

Original Article

<https://doi.org/10.12985/ksaa.2026.34.2.059>
ISSN 1225-9705(print) ISSN 2466-1791(online)

Weibull 분석과 베이지안 네트워크를 활용한 항공기 NFF LRU 신뢰성 예측 연구

문희준*, 김현덕**, 김수진*, 이석종**

A Study on Predicting Reliability for ‘No Fault Found’ LRU using Weibull Distribution Analysis and Bayesian Network

Heejun Moon*, Hyeondeok Kim**, Sujin Kim*, Seokjong Lee**

ABSTRACT

No fault found (NFF) events in aircraft maintenance increase maintenance cost, reduce aircraft availability, and weaken reliability management because removed components are returned without confirmed defects. This study proposes a reliability prediction model for NFF line replacement units (LRUs) by combining Weibull analysis and Bayesian Network modeling. Maintenance records of three avionics LRUs, TCAS Computer, EIVMU, and SPB, were analyzed using failure data and censored data. Weibull analysis was applied to estimate shape parameters and NFF occurrence probabilities at specific time-since-task values, while Bayesian Networks were used to infer NFF probabilities under operational and maintenance conditions. The proposed framework provides a practical decision-support tool for maintenance organizations by integrating time-based reliability characteristics with probabilistic casual inference, supporting NFF screening, decision-making, and reliability-centered component management.

Key Words : No Fault Found(원인불명결함), Line Replacement Unit(교환가능부품), Weibull Analysis(웨이불 분석), Bayesian Network(베이지안 네트워크), Reliability Prediction(신뢰성 예측), Aircraft Maintenance(항공정비)

1. 서 론

현대 항공기의 고장 부품에서 발생하는 No Fault Found(NFF, 원인 불명 결함) 현상은 유지보수 비용의 증가와 정비 효율성 저하를 초래하는 주요 문제이다. LRU(line replacement unit) 전자 부품들이 정비 과

정에서 NFF가 빈번하게 발생하고 있으며, 항공전자 계통(avionics)의 비중이 증가함에 따라 그 복잡성도 심화되고 있다(Khan et al., 2014).

NFF는 항공기 상에서 결함이 의심되어 부품이 장탈된 후 MRO(maintenance repair overhaul) 수리업체에서 테스트 결과 어떠한 결함도 발견되지 않는 것을 의미한다. British Airways의 1992년 데이터에 따르면 전체 비계획 장탈의 약 14%가 NFF였으며, 연간 비용이 2,000만 파운드에 달했다. ATA 보고서에 의하면 200대 항공기 운항 항공사의 연간 NFF 비용은 약 2,000만 달러로 추산된다.

기존의 예지 정비(predictive maintenance)와 신

Received: 2. Jun. 2026, Revised: 8. Jun. 2026,
Accepted: 11. Jun. 2026

* 한국항공대학교 운항관리학과 대학원생

** 한국항공대학교 항공운항학과 교수

연락처 E-mail : afleesjgo@kau.ac.kr

연락처 주소 : 경기도 고양시 덕양구 항공대로 76,
본관 414호

뢰성 기반 정비(RCM) 체계에도 불구하고 NFF 문제는 해결되지 않고 있다. 본 연구는 실제 K항공사의 LRU 부품 데이터를 기반으로 Weibull 분포 생존 분석과 베이저안 네트워크 모델을 결합하여 NFF 발생확률을 예측하고, 정비 의사결정의 정확성을 높이는 방법론을 제시한다.

II. 본 론

2.1 이론적 배경

2.1.1 NFF(No Fault Found) 개념 및 원인

NFF는 NTF(no trouble found), CND(can not duplicate), RTOK(retest OK) 등 다양한 명칭으로 불리며, 항공 산업에서 가장 빈번하게 사용되는 용어는 no fault found이다(Khan et al., 2013). NFF의 주요 원인은 James et al.(2003)에 의해 10가지로 분류되었으며, 간헐적 고장, 부정확한 진단, 조직 문화적 요인 등이 포함된다. 특히 간헐적 고장은 군용기에서 전체 고장의 50%에 달할 수 있다고 보고된다(Haiyu Qi, 2008).

2.1.2 Weibull 분포와 생존 분석

Weibull 분포는 부품의 고장 시간 분포를 모델링하는 데 널리 사용된다. 형상 모수(β)는 고장 패턴을 결정하며, $\beta < 1$ 은 초기 고장, $\beta = 1$ 은 우발 고장, $\beta > 1$ 은 마모 고장을 의미한다. 본 연구에서는 2-parameter(2P)와 3-parameter(3P) Weibull 모델을 비교하며, 최대 우도추정법(MLE)을 통해 파라미터를 추정한다. 검열 데이터(censored data)를 포함한 생존 분석을 적용하여 NFF 발생확률을 예측한다.

2.1.3 베이저안 네트워크 모델

베이저안 네트워크(BN)는 변수 간의 조건부 독립성과 확률적 관계를 방향성 비순환 그래프(DAG) 형태로 표현하는 확률론적 모델이다. MCMC(Markov Chain Monte Carlo) 샘플링을 적용하여 사후 확률 분포를 추정하며, 이를 통해 NFF 발생에 영향을 미치는 다양한 변수 간의 인과 관계를 분석한다. 모델 성능은 AUC, 정확도, 정밀도, 재현율, F1 score 등으로 평가한다.

2.1.4 NFF 관리 프로세스

NFF는 단순히 부품의 장탈 여부만으로 판단되는 것이 아니라, 장탈 이후 시험 및 수리 결과를 기반으로 분류된다.

일반적인 NFF 관리 프로세스는 결함 탐지(fault detection), 부품 장탈(removal), 수리기관 점검(shop finding), NFF 판정, 원인 분석(root cause analysis) 단계로 구성된다.

