

국내 사례로 살펴보는 국방체계 개발의 신뢰성을 높이기 위한 시험평가 방안

서경민^{1,†} · 이찬영¹ · 방경운¹ · 이동철¹ · 최우영¹ · 김탁곤²
대우조선해양 중앙연구원 특수성능연구소¹
한국과학기술원 전기 및 전자 공학과²

Effective Test and Evaluation Approaches for Reliable Defense Systems Development examined through Domestic Defense Cases

Kyung-Min Seo^{1,†} · Chan Young Lee¹ · Kyoung Woon Bang¹ · Dong Chul Lee¹ · Woo Young Choi¹ · Tag Gon Kim²
Daewoo Shipbuilding & Marine Engineering(DSME), Naval & Energy Systems R&D Institute¹
Korea Advanced Institute of Science and Technology(KAIST), Electrical Engineering²

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents practical issues for test and evaluation(T&E) methods to develop defense systems. Our argument is motivated by several domestic defense cases and the cases lead us to discuss two main factors for reliable defense systems development: 1) statistical approaches and 2) technical schemes. Specifically, statistical approaches enable to provide credible interpretations about T&E results in the decision-making process. With practical T&E results of the "Red Shark" torpedo, we performed statistical hypothesis tests and suggest a minimum sample size to accept the hypothesis. Next, technical schemes have more direct effects on improving reliability of developed defense systems and we shortly introduce tools development for systems verification that is required to integrate several sub-systems, e.g., combat, sensor, weapon, and communication systems, within a defense system. We additionally summary some domain cases using modeling and simulation techniques for successful T&E. In closing, we expect that the paper shows empirical investigation and lessons learned with these two practical issues, which provides a guide those who desire to make decisions about reliable defense systems development.

Keywords : Test and evaluation(시험평가), Statistical hypothesis testing(통계적 가설 검정), Combat system(전투 체계), System verification(체계 검증), Simulation(시뮬레이션)

1. 서론

전투/무기체계와 같은 국방체계의 획득(Acquisition)은 요구 성능(ROC: Requirement of Capability)을 만족하는 체계를 계획된 비용과 기간 이내에 개발하여 군에 전력화 하는 것을 목표로 한다. 체계 획득에 필요한 정확한 의사 결정을 위해 체계 개발기관과 운용기관은 시험평가(T&E: Test and Evaluation)를 실시하는데, 이러한 시험평가는 실시하는 기관과 목적에 따라 크게 개발 시험평가(DT&E: Development T&E)와 운용 시험평가(OT&E: Operational T&E)로 구분된다.

최근의 국내 시험평가 사례로는 지난 2009년에 전력화 되었

던 홍상어 어뢰에 대한 운용 시험평가가 대표적이다. 홍상어는 대잠수함 유도 무기체계로 일반 경어뢰의 사정거리를 극복하고 원거리에서 잠수함을 공격하기 위해 개발되었다. 홍상어의 운용 시험평가는 2012년 10월부터 약 1년간 두 차례에 걸쳐 실시되었지만 총 12발 중 8발만 표적에 명중함으로써 시험평가 적합판정 기준인 75%를 달성하지 못해 양산 재개가 보류되었다. 이후 2014년 5월에 추가적인 운용 시험평가를 통해 홍상어 3발이 모두 표적에 명중하였고 전체 시험평가 결과가 73.3%로 적합판정 기준에 근접하여 양산이 재개되었다.

세 차례의 홍상어 운용 시험평가는 신뢰할 수 있는 국방체계 개발의 필요성을 재조명하는 계기가 되었고, 이를 해결하기 위한 다양한 논의가 있었다. 실제로 기술 세미나와 언론을 통해 홍상

어 운용 시험평가의 실패 원인과 함께 국방체계의 신뢰성(품질)을 개선할 수 있는 방안에 대한 토의가 진행되었는데, 예를 들어 홍상어의 입수 시 충격 전달, 전원 공급 커넥터의 성능문제 등과 같이 체계의 상세 설계 측면에서부터 부족한 예산과 사업기간, 그리고 불충분한 초기 시험평가 횟수 등 정책적 측면에까지 다양한 의견들이 제시되었다.

본 논문에서는 신뢰성 높은 국방체계 개발을 위해 시험평가 측면에서 다음의 두 가지 쟁점을 논의한다. 첫 째는 시험평가를 수행할 시에 그 기준과 결과에 대하여 통계적 의미를 고려하는 것이다. 시험평가는 비용 측면으로 인해 제한된 범위 내에서 수행될 수밖에 없다. 그러므로 주어진 시험평가 결과에 대하여 통계적 추정/검정, 신뢰구간 분석과 같은 다양한 통계적 해석은 국방체계 획득의 의사 결정에 있어 신뢰성을 높일 수 있다. 두 번째는 운용 시험평가를 수행하기에 앞서 개발 단계에서부터 체계 결함을 식별하고 기능/성능 상의 기술 개발 목표를 달성하였는지를 판별하는 것이다. 이는 개발 시험평가를 통한 체계 검증 과정이 필수적임을 의미하며 효과적인 검증 기술의 활용은 국방체계 개발의 신뢰성을 높이는 보다 근본적인 대안이 된다.

본 논문은 이러한 두 쟁점을 통계적 접근 방안과 기술적 접근 방안으로 나누어 실증적인 사례를 제시한다. 먼저, 통계적 접근 방안을 설명하기 위해 홍상어 시험평가 결과를 바탕으로 소요군의 적합판정 기준에 대한 달성 여부에 대하여 모비율 가설 검정을 수행한다. 가설 검정의 신뢰성을 높이기 위해서는 적절한 표본의 크기를 확보하는 것이 필요한데, 본 논문에서는 표본의 크기에 따른 가설 검정 추이를 고찰하고, 홍상어 시험평가가 통계적으로 어떤 의미를 전달하는지를 살펴본다.

