 기준으로 추정하는 내용이 적합한지를 통계 기법으로 증명하는 것이다.

4. V&V 과정에서의 통계 기법

4.1 V&V의 주요 기법

미국 국방부 실무 참고 가이드북^[3]에서 제시하는 V&V의 주요 기법은 비정형(Informal), 정형(Formal), 정적(Static), 동적(Dynamic) 기법 등 4가지 범주로 분류하고 있다.

비정형 V&V 기법은 도구와 접근 방식이 엄격한 수학적 형식보다는 인간의 추론과 주관성에 크게 의존하는 기법이다.

정형적 V&V 기법은 형식을 갖춘 수학적 증명으로 정확성을 확보하는 방법으로 상당한 노력이 필요하고 간단한 방법으로는 처리할 수 없는 복잡한 문제의 적용에는 필수적이다.

정적 V&V 기법은 정적 모델 설계 및 소스 코드의 정확도를 평가하는 것으로 시뮬레이션 모델의 직접 실행을 하지 않지만 전용 프로그램 자동화 도구를 사용할 수 있다. 일반적으로 정적 기법은 모델의 구조, 사용된 모델링 기법, 모델 내의 데이터 및 제어 흐름, 구문 정확도에 대한 다양한 정보를 확인하는 방법이다.

동적 V&V 기법은 모델 실행 동작에 따라 모델을 평가한다. 입출력이 가능한 모델을 실행하여 모델 출력을 분석하고 동적 모델 동작을 평가하는 것이다. 운용 시나리오에 의한 실행이전에 필요한 기본적인 시험이다.

시험평가용 M&S의 검증 및 확인 대상을 고려할 때 검증에는 정적 기법 중 추적성 검토(Review), 확인은 비정형 기법 중 관련 전문가에 의한 확인(Face-Validation)이 일반적으로 활용되고 있다. 시뮬레이션 결과 확인은 정형적 기법과 동적 기법을 병행하여 다양한 통계적 기법이 사용되어야 한다.

4.2. M&S를 활용하는 시험평가에서의 V&V

국내 무기체계 획득간 실제 시험 평가가 제한된 여건과 이 때 M&S 결과로 시험평가하는 방법을 정리하면 다음과 같다.

첫째 최고 성능, 극한 조건 등에 대한 시험이 필요하나, 제한된 실제 시험만 가능한 경우이다. 무기 체계 개발 중 실제 시험 평가가 제한된 경우, 예를 들어 100km/s 속도로 시험할 수 없고 70km/s 속도도만 가능한 경우, 해당 실제 시험 데이터를 활용하여 M&S 모델을 검증하고 이를 바탕으로 100km/s의 성능을 추정하는 방법으로 통계적 기법 적용이 가능하다. 이 방법은 실제 데이터를 사용해 M&S의 신뢰성을 높이고, 이를 통해 추가적인 성능 추정을 가능하게 한다.

둘째, 충분한 시험 횟수, 4계절 시험 등의 반복이 필요한 경우로 예를 들면 최소 30회 이상의 시험 평가 결과 추정이 필요하나 일반적으로 무기 체계 시험에서는 물리적 시험이 불가능한 경우 M&S를 통해 30회 이상 시험한 결과로 시험평가결과를 추정할 수 있을 것이다. 이는 제한된 실제 시험 결과와 시뮬레이션 결과를 결합한 통계적 기법의 적용으로 신뢰도를 보다 향상시키게 될 것이다.

4.3 통계 기법 적용 분석

회귀 분석은 외삽법이 필요한 체계에서 제한된 실제 데이터를 기반으로 성능을 추정하는 데 매우 유용한 기법이다. 제한된 실험 데이터를 사용하여 다양한 조건에서 성능을 추정해야 하는 경우 회귀 분석을 통해 획득된 실제 데이터 기반으로 추정된 결과 간의 관계가 통계적으로 유의미하지 않는다면, 이를 통해 M&S가 실제 성능을 신뢰성 있게 반영하고 있음을 확인할 수 있다.