ARINC(2008)는 이러한 절차를 통해 불필요한 부품 교체를 감소시키고 신뢰성 데이터를 개선할 수 있다고 제시하였다.

2.2 연구 방법론

2.2.1 연구 대상 및 데이터 수집

본 연구는 K항공사의 2011년부터 현재까지 누적된 LRU 부품 데이터를 활용하였다. NFF 비율이 높고 간헐적 고장 경향이 뚜렷한 3가지 부품을 선정하였다: TCAS computer(공중충돌방지장치), EIVMU(engine interface/vibration monitoring unit), SPB(seat power box). 데이터는 부품 정비 기록(shop finding report)을 기반으로 수집되었으며, 이상치 제거 및 전처리 과정을 거쳤다. Table 1은 연구 대상의 주요 특성을 나타낸다.

첫 번째 순서인 데이터 기록은 세부적인 결함 정보(PFR/TSD, crew de-briefing 등), 결함 해소를 위한 정비 작업 과정, 수리업체의 결함 분석 내용(shop finding) 등을 데이터 관리 프로그램을 통해 별도로 운영하는 것이다. 두 번째 순서는 앞서 수집 및 기록된 데이터와 정비 매뉴얼(AMM/FIM/TSM/CMM/FTD 등)을 기반으로 실제 결함 사항과 정비 과정 및 수리 내역의 상관관계를 검토하는 것이다. 이를 통해 NFF가

Table 1 .Characteristics of sample LRU components

항목	TCAS computer	EIVMU	SPB
Weibull 데이터	56개	101개	420개
고장 데이터	32개	43개	356개
검열 데이터	24개	58개	64개
BN 데이터 수	285개	520개	2,100개
NFF Rate	0.4286	0.5742	0.1524
특정 TST(hr)	3,777	29,487	32,673

적절한지 또는 부적절한지를 판단 내릴 수 있고, NFF가 부적절한 경우 수리업체에 기술 분석을 기반으로 한 시정조치를 요구할 수 있다. 마지막 순서인 시정조치는 세 가지로 분류된다. 수리업체에 상세 조사(detail investigation)를 요청한다.

즉, 단순 test(BITE, ATE 등)만 수행하여 선부는 NFF 판정으로 결론을 내리지 않고 결함의 근본 원인(root cause)이 발견될 때까지 상위 레벨의 세부적인 test(ESS test 등) 수행을 요청한다. 결함이 발견될 경우, 수리 조치가 가능하다. 다른 경우는 결함 재발 가능성이 있는 주요 부분품을 선제적으로 예방 수리하는 것이다. 기존 선례를 기술 분석 및 예측하여 결함 원인을 일으킬 수 있는 주요 부분품을 수리한다. 마지막은 NFF 재발이 강력히 의심되는 부품을 교환(exchange)한다. 이러한 프로세스를 거쳐 최종적으로 결함 장탈 부품의 NFF 비율을 줄일 수 있다. 이러한 프로세스를 Fig. 1과 같이 알고리즘 순서도로 도식화하였다.

데이터 수집 시 아래 Fig. 2의 로직을 참고해야 한다. 부품이 항공기로부터 제거된 후부터 수리 Shop 정비 기록(shop finding report)이 나오기까지 일련의 과정을 기준에 따라 분류하였다. 이를 기반으로 부품의 사용 시간, 계획 또는 비 계획 장탈 사유, NFF 여부 등 여러 정보 각 기준에 따라 매칭하여 데이터화 할 수 있다. 이렇게 수집된 데이터 중에서 고장데이터(failure data)와 관측데이터(censored data)는 Weibull 분석에 필요한 입력값이 된다.

2.2.2 Weibull 분석 모델 설계

Weibull 분포는 항공기 부품의 고장 특성을 분석하는 데 가장 널리 사용되는 신뢰성 분석 기법 중 하나로

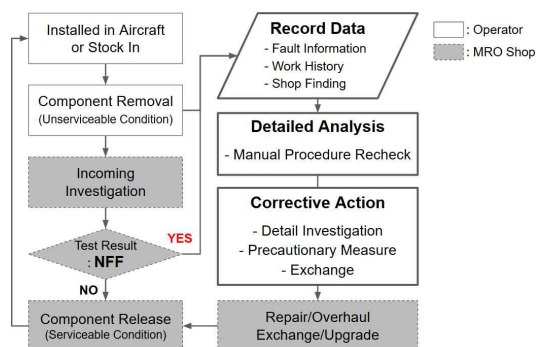


Fig. 1. NFF data process

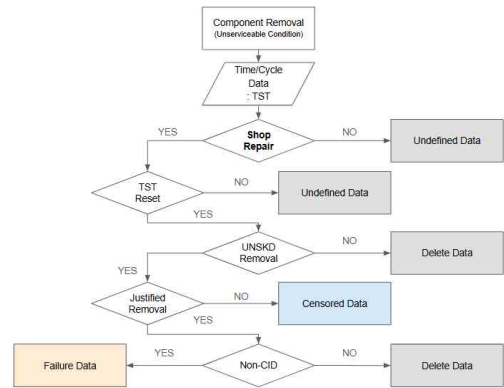


Fig. 2. Classification algorithms for data collection

Weibull 분포의 누적 분포 함수(cumulative distribution function, CDF)는 다음과 같이 정의된다(Weibull, 1951).

$$F(t) = 1 - e^{-\left(\frac{t}{\eta}\right)^\beta}$$

이 식은 특정 시점까지 부품이 고장 날 확률을 나타낸다. 여기서 β 는 형상 모수(shape parameter), η 는 척도 모수(scale parameter)이다. γ 는 위치 파라미터(location parameter)로 일반적으로 $\gamma=0$ 으로 설정되어 생략되며, 이 경우 분포는 0부터 정의된다.