다음으로 국방체계의 신뢰성 향상을 위한 기술적 접근 방안으로 함정의 전투체계통합(Combat system integration) 과정에 필요한 체계 검증용 장비 개발을 소개한다. 개발 단계에서부터 함께 고려된 체계 검증용 장비는 운용 시험평가를 수행하기 이전에 체계 결함을 식별하고 기능/성능 요구 사항에 대한 일치 여부를 판별할 수 있기 때문에 국방체계 개발의 신뢰성을 높이는 근본적인 대안을 제공한다. 그리고 추가적으로 모델링 시뮬레이션 기술(M&S: Modeling and Simulation)이 시험평가에 활용되었던 사례에 대해서 소개한다. 본 논문에 논의된 두 접근 방안은 향후 효율적인 시험평가를 수행할 수 있는 기준 자료로서 활용되어 국방체계 개발의 신뢰성을 높이는 필수 요소가 되기를 기대한다.

논문의 구성은 다음과 같다. 2장에서 홍상어 개발과 시험평가 사례를 살펴보고 본 사례를 바탕으로 3장에서 홍상어 시험평가에 대한 통계적 접근 방안을 살펴본다. 그리고 4장에서 기술적 접근 방안에 대하여 논의하고, 마지막으로 5장에서 결론을 맺는다.

2. 국내 시험평가 사례

본 장에서는 통계적 접근 방안을 설명하기 위해 최근에 실시된 국내 시험평가 사례로 홍상어 어뢰에 대한 운용 시험평가의 과정과 그 결과에 대하여 설명한다.

2.1 홍상어 개발 과정

홍상어 어뢰는 지난 2000년부터 9년간 A 연구소가 주관하여 1천억 원의 예산을 투입해 개발한 대잠수함 유도 무기체계이다. 홍상어는 물속에서 발사되는 일반 경어뢰의 사정거리 문제를 극복하기 위해 로켓추진기관을 이용하여 공중으로 발사된다. 홍상어 발사 운용은 Fig. 1과 같이 잠수함이 숨어있는 해상까지 10여 km를 비행한 후 입수하여 표적을 타격하는 방식으로 탄두는 기존에 개발되었던 경어뢰인 청상어 어뢰를 사용한다.

A 연구소는 L 산업체와 협력하여 홍상어 개발을 진행 하였고 2009년 6월에 개발 완료하여, 2010년부터 1차 사업분 50여 발을 양산한 후에 해군에 인도되어 2012년 7월에 실전 배치되었다. 2차 양산을 위해 2012년 7월에 동해 상에서 성능 검증 목적으로 시험발사를 실시하였으나 목표물을 타격하지 못하고 유실되었다.

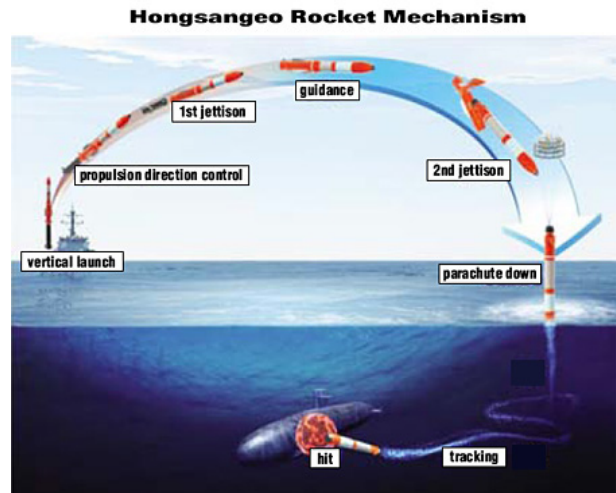


Fig. 1 Operational concept of "Red Shark" torpedo (Yunn, 2014)

2.2 홍상어 운용 시험평가 개요

홍상어의 시험평가 개요를 설명하기 위해 앞서 시험평가의 정의를 살펴보면 다음과 같다. 방위사업청은 시험평가를 무기체계 획득을 위한 구매 또는 연구개발이나 설계 제작이 요구 사항과 일치하는 가를 판단하는 의사 결정 지원 단계로 정의한다. 이러한 시험평가는 크게 개발 시험평가와 운용 시험평가로 구분되는데, 개발 시험평가는 국가 연구기관이나 방위산업체와 같은 체계 개발기관이 실시하는 시험평가로서 제한된 시험 환경에서 체계의 결함을 식별하는 과정이다. 그리고 운용 시험평가는 소요군과 같은 체계 운용/시험기관이 실제와 유사한 전장 환경에서 체계가 전투용으로 적합하고 전력화 준비가 되었는지를 판단하는 과정이다. 시험평가의 정의에 의하면 홍상어 시험평가는 실시하는 기관, 조건, 평가 목적을 고려할 때 운용 시험평가에 해당한다.

Table 1은 지난 3년 간 실시된 홍상어의 운용 시험평가에 대

Table 1 Summary of OT&E(Operational Test & Evaluation) results for “Red Shark” torpedo

T/E stage	Period	Objective	OT/E design	OT/E result	Implication
Test-firing	July, 2012	Performance verification for additional production	One practice unit	One failed	<ul style="list-style-type: none"> Additional production suspended Prehension of technical factors needed
The 1 st OT/E	Sep. 2012 ~ Feb. 2013	Quality improvement	Five practice units Three live units	Three of eight failed (one practice and two live units failed)	<ul style="list-style-type: none"> A criteria for acceptance - 75% for hit ratio OT&E results - 66.7% for hit ratio long-delayed mass production occurred
The 2 nd OT/E	July, 2013 ~ Oct, 2013	Quality improvement	Two practice units Two live units	One of four failed (one live unit failed)	
The 3 rd OT/E	May, 2014	Quality improvement	One practice unit Two live units	No failed All units succeeded	<ul style="list-style-type: none"> Total OT&E results - 73% for hit ratio Mass production will begin soon

한 개요이다(Park, 2014). 본 자료는 언론을 통해 공식적으로 알려진 사실을 바탕으로 비공개 운용 시험평가가 있을 경우를 배제하고 작성되었다. 먼저, 운용 시험평가를 수행하기 전인 2012년 8월에 1발의 기능 시험발사를 실시하였는데 표적 명중에 실패하자 추가 양산을 중단하고 품질 개선 작업을 진행하였다(본 시험발사는 운용 시험평가에서 제외하고 설명한다). 이후 소요군은 2012년 9월부터 운용 시험평가를 수행 계획을 수립하고 표적 명중률 75% 이상을 전투용 적합판정 기준으로 제시하였다.