회귀 분석의 장점은 연속형 변수를 다루는 데 있어 추정 능력이 뛰어나며, 다양한 시나리오에서 성능을 확장하여 평가할 수 있다는 점이다. 특히 실제 시험이 비용이나 시간의 제약으로 인해 제한될 경우, 회귀 분석은 추정의 정확도를 높이는 도구로 효과적으로 활용될 수 있다. 다만, 충분한 양질의 데이터를 확보할수록 추정의 신뢰성이 더욱 높아진다.

카이제곱 검정은 범주형 데이터 간의 일관성을 평가하는 데 적합한 통계 기법으로, 성공 실패를 평가하는 경우 M&S 결과와 일부 실제 데이터를 처리하여 사용할 수 있다. 이 기법은 M&S가 추정한 성공률과 실제 성공률 간의 차이가 통계적으로 유의미한지를 분석하여, 체계의 성능을 검증하는 기법으로 활용할 수 있다. 통계적으로 유의미하지 않음이 확인되면 이는 M&S가 실제 성능을 정확하게 반영하고 있음을 검증한다.

카이제곱 검정의 장점은 이진 데이터나 범주형 데이터를 효과적으로 다룰 수 있다는 점이며, 무기체계의 성공이나 실패 데이터를 평가하는 데 매우 적합하다. 다만, 이 기법은 연속형 데이터에는 적용할 수 없으며, 데이터가 충분히 확보되지 않은 경우 통계적 검증의 한계가 있을 수 있다.

분산 분석(ANOVA)은 세 그룹 이상의 평균 차이를 비교하는 데 유용한 통계 기법으로, 무기 체계의 다양한 거리 및 기상 조건에서 성능을 평가하는 데 적용할 수 있다. 이 기법은 검증된 M&S로 다양한 조건에 대한 추정 결과를 제시하고 다양한 조건에서 분석한 결과, 특정 조건의 차이를 볼 수 있다면 전반적으로 M&S가 실제 성능을 잘 반영하고 있음을 검증할 수 있다.

분산 분석의 강점은 여러 그룹 간의 평균 차이를 효과적으로 비교할 수 있다는 점이다. 이는 다양한 환경 조건에서 성능 평가가 필요한 경우 매우 유용하다. 그러나 분산 분석은 그룹 간의 분산이 클 경우 결과 해석이 어려워질 수 있으며, 데이터가 균등하게 분포되어 있지 않은 경우 통계적 정확도가 떨어질 수 있다는 한계가 있다.

베이지안 추정은 기존 데이터와 새로운 데이터를 결합하여 추정 성능을 강화하는 통계 기법으로 활용할 수 있다. 예를 들어 10회의 실제 사격 시험 데이터를 바탕으로 500회 M&S 추정 데이터를 보완하여 타격 정확성을 더욱 신뢰성 있게 추정한다. 베이지안 추정의 주요 장점은 사전 확률과 새로운 데이터를 결합하여 불확실성을 줄이고 추정의 신뢰도를 높일 수 있다는 것이다. 이는 특히 제한된 실제 데이터를 보완하는 데 매우 유용하며, M&S의 성능을 강화하는 데 효과적이다. 그러나 베이지안 추정은 사전 확률 설정에 오류가 있는 경우, 초기 가정이 결과에 미치는 영향을 신중히 고려해야 한다는 한계가 있다.

5. 무기체계별 적용 예시

본 장에서는 3.4장에서의 이론적이고 개념적으로 정립한 통계 기법 적용에 대해 보다 명확한 이해를 돕기 위하여 예시를 제시하는 것으로 실제 시행한 결과가 아니다. 제안한 방법론 적용을 증명하기 위해서는 실제 시행 결과를 제시해야 하나 통계적 기법의 타당성은 증명하지 않아도 되는 방법이며, 다만 현재 V&V 활동에서 대부분 경험적 기법, 전문가에 의한 확인 등으로만 시행되고 있다는 점에서 통계적 기법을 활용하는 것을 제안하는 것이다.

예시에 사용되는 무기체계와 이의 성능 항목과 수치는 모두 가상의 내용이며, 향후 본 연구의 결과를 활용할 때도 해당 무기체계보다는 성능 항목과 방법을 적용하는 여건과 기대하는 결과에 대한 예시로 활용하는 것이 적절하다.