특히 형상모수(β)는 크기에 따라 고장 발생 패턴을 결정하는 고장률의 형태를 결정한다. $\beta < 1$ 인 경우 초기 고장(decresing failure rate(DFR), infant mortality failure), $\beta = 1$ 이면 무작위고장(constant failure rate(CFR), random failure), $\beta > 1$ 이면 마모고장(increasing failure rate(IFR), wear-out failure)으로 구분할 수 있어 항공기 LRU의 고장 특성을 직관적으로 해석할 수 있다. 또한 우중도절단(censored data)을 포함한 운용 데이터를 효과적으로 처리할 수 있다는 장점이 있다.

MLE 기법이 적용된 Weibull 시뮬레이션을 통해 형상 모수(β)와 척도 모수(η)를 추정하고, 특정 TST 조건에서 NFF 발생확률을 추정한다. 위치 모수를 0으로 가정한 기본 2-parameter 타입과 위치 모수가 추가된 3-parameter 타입의 Weibull 분석을 비교하여 최적의 모델을 찾는다. 그리고 형상 모수의 형태와 해당 부품의 NFF 비율 대비 특정 시점에서의 NFF 발생확률을 비교 분석하여 NFF 적합성 여부와 원인에 대해 추정할 수 있다. 이를 통해 NFF를 허용할지 또는 정비 조치를

지시할 수 있는 의사 결정을 할 수 있게 해준다.

2.2.3 베이지안 네트워크 모델 설계

Bayesian network는 변수 간 인과관계를 확률적으로 모델링할 수 있는 대표적인 Probabilistic Graphical Model이다(Koller and Friedman, 2009). 이 모델을 통해 복잡한 시스템 내의 인과관계를 명시적으로 표현하고, 예측 정확도를 향상시킨다.

베이지안 네트워크 모델의 경우 5가지 변수가 사용되며 수집된 데이터는 결측치 처리, 이상치 처리, 범주형 변수의 인코딩 등 전처리 과정을 거친다. 베이지안 네트워크 모델 학습을 위해 각 변수 노드에 해당되는 데이터를 수집하여 특성 스케일링이 수행된다.

고장 발생 시간과 관련된 시간 변수는 부품의 TST 값을 그대로 가져온다. 장탈 사유는 자율형 정성적 기술 형식으로 정해진 양식은 없으나, Detailed Analysis 단계에서 공통적인 결함 양상을 파악하고 군집화하여 텍스트 마이닝으로 장탈 사유의 주요 키워드를 추출한다. 단, 부품 자체의 고장 양상이 아닌 외부적인 요인이 개입된 결함은 제외한다. NFF 이력은 부품이 Shop 테스트 진단 결과 NFF면 'Y'이고 결함이 발견되어 수리되었다면 'N'으로 입력한다. 해당 연구 모델에서 입력값은 Y는 1, N은 0으로 이진형(binary) 변수 형식으로 변경되어 적용된다. 장탈 항공기 기번은 결함이 확인되어 장탈된 항공기의 등록번호를 입력한다. 그리고 반복 장탈 여부는 부품의 일련번호(serial number)를 통해 결함이 처음으로 발생된 사례인지 또는 2회 이상 반복적으로 결함이 확인되어 장탈된 사례인지를 구분한다. 반복 고장일 경우에는 'Y'로, 처음 고장일 경우에는 'N'으로 입력한다.

물론 Weibull 분석과 마찬가지로 해당 데이터들은 전처리를 해야 한다. TST가 '0'인 이상치는 제외하고 동일한 수리업체가 조건이어야 한다. 독립적인 조건부 변수 데이터 외의 다른 영향을 미치는 요인이 있어서는 안 된다. 이렇게 확보된 데이터를 베이지안 네트워크 모델의 최종 입력값으로 설정한다. Fig. 3은 베이지안 네트워크 입력값의 데이터 전처리 후 셋팅된 샘플이다.

본 연구의 베이지안 네트워크 모델 설계에 적용되는 변수 노드는 사용 시간 TST, NFF 이력, 장탈 사유(removal reason), 반복 장탈(repeated removal), 장탈 기번(aircraft ID)이며 NFF에 영향을 줄 수 있

TST	REMOVAL REASON	NFF	AIRCRAFT ID	REPEATED REMOVAL
17159 EIU		Y	HL77XX	N
37448 EIVMU		Y	HL75XX	N
16962 VIB		N	HL75XX	N
6284 FADEC		Y	HL75XX	N
2266 EIU		Y	HL75XX	N
41287 VIB		N	HL75XX	N
21647 VIB		N	HL75XX	N
41642 VIB		N	HL75XX	N
4388 VIB		Y	HL75XX	N
25685 EIVMU		Y	HL77XX	N
58 VIB		N	HL77XX	Y
34449 VIB		Y	HL77XX	N
2124 VIB		Y	HL75XX	Y
7452 ENG		Y	HL82XX	N
721 VIB		N	HL82XX	Y
3600 EIVMU		N	HL75XX	Y
52058 VIB		Y	HL75XX	N
1522 ENG		Y	HL75XX	Y
2718 VIB		Y	HL82XX	Y
91629 FADEC		Y	HL77XX	N

Fig. 3. Variable data input for Bayesian network

는 각 독립적인 변수의 상호 인과 관계를 고려하여 아래와 같이 베이지안 네트워크 구조를 설계하였다. 이를 DAG(방향성 비순환 그래프) 형태로 나타내면 다음과 같이 구성될 수 있다. 이러한 독립적 변수들은 서로 연관되어 있으며 베이지안 네트워크를 통해 변수 간의 확률적 의존 관계를 모델링할 수 있다. 아래 Fig. 4는 본 연구의 각 노드별 상관관계에 따른 베이지안 네트워크 구조 형태이다.