세 차례에 걸친 운용 시험평가의 개요는 다음과 같다. 먼저 2012년 9월부터 2013년 10월까지 두 차례의 운용 시험평가를 수행하였으나 총 12회의 시험평가 중 8회가 표적에 명중하였고 66.7%의 표적 명중률을 기록함으로써 적합판정 기준(75%)을 달성하지 못해 추가 양산이 보류되었다. 품질 개선 작업 후 2014년 5월에 실시된 세 번째 운용 시험평가에서는 3회 모두 표적에 명중하였고 세 차례의 운용 시험평가의 전체 명중률은 73%를 기록하였다. 운용 시험평가의 전체 명중률이 전투용 적합판정 기준에 미치지 못하는 못하였지만 방위사업청과 소요군은 품질 개선 작업이 마무리된 이후 시행된 운용 시험평가에서 홍상어 3발이 모두 표적에 명중했기 때문에 품질 향상이 이루어 졌다고 판단하였고 양산 재개에 문제가 없다는 의사 결정을 하였다.

3. 통계적 방안

본 장에서는 2장에서 설명한 홍상어 시험평가 사례를 바탕으로 시험평가 기준과 결과에 대한 통계적 해석을 제시한다. 통계 분석을 위해 먼저 소요군의 적합판정 기준에 대한 달성 여부를 가설로 설정하고 시험평가 결과에 대하여 가설 검정을 수행한다. 그리고 신뢰성 있는 의사 결정을 위해 필요한 시험평가 횟수에 대한 통계적 의미를 고찰한다.

3.1 모비율 가설 검정

통계적 가설 검정은 모집단으로부터 추출된 표본(Sample)을 바탕으로 가설(Hypothesis)의 합당성 여부를 통계적으로 판단하는 과정으로 여기서 가설이란 모집단에 대하여 주장하는 바를 표현한 명제를 말한다. 통계적 가설 검정은 일반적으로 검정하고자 하는 모수(Parameter)에 따라 모분산/모평균/모비율 가설 검정으로 나뉘는데 홍상어 시험평가의 경우에는 결과가 명중 여부에 대한 비로 표현되는 명중률을 평가하는 것이기 때문에 모비율(Proportion) 가설 검정에 해당한다.

Fig. 2를 통해 홍상어 운용 시험평가에 대한 모비율 가설 검정 단계를 순서도로 나타내었다. 가설 검정은 일반적으로 연구가설을 통하여 규명되어야 하는 귀무 가설(Null hypothesis) H_0 과 귀무 가설과 대립되는 대립 가설(Alternative hypothesis) H_1 로 분류되고, 검정하고자 하는 모수와 추정 값(Estimated value)의 대소 관계를 통해 가설 검정(H_0 의 채택 여부)을 수행한다. 한 예로, 홍상어 운용 시험평가의 경우 H_0 는 홍상어 전체의 명중률 p 는 적합판정 명중률 p_0 이상인 경우로 수립 가능하다.

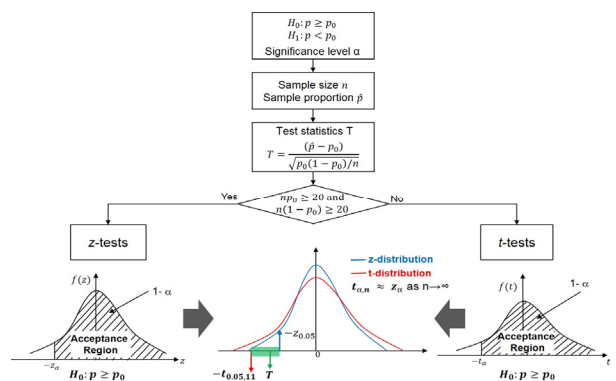


Fig. 2 Two hypothesis testings depending on sample size: t -procedures and z -procedures

가설이 수립되면 가설 검정에 필요한 표본의 크기(Sample size) n 과 유의 수준(Significance level) α 를 결정해야 한다. 여기서 유의 수준은 H_0 가 참임에도 불구하고 그것을 기각할 확률을 의미하는데 많은 공학 응용 사례에서 0.01(1%)과 0.05(5%)를 사용한다. 예를 들어 유의 수준이 0.05이면 H_0 가 참이지만 기각할 확률이 95%임을 의미한다. 다음으로 가설 검정을 위한 실험을 수행하고 주어진 실험 결과를 바탕으로 표본 비율(Sample proportion) \hat{p} 과 검정통계량(Test statistics) T 를 산정한다. 마지막으로 가설 검정은 np_0 와 $n(1 - p_0)$ 의 크기에 따라 t -분포를 이용하는 소표본 검정과 z -분포(정규 분포)를 이용하는 대표본 검정을 나누어 실시된다.