5.1 기동 체계(도하 장비)

제한된 실제 도하 시험 데이터를 활용하여, 다양한 환경에서 도하장비의 성능을 추정하는 M&S에 대해 회귀 분석을 통해 다양한 지형에서 성능을 추정하고, 신뢰 구간을 산출한다

도하 성능 데이터를 바탕으로 고유속 환경에서 이동 성능을 추정하는 데 사용한다. 특히, 제한된 실제 시험 데이터를 활용하여 다양한 유속 조건에서 얼마나 효율적으로 도하할 수 있는지를 추정하고자 한다. 이 방법은 실제 시험이 불가능하거나 위험한 상황에서 M&S를 통해 비교적 정확한 성능 추정을 가능하게 한다.

표 1과 같이 회귀 분석은 도하 장비의 성능을 추정하기 위해 변수 간의 관계를 모델링하여 한 변수가 다른 변수에 미치는 영향을 분석하기 위해 적용한다. 도하 장비의 경우, 도하 성능(시간, 속도 등)을 설명 변수로 사용하고, 고유속 조건에서의 성능을 종속 변수로 하여 회귀 모델을 구축한다. 이를 통해 M&S를 통해 추정된 성능이 실제 데이터와 얼마나 일치하는지를 분석하고, 추가적인 유속 조건에서 성능을 추정할 수 있다.

표 1 기동 체계 통계 기법 적용

Table 1. Procedure for applying statistical techniques to the maneuver system

데이터 수집	실제 도하 성능 데이터를 바탕으로 회귀 분석을 수행한다. 유속이 다른 여러 환경에서 수집된 도하 시간 및 속도 데이터를 기반으로 추정 모델을 구성한다.
모델 구축	고유속 환경에서 도하 성능을 추정하기 위해 회귀 방정식을 생성한다. 예를 들어, 유속(X)과 도하 시간(Y)의 관계를 분석하여, X의 변화에 따라 Y가 어떻게 변화하는지 추정하는 모델을 만든다.
추정	고유속 환경에서 실험이 제한될 경우, 회귀 모델을 활용하여 추정된 도하 성능을 평가하고, 그 결과의 신뢰 구간을 계산하여 모델의 추정력을 보완한다.

실제 도하 성능 데이터를 통해 유속 v_1 서 평균 도하 시간인 t_1 , 유속 v_2 에서 t_2 라고 측정되었다. 이를 바탕으로 회귀 모델

을 구축한 후, 유속이 v_3 일 때 도하 시간이 t_3 일 것으로 추정할 수 있다. 회귀 분석을 통해 다양한 유속에서의 성능을 추정하고, 이를 통해 실제 시험을 대체하거나 보완하는 결과를 도출할 수 있다.

그리고 회귀 분석을 통해 추정된 결과의 신뢰 구간을 산출하여, 추정된 성능이 실제로 어느 범위 내에서 일어날지를 평가한다. 예를 들어, 유속 v_3 에서 도하 시간이 t_3 로 추정되었다면, 신뢰 구간을 계산하여 $\pm t_4$ 내에서 실제 도하 성능이 발생할 가능성을 평가할 수 있다. 이는 추정의 불확실성을 최소화하고, M&S의 신뢰성을 높이는 데 기여한다.

5.2 지휘통제 체계(전술 C4I 체계)

제한된 실제 송수신 데이터를 활용하여 전술 C4I 체계 (지휘 통제 체계)의 성능을 추정하는 M&S에 대해서는 실제 데이터와의 차이를 분석하여 카이제곱 검정(Chi-Square Test)기법을 사용하여 M&S의 신뢰성을 확보한다.

전술 C4I 체계에서 M&S를 통해 추정된 통신 성공률과 실제 네트워크 데이터의 통신 성공률 간의 일관성을 평가하는 것이 목적이다. 이를 통해 M&S가 실제 네트워크 통신 성공률을 얼마나 정확하게 반영하는지 검증하고, 두 데이터 간의 차이가 통계적으로 유의미한지를 확인한다.