본 예측 모델의 기본 구조를 바탕으로 장탈 사유(removal reason), 사용 시간(TST), 반복 장탈 여부(repeated removal), 장탈 항공기 기번(aircraft ID)의 데이터를 기반으로 NFF 확률을 구하는 식은 아래와 같다.

$$P(NFF|정보) = \frac{P(정보|NFF) \cdot P(NFF)}{P(정보)} \quad (1)$$

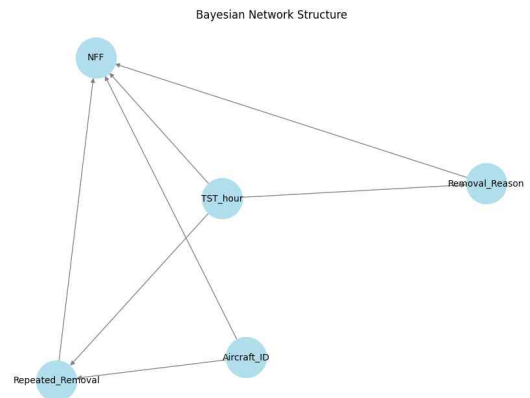


Fig. 4. Bayesian network structure of NFF prediction

여기서 '정보'는 removal reason, TST, repeated removal, aircraft ID의 데이터를 내포한다. MLE 기법을 적용하여 각 변수 간의 관계를 학습하고 조건부 확률을 추정한다. 이를 바탕으로 NFF 발생 가능성을 계산할 수 있다. 특히 각 변수의 특정 조건에서 NFF 발생확률을 추정할 수도 있다.

본 연구에서의 모델을 적용하여 불확실성(uncertainty) 성질의 NFF 발생확률을 정량화하고 예측할 수 있어 정비 의사결정의 신뢰성을 향상시킨다.

따라서 단일 변수 기반 분석보다 다변량 확률모형이 적합하며, Bayesian network는 이러한 복합 요인의 영향을 설명할 수 있는 장점을 가진다.

베이지안 네트워크 모델에는 사용 시간(TST), 장탈 사유(removal reason), NFF 이력, 반복 장탈(repeated removal), 장탈 항공기 기번(aircraft ID) 총 5개의 변수를 사용한다. 각 변수는 DAG 형태로 인과관계 네트워크를 구성하며, MCMC 샘플링(5,000, 10,000회)을 적용하여 조건부 확률을 추정한다. 학습/테스트 데이터를 7:3으로 분할하여 모델 성능을 검증한다.

2.3 연구 결과 및 분석

2.3.1 Weibull 분석 결과

본 연구는 각 부품별 Weibull 분석을 통해 고장 시간 데이터의 분포를 파악하고, NFF 확률을 추정하였다. 우선 실험 대상 부품의 데이터 수, NFF 비율, 조건부 특정 시간을 파악하였다.

TCAS computer의 경우 고장데이터는 32개, 관측 데이터는 24개이며 2-parameter와 3-parameter Weibull CDF (신뢰도 95%) 분석한 결과, 적합도 평가는 Log-likelihood, AIC(Akaike information criterion), BIC(Bayesian information criterion), AD(Anderson-Darling) 수치로 비교 분석하였다. AIC와 BIC는 값이 작을수록 성능이 더 좋은 모델을 의미하고 AD 통계량은 1.0 이하일 때 대체로 양호한 적합도로 판단한다. 아래 Table 2는 두 가지 결과값에 대한 자료이다.

2P의 경우 모델의 적합도는 적절한 수준으로 실제 데이터에 잘 맞는 분포로 해석된다. 3P와 비교하였을 때 2P 모델의 Log-likelihood가 조금 더 작고 AIC 및 BIC가 약간 높지만, AD 값은 1.0 이하이며 K-S Test

Table 2. TCAS computer 2P and 3P Weibull analysis verification indicator

항목	2P Weibull	3P Weibull
Log-likelihood	-329.53	-326.01
AIC	663.43	658.81
BIC	666.10	662.59
AD	0.92	1.53
K-S test statistic	0.1368	0.1854
K-S Test P-value	0.5416	0.1953
형상모수(β)	1.0627	0.731455
95% CI(lower~upper)	0.80589~1.40141	0.54269~0.98587
특정시간 NFF 발생 확률	0.7287	0.6024

P-value 값이 더 높으므로 2P 모델이 더 적합하다. 그러나 아래 Weibull 누적분포함수(CDF) 그래프인 Fig. 5에서 3P 모델의 점들이 추세선에 더 밀집 되어있고, 신뢰구간을 의미하는 파란 구간이 더 좁고 일관된 형태이므로 데이터 적합도는 3P가 더 우수하다. 다시 말해서, 성능 면에서는 2P 모델이 더 우수하나, 데이터 적합도에서는 3P가 약간 더 우수한 편이다.

본 연구에서는 어떤 모델이 더 우수한지 판가름하여 적용하는 것이 아니라 2P와 3P 모델을 비교하여 어떠한 공통성과 패턴이 있는지 확인하여 형상 모수와 NFF 예측 확률의 수치를 근거로 하여 NFF 판정의 적합성 여부에 대해 의사결정을 내리는 것이 중요하다.

TCAS computer의 경우 2P Weibull 기준 형상 모수 β 는 1.0627로 1에 가까워 우발고장 특성을 보였고, 3P Weibull 기준 β 는 0.731455로 초기고장 특성을 보였다. 2P 모델은 K-S P-value가 0.5416으로 양호하고 AD가 0.92로 1 이하였으며, 특정 TST 3,777시간에서 NFF 발생확률은 0.7287로 추정되었다. 이는 대상 부품군의 실제 NFF rate 0.4286보다 높아 해당 사례의 NFF 판정 가능성이 높은 것으로 해석된다.