검정 단계에서 H_0 가 채택되기 위해서는 Fig. 2의 하단에 나타난 그래프와 같이 T 가 채택역(Acceptance region)에 포함되어야 한다. 채택역의 포함 여부는 검정에 사용한 분포의 임계점(Critical point)과 T 의 대소 관계로 판단한다. 예를 들어, Fig. 2의 소표본 가설 검정의 경우 H_0 가 채택되기 위해서는 $T \geq -t_{\alpha, n-1}$ 을 만족해야 한다. 자세한 통계적 가설 검정은 참고 문헌을 참조하기 바란다(Hayter, 2012).

3.2 시험평가의 가설 검정

본 절을 통해 Fig. 2의 검정 단계를 흉상어 운용 시험평가에 적용하여 가설을 검정한다. Table 1에 의하면 소요군은 시기적으로 다른 세 차례의 운용 시험평가를 수행하였지만, 본 논문에서는 이들을 하나의 시험평가로 간주하여 가설 검정을 수행하였다. 본 사례에 대한 가설 검정의 유의 수준은 일반적으로 많이 사용되는 0.05와 0.01로 설정하였고, 가설 검정 단계에 따른 결과는 식 (1)에서 (5)로 정리하였다.

$$H_0 : p \geq 0.75, \quad H_1 : p < 0.75 \quad (1)$$

$$\alpha = 0.05, 0.01$$

$$n = 15, \quad \hat{p} = 0.733 \quad (2)$$

$$T = \frac{(0.733 - 0.75)}{\sqrt{0.75 \cdot 0.25/15}} = -0.152 \quad (3)$$

$$t_{0.05, 14} = 1.791, \quad t_{0.01, 14} = 2.624 \quad (4)$$

$$-0.152 \geq -1.791, \quad -0.152 \geq -2.624 \quad (5)$$

$\therefore H_0$ accept when $\alpha = 0.05, 0.01$

운용 시험평가의 적합판정 명중률이 75%이므로 H_0 는 식 (1)과 같고, 총 15회의 시험평가를 수행했으므로 n 은 15이다. 시험평가 결과를 바탕으로 T 를 계산하면 식 (3)에 의해 -0.152이고, 가설 검정에 필요한 임계점은 주어진 α 와 n 에 따라 1.791과 2.624로 산출되었다. 식 (5)는 가설의 채택 여부를 판별하기 위한 과정으로 T 가 임계점의 음수보다 크므로 H_0 는 유의 수준 0.05와 0.01에서 채택이 가능하다. 결론적으로 본 가설 검정 결과에 의하면 15회의 운용 시험평가를 수행하였을 때 0.05, 0.01의 유의 수준에서 전체 흉상어 어뢰의 명중률은 75% 이상이라 할 수 있다.

3.3 시험평가 횟수에 따른 가설 검정

표본 실험은 모집단에서 추출된 표본의 다양성으로 인해 해당 표본의 크기가 모수 추정의 정확도에 영향을 미친다. 따라서 표본을 추출하기 전에 표본 크기에 따른 가설 채택의 가능성을 역으로 해석하거나 기 결정된 표본에 대하여 실험을 수행한 후에 가설 채택을 위해 추가적으로 필요한 표본의 크기를 산출하는 것은 좋은 분석 방법이 된다.

Table 2 Summary of t-procedures at 12 to 20 sample sizes(n)

$H_0 : p \geq 0.75 \quad (\alpha = 0.05)$		
Minimum number of hits(n_{Hit}) to accept H_0		
n	n_{Hit}	\hat{p} (Hit ratio, %) = $(n_{Hit}/n) \times 100$
20	12	60.00
19	11	57.89
18	11	61.11
17	10	58.82
16	9	56.25
15	9	60.00
14	8	57.14
13	7	53.85
12	7	58.33

Table 2는 식 (1)의 H_0 에 대하여 유의 수준이 0.05일 때 표본 개수 n 을 달리 하면서 가설 검정을 만족하는(H_0 를 채택할 수 있는) 최소 명중 횟수(n_{Hit})를 분석한 것이다. 한 예로, n 이 15일 때 H_0 를 채택할 수 있는 n_{Hit} 은 9이고, 이는 15회의 시험평가에서 9회 이상 명중하면 H_0 를 채택할 수 있음을 의미한다. 여기서 주목해야 할 부분은 n 이 12인 경우이다. 흉상어 운용 시험평가의 경우 2013년 10월까지 시행한 두 차례에서 12회

중 8회가 명중하였는데, Table 2에서 n 이 12인 경우 n_{Hit} 은 7
이므로 두 차례의 시험평가 결과에 대해서도 통계적으로는 유의
수준이 0.05일 때 H_0 를 채택할 수 있는 근거를 제공한다.