표 2와 같이 카이제곱 검정(Chi-Square Test)은 두 범주형 변수 간의 분포 차이를 평가하는 데 사용되는 통계 기법이다. 전술 C4I 체계의 M&S 결과와 실제 데이터에서의 통신 성공 및 실패 횟수를 비교하여, 두 그룹 간의 성공률 분포가 유의미하게 다른지를 검증할 수 있다. 이 기법은 두 집단의 성공/실패 데이터 빈도를 비교함으로써, M&S가 실제 성능을 얼마나 잘 반영하는지 평가하는 데 적합하다.

표 2 지휘통제체계 통계 기법 적용

Table 2. Procedure for applying statistical techniques to the command & control system

데이터 수집	실제 전술 네트워크에서의 통신 성공률과 M&S로 추정된 성공률 데이터를 수집한다. 예를 들어, 여러 통신 시나리오에서의 성공 횟수와 실패 횟수를 기록한다.
모 델 비교	카이제곱 검정을 사용하여 M&S에서 추정된 성공률과 실제 네트워크에서 관측된 성공률을 비교한다. 이 과정에서 성공/실패 데이터를 범주화하여 두 집단 간의 빈도를 비교한다.
유의미성 평가	p값을 통해 M&S 결과와 실제 데이터 간 차이가 통계적으로 유의미한지 평가한다. 만약 p값이 0.05보다 크면 두 집단 간 차이는 유의미하지 않다고 판단할 수 있다.

M&S에서 추정된 통신 성공률이 A%이고, 실제 데이터에서 측정된 성공률이 B%인 경우, 평균 성공률로 단순히 비교하는 것이 아니라 두 그룹 간 성공/실패 데이터를 비교하여 차이가 유의미한지를 검토한다. M&S에서 A1회의 성공과 A2회의 실패가 도출되었고, 실제 데이터에서 B1회의 성공과 B2회의 실패가 측정되었다면, 카이제곱 검정을 통해 두 데이터의 차이가

유의미한지를 확인할 수 있다.

유의수준 0.05에서 p값이 0.05보다 크다면, M&S 결과와 실제 데이터 간의 차이는 통계적으로 유의하지 않으며, 이는 M&S가 실제 네트워크 성능을 잘 반영하고 있음을 의미한다. 반대로 p값이 0.05보다 작으면 반영하지 못한다는 결론을 도출할 수 있다.

5.3 감시정찰 체계(레이더)

레이더 체계의 M&S를 통해 다양한 거리와 기상 조건에서 다수의 표적을 탐지하는 성능을 추정하기 위해 분산 분석(ANOVA, Analysis of Variance)을 통해 M&S 신뢰성을 평가할 수 있다.

레이더 체계에서 실제의 제한된 거리에서 표적을 탐지한 데이터를 바탕으로 검증된 M&S를 통해 다수 표적을 탐지하는 성능을 추정하는 것이다. 이를 통해 다수 표적 탐지 결과와 다양한 거리 탐지의 일관성을 확보하고, 다수 표적 탐지 M&S 신뢰성을 확인한다.

표 3과 같이 분산 분석(ANOVA)은 세 그룹 이상의 평균 차이를 분석하는 통계 기법으로, 레이더 체계의 탐지 성능을 거리, 기상 조건 등 여러 요인에 따라 비교하는 데 적합하다. 다양한 거리 표적을 탐지한 데이터를 기반으로 다수 표적 탐지 상황에서 M&S 결과를 추정하고, 그 차이를 분석하여 M&S가 실제 성능을 일관되게 반영하는지 확인할 수 있다.

표 3 감시정찰 체계 통계 기법 적용

Table 3. Procedure for applying statistical techniques to the surveillance and reconnaissance system

데이터 수집	다양한 거리(예: 10km, 20km, 30km)와 기상 조건에서 소수 표적을 탐지한 탐지 데이터를 바탕으로 다수 표적을 탐지한 M&S 결과를 산출한다.
모 델 비교	M&S 결과와 탐지 데이터를 다양한 조건에서 비교하기 위해 분산 분석을 수행한다. 이를 통해 거리와 기상 조건에 따라 표적 탐지 성능 차이를 평가하고, M&S의 신뢰성을 확인한다.
유의미성 평가	ANOVA를 통해 각 조건에 따른 M&S 결과와 탐지 데이터 간의 성능 차이가 통계적으로 유의미한지 p값을 통해 평가한다. p값이 0.05 미만일 경우, M&S 결과와 탐지 데이터 간의 차이가 유의미하다는 결론을 내릴 수 있다.