Weibull 분석 결과를 토대로 실제 3,777시간 사용된 해당 부품은 수리업체의 test 판정 결과 NFF로 입고하여 운영상 문제가 없음을 확인하였기 때문에 해당 모델의 예측 성능은 적합하다고 판단된다.

EIVMU 경우, 두 모델 간 Log-likelihood 수치는 같으며 AIC 및 BIC가 3P 모델이 더 높으므로 2P 모델

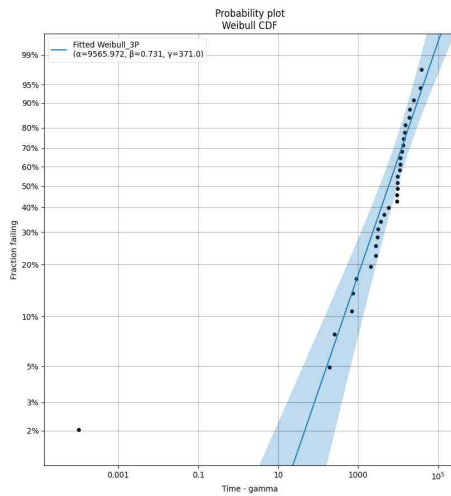
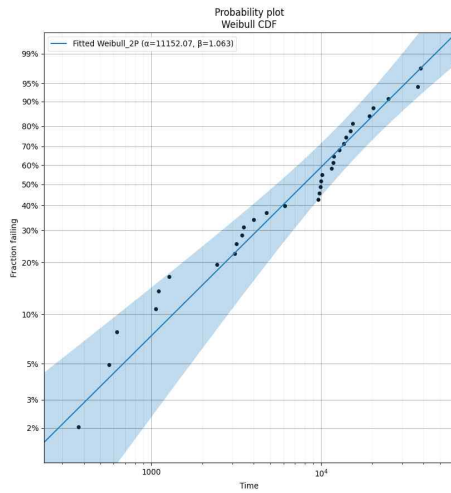


Fig. 5. TCAS computer 2P and 3P Weibull cumulative distribution function

이 더 적합하다고 할 수 있다. 그러나 AD 수치는 2.6으로 1보다 크므로 Weibull 분포가 완벽하게 적합하지 않을 수 있지만 K-S 검정 P -value는 양호한 0.05보다 큰 양호한 수치로 모델을 활용하기에 문제는 없다. 아래 Table 3은 두 가지 결과값에 대한 자료이다.

또한, 아래 Fig. 6 Weibull 누적분포함수(CDF) 그래프 형태는 일치한다. 그리고 두 모델의 점들은 신뢰구간 내에 위치하며, 추세선과 비교 시 비교적 양호한 적합을 보인다. 그러나 좌측 하단의 신뢰구간이 다소 넓게 퍼져 있어, 중간 이상에서 분산이 커지는 것을 볼 수 있다. 이는 두 모델 모두 어느 정도 데이터를 잘 설명하고 있지만 AD 값 및 신뢰구간 폭을 고려하면 완벽한 적합은 아닐 수 있다.

Table 3. EIVMU 2P and 3P Weibull analysis verification indicator

항목	2P Weibull	3P Weibull
Log-likelihood	-480.658	-480.658
AIC	965.546	967.786
BIC	969.33	973.337
AD	2.61666	2.61666
K-S test statistic	0.1647	0.1647
K-S test P -value	0.1733	0.1733
형상모수(β)	0.561189	0.561189
95% CI(lower~upper)	0.42988~0.73259	0.42988~0.73259
특정시간 NFF 발생 확률	0.3836	0.3836

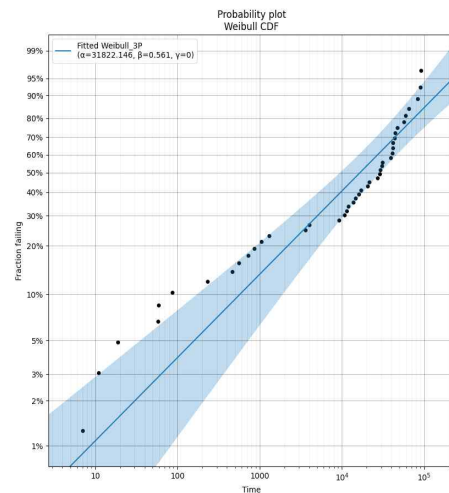
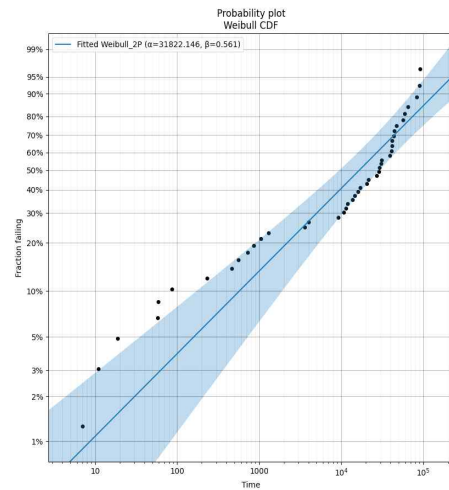


Fig. 6. EIVMU 2P and 3P Weibull cumulative distribution function

결론적으로, K-S 검정 P -value는 0.05보다 크고 적절한 신뢰구간(CI)으로 통계적으로 모델 활용은 가능하나, AD 값이 2.6 이상으로 매우 높으므로 분포 적합도가 낮다고 볼 수 있다. 그러나 이는 좌측 하단 500시간 미만 데이터 부분의 적합도가 다소 떨어지지만, 본 연구에서 관측하고자 하는 데이터와는 다소 이격이 있어 해당 모델을 활용하기에는 큰 문제는 없다.