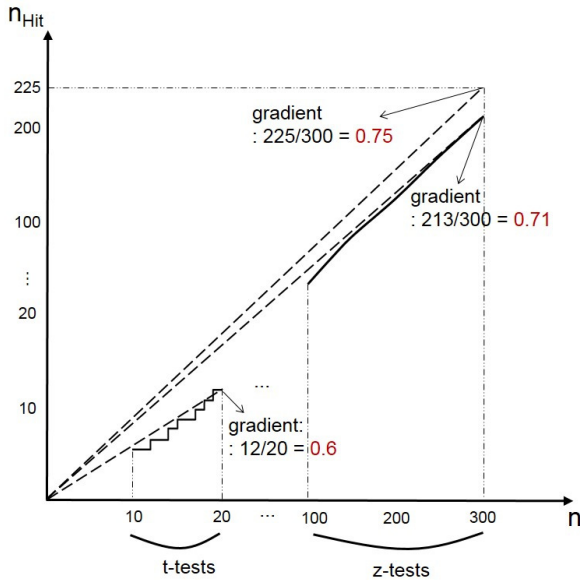


Fig. 3 Tendency analysis for minimum number of hits to accept H_0

Table 2를 통해 n 이 작은 경우 H_0 가 채택 가능한 표본 비율 \hat{p} 과 모비율 p 은 큰 차이를 보임을 알 수 있다. 이러한 현상을 좀 더 구체적으로 분석하기 위해 n 의 크기에 따른 \hat{p} 의 경향성을 Fig. 3에 나타내었다. Fig. 2에서 제공한 순서도에서 판단 기호의 내부 수식에 따라 n 이 [10, 20] 인 경우에는 t -분포를 이용한 소표본 검정을 수행하였고 n 이 [100, 300] 인 경우에는 z -분포를 이용한 대표본 검정을 수행하였다. Fig. 3 그래프의 평균 기울기(Average gradient)는 해당 표본 개수 n 에서의 \hat{p} 을 나타낸다. 예를 들어, 300회의 실험을 수행했다면 H_0 를 채택하기 위해 213회의 성공이 필요하고, 이 경우에 표본에 대한 명중률은 71%로 모비율인 75%에 근접하게 된다.

3.4 통계적 접근 방안에 대한 고찰

앞서 설명하였듯이 가설 검정은 표본을 추출하여 전체 모집단에 대한 가설을 검정하는 과정이다. 대부분의 경우 모집단 전체에 대한 실험은 불가능하기 때문에 표본 집단을 이용한 가설 검정을 수행한다. 그러나 Table 2와 Fig. 3을 통해서 표본의 크기가 작은 경우에는 가설 채택이 가능하더라도 표본 비율이 모비율과 상대적으로 큰 오차를 보임을 확인하였다. 홍상어 운용 시험평가의 경우를 살펴보면 두 차례의 운용 시험평가 결과만으로도 전체 홍상어 명중률이 75% 이상이라는 가설을 채택할 수 있지만 표본 크기(12회)가 작은 만큼 신뢰 구간(Confidence interval)이 넓어져서 가설 검정의 의미가 약해진 경우로 볼 수 있다.

$$n = \frac{z_{1-\alpha}^2 p(1-p)}{\delta^2} \tag{6}$$

가설 검정의 신뢰성을 높이기 위해서는 적절한 표본의 크기를 확보하는 것이 필요한데, 통계학에서는 가장 보수적인 방법으로 표본의 크기를 산출하는 공식이 식 (6)으로 알려져 있다(Hayter, 2012). 식 (6)은 정규 분포를 활용한 식으로 δ 는 모비율 p 에 대한 오차 범위를 의미한다. 예를 들어 표본 비율 \hat{p} 이 p 와 $\pm 5\%$ 이내의 오차 범위를 가진다고 가정하면 δ 는 0.10이다. δ 가 0.10이고 유의 수준이 0.05인 경우에 대하여 홍상어 운용 시험평가를 식 (6)에 적용하면 n 은 약 50.70이다. 이는 51회의 운용 시험평가를 실시하였을 때 가장 보수적으로 신뢰할 만한 통계적 자료를 추출할 수 있음을 의미한다.

이와 같이 통계적 관점에서 표본 크기의 증가는 가설 검정력을 높일 수 있지만 현실적으로 시간과 비용의 문제를 야기한다. 예를 들어, 1회 발사에 약 1억 원이 투입되는 홍상어 시험평가를 고려해 볼 때 50회 이상의 시험평가는 시험평가 비용에만 약 50억 원 이상이 소요된다. 본 사례는 시험평가의 신뢰성과 비용은 서로 상충(Trade off) 관계에 있기 때문에 시험평가의 범위 설정 시 지나친 엄격성을 적용하는 경향을 방지하고 적정 수준의 시험평가 수행을 위한 검토가 필요함을 암시한다.

따라서 통계적 분석 결과는 시험평가의 절대적인 지표를 제시하는 것이 아니라 다양한 통계량을 바탕으로 의사 결정 과정에 신뢰성을 제공할 수 있는 근거 자료로 활용 되어야 한다. 예를 들어, 표본을 추출하기 전에 표본 크기에 따른 가설 채택의 신뢰 구간을 분석하거나 실험을 수행한 후에 추가적으로 필요한 통계적 자료의 분석은 의사 결정자가 국방체계 획득을 결정하는 데 있어 활용될 수 있는 공학적 근거 자료가 된다.

정리하면 통계적 접근 방안은 적정 수준의 시험평가 수행을 위해 검토해야 할 정보를 제공하고 해당 정보는 국방체계 획득 과정에서 의사 결정의 신뢰성을 높이는 방안으로 활용되어야 한다.

4. 기술적 방안

3장에서 살펴 본 접근 방안은 통계적 특성을 내포하는 자료를 수집/분석하여 일정한 체계성을 규명하는 데이터 분석 과정이다. 이러한 통계적 접근은 국방체계의 결함을 찾아서 보완하는 등 기술적 신뢰도를 높인다기보다 실행한 결과에 대한 의사 결정의 신뢰성을 높이는 방안이다.

본 장은 보다 근본적으로 국방체계 개발의 신뢰성을 높일 수 있는 기술적인 접근 방안을 소개한다. 이를 위해 먼저 함정전투 체계(Naval combat system)의 체계 검증(System verification)을 위해 개발되고 있는 검증 장비를 설명하고, 부가적으로 개발/운용 시험평가에 M&S 기술을 활용한 사례를 소개한다.

4.1 체계 검증용 장비 개발

D 산업체가 진행하고 있는 함정 성능개량사업은 함 운용 기간

중에 국내 연구/개발 기술로 함정의 통합전투성능을 개량하는 선진국형 사업으로 성능 개선된 함정전투체계의 개발 및 통합이 본 사업의 핵심 요소이다(DSME, 2014). 함정전투체계는 Fig. 4의 좌측 그림과 같이 전투관리/센서/항해/무장/통신 체계와 같은 이기종 하위 체계들의 복합 체계(System of systems)로 구성된다(Kim, et al., 2013). 복합 체계 기반의 함정전투체계는 성능개량 사업을 통하여 Fig. 4의 우측 그림과 같이 기존의 주요 탑재된 하위 체계를 개선된 체계로 교체하거나 신규 체계를 탑재함으로써 통합전투성능을 향상시킬 수 있다.