예를 들어, 탐지 데이터에서 R1km, R2km, R3km 거리에서 각각 소수 표적을 탐지한 데이터가 수집되었고, M&S에서는 다수 표적을 탐지한 결과를 산출하면, ANOVA를 통해 거리별 성능 차이를 분석할 수 있다. 탐지 데이터에서 R1km에서 탐지 성공률이 S1%, R2km에서 S2%, R3km에서 S3%로 나타났고, M&S에서는 R1km에서 s1%, R2km에서 s2%, R3km에서 s3%로 산출되었다면, 두 데이터 간의 성능 차이를 ANOVA로 분석하여 M&S의 신뢰성을 평가한다.

p값이 0.05 미만이면, 조건에 따라 M&S 결과와 실제 데이터 간의 성능 차이가 유의미함을 의미하며, M&S의 결과가 실제

성능을 충분히 반영하지 못할 가능성이 있음을 시사한다.

추가적인 고려사항으로 실제 데이터와 M&S 데이터를 비교하는 경우에서 실제 시험데이터가 현저히 부족한 경우에는 두 데이터 세트를 직접 비교하는 것은 새로운 통계적 오류가 있을 수 있다. 예를 들어, 실제 표적탐지는 10개 표적 10회의 결과 데이터와 시뮬레이션 10개 표적 각 100회의 데이터를 비교한다면 표본의 숫자가 다르므로 이를 위한 전처리가 필요하고 가능한 전처리 방법에 따라 동시신뢰구간(Simultaneous Confidence Intervals) 기법을 적용하여 모델의 정확도를 확보하는 방법을 검토한다.

5.4 정밀 타격 체계 (대공 유도 무기)

대공 유도 무기는 제한된 사격 시험 데이터를 기반으로 M&S를 통해 유도 무기의 타격 정확성을 검증하며, 다양한 공격 시나리오에서 베이지안(Bayesian) 분석을 통해 성능을 평가한다.

대공 유도 무기의 타격 정확성을 다양한 속도 및 환경에서 평가하는 것이 목적이다. 제한된 실제 사격 데이터를 활용하여 M&S 결과를 업데이트하고, 베이지안 추정을 통해 추정 성능을 개선함으로써 유도 무기의 타격 정확성을 더욱 신뢰성 있게 추정한다.

표 4와 같이 베이지안 추정은 기존 데이터와 새로운 데이터를 결합하여 성능을 추정하는 통계 기법이다. 제한된 실제 사격 시험 데이터를 활용하여 M&S 결과를 보완함으로써, 추정 성능을 강화할 수 있다. 대공 유도 무기의 경우, M&S에서 n 회의 실험 데이터를 기반으로 타격 정확성을 추정하는 경우, 실제 사격 데이터 N 회의 결과를 추가하여 추정 결과를 업데이트함으로써 보다 신뢰성 있는 성능 평가를 수행할 수 있다.

표 4 정밀타격 체계 통계 기법 적용

Table 4. Procedure for applying statistical techniques to the precision strike system

데이터 수집	M&S를 통해 500회의 타격 정확성 데이터를 수집하고, 실제로 수행된 10회의 사격 시험 데이터를 수집한다.
베이지안 추정	제한된 실제 데이터를 바탕으로 M&S 결과를 업데이트하고, 새로운 성능을 추정한다. 이 과정에서 기존의 M&S 결과와 새로운 실제 데이터를 결합하여 추정을 강화하고, 신뢰성 있는 타격 정확성 평가를 수행한다.
유의미성 평가	베이지안 추정을 통해 제한된 실제 데이터를 기반으로 M&S 결과를 업데이트한 후, 예상 성능이 실제 성능과 얼마나 일치하는지 평가한다.