이러한 데이터 적합도에 대해 일부 신뢰도는 떨어질 수 있으나, 여전히 Weibull 분석의 형상 모수 해석이 가능하다. 형상 모수는 1보다 적은 초기고장 형태로 운영상 문제가 있거나 부품의 수리 품질에 영향이 있을 수 있다.

특정 시간인 TST 29,487시간일 때 NFF 발생확률이 38.4%로 예측되었다. 이는 해당 부품군의 NFF 비율이 57.4%임을 고려하였을 때, 실험 대상인 특정 부품의 NFF 발생확률이 낮게 예측되었고 초기고장 형태를 따르는 점을 판단하였을 때 NFF 판정이 부적합할 가능성이 크다.

한편 SPB 분석 결과, 두 모델 간 Log-likelihood는 동일하므로 데이터에 대한 적합도는 같다. AIC 및 BIC는 3P 모델이 더 높으므로 2P 모델이 통계적으로 더 우수하다. AD 값은 2.9756으로 1보다 크므로 일반적으로 높은 값에 해당되어 Weibull 모델이 완벽하게 맞지 않을 수 있지만 K-S 검정 P -value는 0.8888로 0.05보다 큰 양호한 수치로 적합도 통계적으로 모델 수용이 가능하다. 그리고 3P 모델의 Gamma(위치 모수)가 0으로 최적화되어 사실상 2P와 동일한 형태로 수렴하였다. 아래 Table 4는 두 가지 모델의 결과값에 대한 자료이다.

여기서 형상 모수는 1.76으로 1보다 큰 마모 고장 형태이다. 이는 부품의 노후화(aging)에 따른 고장이 발생되는 추세로 확인된다. 또한 Fig. 7의 2P 및 3P 모델 누적분포함수의 그래프 형태를 보면 신뢰구간 범위는 좁고, 데이터를 의미하는 점의 분포는 실선과 거의 일치하므로 모델의 신뢰도와 데이터 적합도는 우수하다(Fig. 7).

즉, EIVMU와 마찬가지로 K-S 검정 P -value는 0.05보다 크고 안정적인 신뢰구간(CI)으로 통계적으로 모델 활용은 가능하여 형상 모수에 대한 해석은 가능하다. 그러나 AD 값이 1 이상인 2.98로 매우 높으므로 분포 적합도가 낮다고 부적절하다고 볼 수 있다. 그러나 이는 좌측 하단 1,000시간 미만 데이터 부분의

Table 4. SPB 2P and 3P Weibull analysis verification indicator

항목	2P Weibull	3P Weibull
Log-likelihood	-480.658	-480.658
AIC	965.546	967.786
BIC	969.33	973.337
AD	2.61666	2.61666
K-S test statistic	0.1647	0.1647
K-S test P-value	0.1733	0.1733
형상모수(β)	0.561189	0.561189
95% CI(lower~upper)	0.42988~0.73259	0.42988~0.73259
특정시간 NFF 발생 확률	0.3836	0.3836

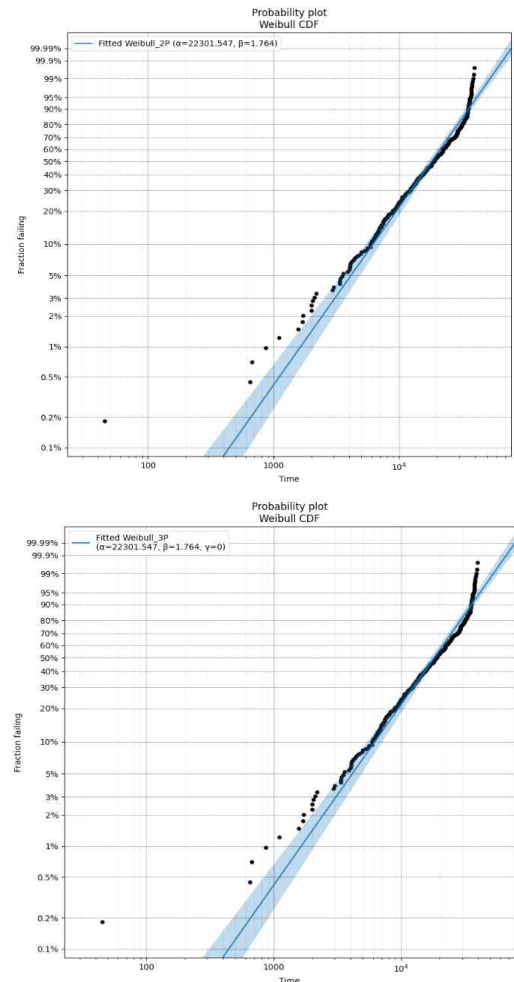


Fig. 7. SPB 2P and 3P Weibull cumulative distribution function

적합도가 다소 떨어지지만, 본 연구에서 관측하고자 하는 특정 시간과는 다소 이격이 있어 해당 모델은 여전히 유효하다. 따라서 해당 Weibull 분포로 형상 모수 해석이 가능하므로 NFF 예측 확률과 함께 NFF의 발생 원인을 탐구할 수 있다.

2.3.2 베이지안 네트워크 모델 결과

본 연구에서는 TST, 장탈 사유, NFF 이력, 장탈 항공기 기번, 반복 장탈 여부를 주요 변수로 설정하고, 변수 간 조건부 의존관계를 베이지안 네트워크로 구성하여 LRU별 NFF 발생확률을 추정하였다. MCMC 샘플링은 5,000회와 10,000회로 반복 수행하였으며, 두 샘플 크기에서 유사한 추정값이 도출되는지를 함께 검토함으로써 추론 결과의 안정성을 확인하였다(Table 5).