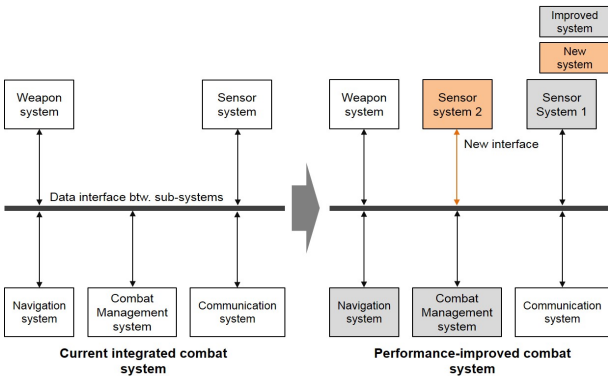


Fig. 4 Concept of performance improved plan for combat system integration

함정전투체계의 통합과 더불어 시험평가 관점에서 통합전투성능을 극대화하기 위해 전투체계통합(CSI: Combat System Integration) 과정이 제안되었다(Lee, 2013). 전투체계통합은 전투체계 통합 및 연동을 수행하기 위해 계획 수립에서부터 탑재 장비별 성능/기능 요구조건과 연동 요구조건을 분석하여 탑재 장비를 선정하고 연동계통을 설계하는 일련의 활동을 의미한다.

이러한 전투체계통합은 체계 오류 원인 식별 및 통합성능 검증과 같은 시스템 레벨의 검증 과정을 반드시 필요로 한다. 구체적으로 시스템 레벨의 통합전투체계 검증하기 위해서는 이기종(Heterogeneous) 체계들이 통합되어 상호작용하는 연동 상황을 확인 할 수 있어야 하고, 일부 하위 체계들에 대해서 통합전투성능을 측정하여 해당 체계가 주어진 성능 지수를 달성할 수 있는지에 대한 성능 분석을 수행해야 한다. 본 사업에서는 이러한 체계 검증을 위해 연동 검증장비와 성능 분석장비로 나누어 개발된다.

Table 3 Summary of M&S techniques for efficient T&E

Previous works	Target system	Objectives	Proposed M&S techniques
Cho, et al., 2012	Mid-range surface to air guided missile (Cheongung)	<ul style="list-style-type: none"> Cheongung system interoperability test Hostile fighter and guided missile generation(simulation) Radar signal generation(simulation) 	<ul style="list-style-type: none"> Simulation-based integrated test system Simulation environment for V(Virtual) and C(Constructive) systems interoperation
Kim, et al., 2014	Ship to ship missile	<ul style="list-style-type: none"> Engagement effectiveness analysis Utilization of system design, construction, and integration 	<ul style="list-style-type: none"> Simulation environment for engagement effectiveness and performance analysis Missile display information system
Lee and Seo, 2014	General weapon system with specific performance index	<ul style="list-style-type: none"> Evaluation and visualization of weapon systems' operational performances 	<ul style="list-style-type: none"> Process-based performance evaluation model Process-based operational test and evaluation systems for weapon systems

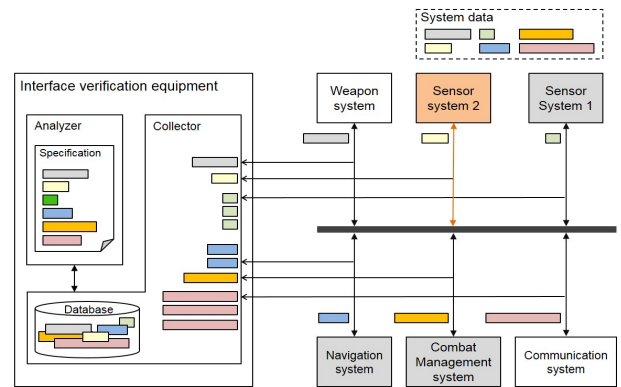


Fig. 5 Conceptual structure of interface verification equipment for systems verification

먼저, 연동 검증장비는 전투관리 체계가 센서/항해/무장 체계들과 연동 데이터를 교환할 때 기 정의된 명세(Specification)에 맞게 해당 연동 데이터를 주고받는 지를 검증한다. 이를 위해 연동 검증장비는 체계 간 인터페이스에 연결되어 연동 데이터(시스템 데이터)를 수집하는 수집기와 수집된 데이터를 바탕으로 분석 및 검증을 수행하는 분석기로 구성된다. 그리고 연동 데이터에 대한 명세는 디지털 정보를 교환하기 위한 통신 기술과 표준 메시지 형식을 정의하고, 이를 바탕으로 연동 데이터를 해석 가능한 형태로 변환 하여 최종 검증 작업을 수행한다.

Fig. 5는 연동 검증장비의 구성을 도식화하였다. 분석기 내부에는 하위 체계들이 송수신하는 연동 데이터에 대한 명세를 서로 다른 길이와 색을 가진 블록으로 표시하였고, 같은 방법으로 하위 체계들 사이에서 실제 송수신하는 연동 데이터는 Fig 5의 우측 범례에 표현하였다. 여기서 블록의 길이는 연동 데이터의 형식을 의미하고 색은 데이터 값의 범위를 의미한다. 하위 체계는 각기 구별된 연동 데이터를 전송하고 연동검증장비의 수집기는 이러한 연동 신호를 분기하여 수집한다. 분석기는 기 정의된 연동 문서에 따라 연동 관련 명세를 저장하여 명세와 실제 수집된 연동 정보가 일치하는 지 여부를 판단하여 검증 작업을 수행한다. Fig 5의 예에서는 센서 체계 1이 전송하는 연동 데이터에서 분석기에 설정된 명세와 비교하였을 때 블록의 색이 다르므로 데이터의 형식은 동일하지만 범위가 상이한 경우이고, 통신 체계의 경우에는 블록의 길이가 다르므로 데이터의 값은 동일하지만 데이터의 형식이 잘못된 경우를 나타낸다.