예를 들어, M&S에서 몬테카를로 시뮬레이션으로 n 회의 타격 정확성을 추정한 결과가 있다. 실제로는 N 회의 사격 시험이 수행되어 각각 타격 정확성이 C1%, C2%, C3%로 나타났다. 베이지안 추정을 통해 M&S에서 추정된 n 회의 성능 데이터를 실제 N 회의 데이터와 결합하여 새로운 성능을 추정한다. 이 방법은 기존 M&S 결과를 업데이트하고, 추정의 정확성을 개선하는 데 유용하다.

이를 통해 M&S에서 추정된 타격 정확성(예: s1%)이 실제 성능과 얼마나 일치하는지 평가할 수 있으며, M&S 결과가 신뢰할 만한 추정임을 확인할 수 있다. 만약 추정된 성능이 s2%로 도출된다면, 이는 M&S와 실제 데이터의 결합으로 인해 성능 추정이 더 정확해졌음을 의미한다.

6. 통계기법 적용의 장점과 한계

장점으로는 첫째, 비용 절감과 시간 효율성이다. T&E 과정에서 실제로 무기 체계를 시험하는 데에는 많은 비용과 시간이 소요된다. 반면, 통계적 기법을 활용한 M&S는 제한된 실제 데이터를 바탕으로 추가적인 성능 추정을 가능하게 하며, 이는 실험 비용을 크게 절감할 수 있다. 예를 들어, 실제 사격 시험이 제한적인 대공 유도 무기에서 베이지안 추정을 활용하여 성능을 추정할 수 있는 것은 M&S의 효율성을 극대화하는 사례라 할 수 있다.

둘째, 제한된 데이터의 보완이다. V&V 과정에서 M&S 결과를 검증할 실제 데이터가 제한될 수 있지만, 회귀 분석, 분산 분석(ANOVA), 베이지안 추정과 같은 통계 기법은 이러한 제한된 데이터를 보완하여 신뢰성 있는 성능 추정을 가능하게 한다. 예를 들어, 도하 장비의 성능 추정에서 소수의 실제 도하 데이터를 통해 다양한 유속 조건에서의 성능을 회귀 분석으로 추정할 수 있었다.

셋째, 불확실성 관리이다. 특히 베이지안 추정은 기존의 M&S 데이터를 실제 데이터를 통해 보완함으로써 불확실성을 줄이고, 추정의 신뢰성을 높일 수 있다. 이는 M&S를 통해 다양한 시나리오에서 추정된 결과가 실험 데이터와 결합되어, 보다 정확한 결과를 제공할 수 있게 한다.

한계점으로는 첫째, 실제 데이터의 품질과 양에 의존한다는 한계가 있다. 통계 기법은 제한된 데이터를 보완할 수 있지만, 데이터의 품질과 양이 부족할 경우 추정 결과의 신뢰성이 떨어질 수 있다. 예를 들어, 대공 유도 무기의 경우 10회의 실제 시험 데이터를 기반으로 M&S를 보완했지만, 더 많은 데이터가 확보되면 추정의 신뢰성이 더욱 강화될 수 있다.

둘째, 모델의 정확성이다. 통계 기법을 통해 M&S 결과를 검증한다고 해도, M&S 자체가 부정확한 모델을 기반으로 하고 있다면 그 결과는 잘못된 방향으로 흐를 수 있다. 따라서, V&V 과정에서 통계적 기법이 적용되기 전에, M&S 모델 자체의 정확성을 검증하는 것이 중요하다.

셋째, 복잡한 체계에서의 해석 어려움이다. 복잡한 무기 체계일수록 다양한 변수를 고려해야 하므로, 이를 단순한 통계 기법으로 평가하는 데 한계가 있을 수 있다. 예를 들어, 전술 C4I 체계의 경우 통신 네트워크의 복잡성으로 인해 카이제곱 검정만으로는 모든 성능 평가를 포괄하지 못할 수 있다.

따라서 통계적 V&V 기법은 M&S의 신뢰성을 높이고 T&E 과정에서 실험적 제약을 극복하는 데 중요한 도구로 사용될 수 있지만, 데이터의 품질과 모델의 정확성에 대한 지속적인 노력이 필요하다.