연구 결과, 4번 조건인 특정 항공기와 TCAS computer 사용 시간 3,777시간에 따른 조건에서 다른 조건들에 비해 가장 높게 약 73~74% 정도 NFF 발생확률을 예측하였다. 반대로 2번 조건의 장탈 사유인 특정 결함과 반복 장탈 조건일 경우, NFF 확률이 가장 낮다는 것을 알 수 있다. 이는 고장 내용과 반복 장탈의 단서만으로 예측력이 떨어지고, NFF 예측을 분석할 때 제한적이다. 그 외 나머지 조건들도 거의 NFF 예측 확률이 50%대로 중간 경계선에 모호한 값이므로 NFF 발생확률에 가장 큰 요인으로 보기는 어렵다. 따라서 해당 조건들은 참고 보조 지표로 활용해야 할 것이다.

Table 5. TCAS computer NFF occurrence probability for each condition

조건	Sample size	NFF 확률	CI	
			Lower	Upper
1. 장탈사유 & 사용시간	5,000	0.5236	0.509756	0.537444
	10,000	0.5332	0.523422	0.542978
2. 장탈사유 & 반복장탈	5,000	0.5050	0.491141	0.518859
	10,000	0.4953	0.485500	0.505100
3. 반복장탈 & 사용시간	5,000	0.5204	0.506552	0.534248
	10,000	0.5218	0.512009	0.531591
4. 항공기기번 & 사용시간	5,000	0.7404	0.728248	0.752552
	10,000	0.7329	0.724228	0.741572
5. 장탈사유 & 항공기기번	5,000	0.5228	0.508995	0.536645
	10,000	0.5246	0.514812	0.534388

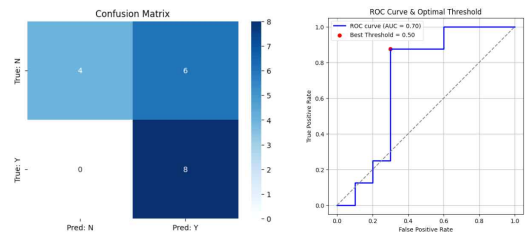


Fig. 8. TCAS computer confusion matrix and ROC curve for NFF prediction model

다음은 해당 모델의 성능도를 평가하고 검증하였다. 검증 방식은 NFF의 사후확률에 대한 샘플링을 기반으로 동일한 조건에 대해 여러 번 반복 예측하고 평균 확률을 사용하는 방식을 적용하여 검증에 대한 예측을 안정화하도록 하였다. 충분한 샘플 수로 반복 예측하고, 표준오차 2% 이내 수렴 조건을 만족하는 결과만 사용하였다.

Fig. 8의 혼동행렬과 ROC 곡선의 AUC 지표를 확인할 수 있었다. 총 18건의 평가 샘플 중 12건을 정확히 분류하였으며, 전체 분류 정확도(accuracy)는 약 66.7%로 매우 높은 편은 아니지만 이 모델이 클래스가 불균형한 상황에서도 일정 수준 이상의 예측력을 가지고 있다는 것을 의미한다.

혼동행렬(confusion matrix)에서 고장을 의미하는 클래스 'N'을 정확히 'N'으로 예측한 true negative (TN)은 4건, 고장 'N'을 NFF를 의미하는 클래스 'Y'로 오 예측한 false positive(FP)는 6건이다. 그러나 'Y'를 'N'으로 오 예측한 false negative(FN)는 0건, 'Y'를 정확히 'Y'로 예측한 true positive(TP)는 8건이다. 이는 실제 NFF를 모두 예측한 것이다.

한편, ROC 곡선에서 볼 수 있듯이 최적의 threshold는 0.5로 기준과 일치한다. ROC 곡선은 대체로 직선인 0.5 이상의 면적에 위치하며, AUC가 0.70으로 확인되었다. 이는 모델이 'Y' 클래스와 'N' 클래스를 잘 구분할 수 있는 성능을 가지고 있음을 의미한다.

다른 LRU인 EIVMU는 'Engine Fault' 결함과 함께 SOS 격리 대상으로 선정되어 장탈되었으며, 장탈 되기까지의 사용 시간 TST는 29,487시간이다. 또한 과거에도 반복적으로 장탈된 이력이 있는 부품이다.

이런 특정 조건을 바탕으로 베이지안 네트워크 분석 결과 Table 6과 같이 각 조건 별 NFF 발생확률과 95% 신뢰구간의 지표가 확인되었다.

예측 지표 결과, NFF 발생에 가장 영향력이 큰 조

Table 6. EIVMU NFF occurrence probability for each condition

조건	Sample size	NFF 확률	CI	
			Lower	Upper
1. 장탈사유 & 사용시간	5,000	0.5344	0.520574	0.548226
	10,000	0.5232	0.513411	0.532989
2. 장탈사유 & 반복장탈	5,000	0.4962	0.482341	0.510059
	10,000	0.4974	0.487600	0.507200
3. 반복장탈 & 사용시간	5,000	0.5182	0.504350	0.532050
	10,000	0.5080	0.498201	0.517799
4. 항공기기변 & 사용시간	5,000	0.7384	0.726218	0.750582
	10,000	0.7363	0.727663	0.744937
5. 장탈사유 & 항공기기변	5,000	0.5014	0.487541	0.515259
	10,000	0.5105	0.500702	0.520298

건부 변수는 특정 항공기 기변과 해당 부품의 사용 시간이다. 즉, 특정 기변에서 사용 시간이 약 30,000시간 내외의 조건일 경우 NFF 발생확률이 약 73%로 NFF 예측에 가장 유의미한 변수가 된다. 반면에 2, 3, 5번 조건은 임계값인 50%의 모호한 경계면에 있으므로 NFF 예측 분석에 활용하기에는 부적절하다.