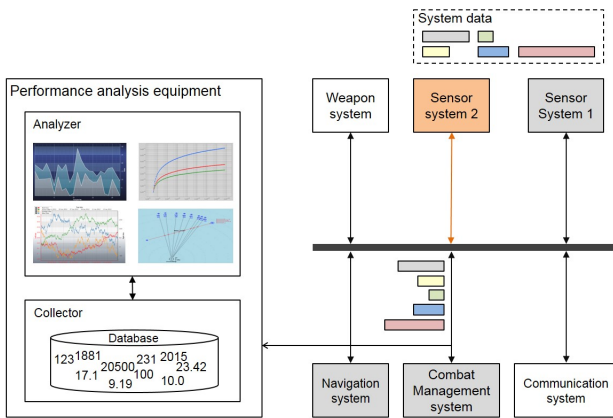


Fig. 6 Conceptual structure of performance verification development for systems verification

다음으로 성능 검증장비는 신호 처리가 완료된 센서 데이터를 실시간으로 수집하여 운용자가 소나, 항해, 레이더 센서 체계들의 방위 정확도 등의 체계성능을 확인 할 수 있도록 한다. 성능 검증장비는 연동 검증장비와 유사하게 시스템 데이터를 수집하는 수집기와 수집된 데이터를 바탕으로 분석 및 성능평가를 지원하는 분석기로 구성된다.

성능 검증장비의 운용 예를 살펴보면 Fig. 6과 같다. Fig. 6에서 전투관리 체계는 각 센서 체계로부터 연동 데이터를 전달받는데, 각 연동 데이터는 센서 체계가 수집한 수치 데이터를 포함한다. 수치 데이터는 구체적으로 센서 체계로부터 수집된 표적 방위의 오차를 분석하여 센서 체계에 대한 성능을 검증하거나 다수의 센서 체계들로부터 수신된 정보를 바탕으로 전투 체계의 표적기동분석(TMA: Target Motion Analysis)에 대한 성능을 검증한다. 이를 위해 성능 검증장비는 수신한 센서 데이터들을 바탕으로 다양한 통계적 해석을 가시화하여 보여 준다. 본 논문에서는 언급된 체계 검증용 개발 장비에 대하여 간단한 내용을 소개하는 것으로 제한하고 보다 자세한 내용은 추후 연구들에서 수행토록 한다.

이러한 체계 검증용 장비들은 실제 함정통합을 수행하기 이전에 체계 검증을 위해 일차적으로 육상시험체계(LBTS: Land Based Test System)에 설치되어 개발 시험평가를 수행한다. 육상에서 개발 시험평가를 완료한 체계는 함정에 탑재되어 전투체계통합을 수행하게 되는데, 개발 단계에서부터 함께 고려된 체계 검증용 장비는 운용 시험평가를 수행하기 이전에 체계 결함 식별하고 기능/성능 요구 사항에 대한 일치 여부를 판별할 수 있기 때문에 국방체계 개발의 신뢰성을 높이는 근본 원인을 제공한다.

4.2 모델링 시뮬레이션 기술 활용

국방체계 개발의 신뢰성을 높이기 위한 또 다른 방안으로 M&S 기술을 시험평가에 활용할 수 있다. M&S 기술 (Zeigler, et al., 2000)의 활용은 최근 국방 분야에서 모의기반획득(SBA: Simulation Based Acquisition)이 시험평가에 필수적인 정책으로 인식되었음을 반영한 결과이다.

표 3은 개발/운용 시험평가에 M&S 기술을 직·간접적으로 활용하였던 몇 가지 연구 사례를 정리한 것이다. 먼저, Cho, et al.(Cho, et al., 2012)은 중거리 지대공 유도 무기에 대한 시험 평가를 수행하기 위해 M&S 기술을 활용하였다. M&S 활용 목적은 크게 천공 체계의 연동검증, 실시간 공중위협 및 유도탄 모의, 그리고 레이더 신호 생성을 위함이었으며, 이를 위해 천공 통합시험체계를 개발하였다. 개발된 통합시험체계는 체계 시험 평가 단계에서 실 무기체계와 모의장비의 연동을 통해 HILS(Hardware-In-the-Loop Simulation)를 수행하여 체계의 신뢰성을 검증하였다. 제안하는 M&S 기술은 비용 및 환경 제약 등으로 실제 사격 시험이 어려운 시험항목에 대한 평가가 가능하고, 개발기간 단축 및 비용절감 효과를 기대할 수 있으며, 향후에는 사용자 운용성 평가를 위한 MILS(Man-In-the-Loop Simulation)에도 활용이 가능하다.

다음으로, Kim, et al.(Kim, et al., 2014)은 함대함 유도 무기 체계에 대하여 M&S 기술을 활용한 사례를 소개하였다. 구체적으로는 교전 효과도 분석을 위해 일대일 교전 시뮬레이션을 수행하여 대략적인 효과도 분석을 수행하고, 요구 분석, 체계 설계 및 개발, 시험 평가 등에 활용하기 위해 체계 성능 시뮬레이터를 개발하여 이용하였다. 그리고 실제 체계의 설계 및 제작, 그리고 통합에는 디지털 목업 시스템을 활용하였다. 본 연구의 M&S 활용은 부체계를 개발하기 위한 보조도구 혹은 실제 운용시험평가 시에 평가를 위한 보조 데이터 산출을 위한 도구로 제한적으로 활용되었다.