7. 결 론

본 연구에서는 무기체계 시험평가 과정에서 모델링 및 시뮬

레이션(M&S)의 신뢰성을 높이기 위한 검증 및 확인 방법의 통계적 기법을 제안했다. 도하 장비, 전술 C4I 체계, 감시 레이더 체계, 대공 유도 무기 등을 대상으로 회귀 분석, 카이제곱 검정, 분산 분석, 베이지안 추정과 같은 통계 기법을 적용하여 M&S 결과와 실제 데이터 간의 일관성을 평가하는 방법을 제시하였다. 본 연구에서 예시는 적용방법의 이해를 위하여 무기 체계별로 구분하였고 예시된 데이터나 방법을 매우 간략하게 정리하였다. 그러나 각 통계 기법은 무기체계의 시험항목 특성에 따라 M&S 결과의 정확성을 제고하는 방향으로 선정하고 분석하는 절차도 매우 심층깊게 적용하여야 한다.

향후 연구에서는 다음과 같은 과제가 필요하다. 첫째, M&S와 실제 데이터를 결합하는 하이브리드 모델에 대한 연구가 필요하다. 이러한 모델은 실제 데이터를 통해 M&S를 보완하고, M&S의 추정 결과를 더욱 정확하게 만드는 데 기여할 수 있다. 둘째, 다양한 무기 체계의 복잡한 상호작용을 반영할 수 있는 다차원적인 V&V 기법이 요구된다. 이는 여러 시스템이 통합되어 작동하는 현대 무기 체계의 특성을 반영하여 보다 종합적인 성능 평가가 가능하게 할 것이다. 마지막으로, AI와 기계 학습을 활용한 자동화된 V&V 기법에 대한 연구가 필요하다. 이러한 기법은 M&S의 정확성을 실시간으로 평가하고 개선할 수 있는 강력한 도구로 자리매김할 수 있다.

References

- [1] 미 국방부 M&S VV&A RPG[3] U.S.DoD, "M&S VV&A RPG Core Document: An Overview of VV&A of Legacy", 2011.
- [2] Defense Acquisition Program Administration Manual No. 2019-10 (2019. 11. 19.)
- [3] U.S.DoD, "M&S VV&A RPG Core Document: V&V Techniques", 2011.
- [4] Doo Hyun Kim, Sang Hoon Kim, "A Study on the Comparison of Test & Evaluation Method for the Guided Missile," Journal of the Korea Association of Defense Industry Studies, Vol.22 No.1 pp. 117-130, 2015
- [5] Moon-gul Lee, Seung-hoon Hwang, Seung-Ryoung Baek, "Methodologies to Decide the Number of Samples and to Verify an Accuracy Rate of the Precise Guided Missiles for Test and Evaluation," Journal of the KIMST, Vol.18 No.5 pp. 558-565, 2015
- [6] Bo-Gil Seo*, Seok-Jin Hong, "A study of estimating the hit probability and confidence level considering the characteristic of Precision Guided Missile," Journal of the Korea Academia-Industrial cooperation Society, Vol. 17, No. 12 pp. 193-197, 2016
- [7] Ojeong Kwon, Hoseok Moon, Chunju Lee, Sunghak Yeon, Doohyun Kim, "A study on method of planning and decision making for test & evaluation using statistical methodology," Journal of the Military Operations Research Society of Korea, Vol.44 No.2 pp. 11-19, 2018
- [8] Sung-Mee Park, Hyun-Jin Han, "A study on the proper number of replication to reach reliable results of AWAM simulation," Journal of the Military Operations Research Society of Korea, Vol.43 No.2 pp. 1-10, 2017
- [9] Sueng-Won Chung, Il-Kuk Jeon, "Statistical Approach for Determining the Number of Tests to Enhance the Reliability of Test and Evaluation Items," Journal of Korea Academia-Industrial cooperation Society, Vol.23 No.11 pp. 352-358, 2022
- [10] Kyung-Min Kang, Joon Huh, "A Study on the Rational Judgement Way of the Test Results for the Weapons System," Annual Korea Institute of Military Science and Technology Conference, pp. 650-651, 2019