마지막으로, Lee & Seo(Lee and Seo, 2014)는 프로세스 기반의 모델링을 통한 무기체계 운용시험평가 시스템을 개발하였다. 제안하는 프로세스 기반 성능평가 모델은 성능평가의 대상이 되는 무기체계를 물리 정보와 행위 정보로 나누어 구성한 후 데이터베이스화 하여 저장소에 저장한다. 그리고 해당 정보를 다양하게 결합하여 무기체계의 성능을 표현하였고, 다른 무기체계들과의 상호 작용을 프로세스 형태로 표현하여 전체 성능평가를 수행하였다. 또한, 성능평가 모델을 상업용 게임 엔진에 적용하여 3차원으로 가시화함으로써 무기체계의 운용성능을 시각적으로 표현이 가능한 프로세스 기반의 무기체계 운용시험평가 시스템을 제안하였다. 본 연구는 데이터베이스화 되어 있는 기존의 정보가 부족하고 실제 설계할 수 있는 프로세스 기반 모델의 충실도(Fidelity)가 낮은 점 등과 같이 아직은 프로토타입 수준이기는 하지만 M&S 기술을 운용 시험평가에 활용할 수 있는 가능성을 실증적 자료를 바탕으로 제시한 사례로 볼 수 있다.

이상과 같이 M&S 기술을 활용한 시험평가 방안은 실 체계 기반의 시험평가를 보다 효율적인 형태로 개선하여 시험평가에 소요되는 시간 및 비용을 절감하고 효율성을 증대시킬 수 있는 대안을 제시한다. 물론 이를 위해서는 M&S 체계가 실 체계와 실시간 연동 시뮬레이션이 가능해야 하거나 M&S를 활용한 시험평가의 결과를 신뢰할 수 있도록 M&S 결과에 대한 W&A(Verification, Validation, & Accreditation) 프로세스를 진행하여 결과의 충실도를 높이는 등의 기술적인 과제들을 해결해야 한다. 따라서 우리나라는 국방체계의 시험평가를 위한 기반시설, 인력, 연구기간 등의

인프라가 국방 선진국에 비해 미흡한 수준이기 때문에 한정된 인프라를 효율적으로 활용하여 신뢰성 있는 시험평가 체계를 구축하기 위한 대책이 필요할 것으로 보인다.

5. 결론

본 논문은 시험평가 관점에서 신뢰성 높은 국방체계 개발을 위한 두 가지 접근 방안을 제시하였다. 먼저, 통계적 접근 방안은 적정 수준의 시험평가 수행을 위해 검토해야 할 공학적 정보를 제공하여 국방체계 획득 과정에서 의사 결정의 신뢰성을 높일 수 있다. 그리고 국방체계 개발의 신뢰성을 보다 근본적으로 해결할 수 있는 기술적 접근 방안으로 연동/성능 검증을 위한 체계 검증용 장비의 개발을 소개하였고, 최근 시험 평가에 많이 활용되고 있는 M&S 기술에 대해서도 몇 가지 연구 사례를 살펴보았다.

시험평가는 획득 전 주기 동안에 수행되어야 하지만 여러 가지 현실적 원인으로 일부 단계에서만 수행되었던 것이 사실이다. 최근 언론과 기술 세미나를 통해 논의되는 국내 개발 전투/무기 체계의 신뢰성 문제 등을 고려할 때 개발 단계에서 양산 과정에 이르기까지 시험평가를 통한 품질 보증과 개선이 가능케 하는 일련의 전문적인 검증 프로세스가 필요하다. 본 논문에 논의된 접근 방안이 향후 이러한 검증 프로세스에 활용될 수 있는 기준 자료가 되기를 기대한다.

References

Cho, K.T. Lee, S.Y. Lee, H.M. Kim, S.H. & Jung, H.M., 2012. Enhancing the Efficiency and Reliability for M&S based Test and Evaluation System Development. *The Korea Society for Simulation*, 21(1), pp.89-96.

Daewoo Shipbuilding & Marine Engineering (DSME), 2014. DSME ordered Jangbogo-I PIP Project. [Online] (Updated 1 September 2014) Available at: http://www.dsme.co.kr/pub/ds/td/dstd030Q.do?dt_type=tod&dt_seq_no=3008¤tPageNo=1

Hayter, A., 2013. *Probability and Statistics for Engineer and Scientists* (4th edition). Chester: Brooks/Core.

Kim, H.H. Lee, C.H. & Cho, N.H., 2009. A Study on Application and Promotion of DM&S W&A (Focused on M&S for Test and Evaluation). *The Korea Society for Simulation*, 18(4), pp.157-164.

Kim, T.G. Kwon, S.J. & Kang, B.G., 2013. Modeling and Simulation Methodology for Defense Systems Based on Concept of System of Systems. *Journal of the Korean Institute of Industrial Engineers*, 39(6), pp.450-460.

Lee, B. & Seo, Y.H., 2014. A Design of Operational Test & Evaluation System for Weapon Systems thru Process-based Modeling. *The Korea Society for Simulation*, 23(4), pp.211-218.

Lee, B.J., 2013. Platform-based CSI methods to fulfill efficient test & evaluation. *2013 Seminar for Test & Evaluation of Weapon Systems*, pp.77-90.

Park, S.C., 2014. Succeed in Test-firing "Red Shark" torpedo, Segye-Ilbo. [Online] (Updated 26 May 2014) Available at: <http://www.segye.com/content/html/2014/05/26/20140526001117.html>

Yuun, S.H., 2010. <2>Torpedo, Kookbang-Ilbo. [Online] (Updated 1 September 2014) Available at: <http://kookbang.dema.mil.kr/>

Zeigler, B.P. Praehofer, H. & Kim, T.G., 2000. *Theory of Modeling and Simulation 2nd Edition*. Academic Press.