조건 4번(항공기 기변과 부품 사용 시간 TST)에서 두 샘플 모두 가장 높은 NFF 확률을 나타내는데 10000 샘플(주황색)이 약간 더 높은 것을 알 수 있다. 또한 1번 조건인 장탈 사유와 부품 사용 시간에 대해서 NFF 발생확률이 52~53%로, 신뢰구간이 임계점보다 우위에 있으므로 NFF 발생원인 분석에 적합하다. 따라서 본 연구 결과, 해당 부품에 대해 NFF 예측 시 특정 기변과 사용 시간이 조건일 때 가장 효과적이고 적합하며, 장탈 사유인 특정 결함과 부품의 사용 시간도 NFF에 영향을 주는 조건이 된다.

본 모델의 검증 평가를 수행할 때 threshold는 0.5로 기준을 설정하여 수행하였다. 모델의 예측 정확도는 62.5%로 중간 수준이다. NFF를 의미하는 'Y' 예측 정밀도는 71.4%로, 양호한 성능을 나타낸다. 그러나 재현율은 55.6%로 낮은 편에 속한다. 이는 NFF 예측 성능이 좋지 않은 것을 의미한다. 그럼에도 Y 클래스에 대한 정밀도와 재현율의 조화 평균인 F1 score는 62.5%로 균형 잡힌 성능을 보였다. Balanced accuracy는 63.5%로 클래스 간 예측 편향이 심하지 않음을 나타낸다.

다. 이는 클래스 불균형 상황에서의 공정성 있는 성능 평가를 제공한다. 또한 모델이 확률 기반 예측에서 얼마나 확실한지를 측정하는 로그 손실의 경우, 69.2%로 중간 수준의 예측 성능 신뢰도를 나타낸다.

본 연구 모델은 threshold 0.5를 NFF 예측 임계 기준으로 설정하고 검증 지표를 통해 활용이 가능한 모델임을 확인하였으나, NFF 예측 성능이 중요하므로 최적의 threshold 0.496으로 재설정하여 모델을 개선하였다. 특히 개선된 모델 재현율이 향상되어 NFF 예측에 좋은 성능을 나타낸다(Table 7).

본 연구 모델은 샘플 5000, 10000에서 각 0.5와 0.496 임계 기준값이 NFF 예측 신뢰구간의 모호한 경계에 위치하여 2, 3, 5번 조건은 NFF 예측 증거 자료로 활용하기 어렵다. 따라서 1번 및 4번 조건에서만 확실하게 NFF 발생 조건을 판단할 수 있다. 반면 임계값 0.496으로 설정하였을 때 신뢰도가 증가하여 1, 3, 4번 조건의 신뢰구간이 threshold보다 위에 있으므로 NFF 발생 원인에 대한 추가 단서를 적용할 수 있다. 전반적으로 NFF 예측 확률이 과반수 초과로 NFF 발생이 높은 부품이다.

세 번째 샘플 LRU인 SPB는 장탈 사유이자 결함 양상이 전부 일관성 있는 고장유형이며, 32,673시간 만에 장탈되었다. 과거에 반복 장탈된 이력이 없는 첫 고장 사례로 NFF 적합성 여부를 분석한 결과 Table 8의 NFF 예측 확률 및 신뢰구간을 확인하였다.

분석 결과, 특정 항공기에서의 사용 시간 TST 32,673 조건일 때 NFF 확률이 약 99.9%로 매우 높

Table 7. Comparison of performance metrics for Bayesian network models of EIVMU

지표	Threshold 0.5		Threshold 0.496	
	TN=10	FP=4	TN=9	FP=5
Confusion matrix	FN=8	TP=10	FN=4	TP=14
AUC	0.643		0.691	
Accuracy	0.625		0.719	
Precision	0.714		0.737	
Recall	0.556		0.778	
F1 score	0.625		0.752	
Balanced accuracy	0.635		0.710	
Log loss	0.692		0.606	

Table 8. SPB NFF occurrence probability for each condition

조건	Sample size	NFF 확률	CI	
			Lower	Upper
1. 장탈사유 & 사용시간	5,000	0.3302	0.317164	0.343236
	10,000	0.3138	0.304705	0.322895
2. 장탈사유 & 반복장탈	5,000	0.3188	0.305883	0.331717
	10,000	0.3215	0.312346	0.330654
3. 반복장탈 & 사용시간	5,000	0.3110	0.298169	0.323831
	10,000	0.3153	0.306193	0.324407
4. 항공기기변 & 사용시간	5,000	0.9986	0.997564	0.999636
	10,000	0.9985	0.997741	0.999259
5. 장탈사유 & 항공기기변	5,000	0.4958	0.481941	0.509659
	10,000	0.4955	0.485700	0.505300

다. 그러나 나머지 조건에서는 NFF일 확률이 과반수 미만이며 특히 일부는 NFF 예측 확률이 30%대로 매우 낮으므로 NFF가 아닐 가능성이 크다. 따라서 4번 조건을 제외한 나머지는 NFF 발생에 인과관계가 적은 인자이다(Table 8).

전반적으로 성능이 저조한 모델을 개선하고자 최적의 threshold 0.4896을 재설정 후 분석하였다. Table 9. 처럼 기존 threshold 0.5와 비교하였을 때, AUC는 유사하며, 정확도는 다소 줄어들었다. 전반적인 예측 신뢰도는 기존 모델이 조금 더 나을 수 있으나 정밀도,

Table 9. Comparison of performance metrics for Bayesian network models of SPB

지표	Threshold 0.5		Threshold 0.4896	
	TN=66 FN=11	FP=41 TP=8	TN=38 FN=1	FP=69 TP=18
AUC	0.6316		0.6264	
Accuracy	0.587		0.444	
Precision	0.151		0.207	
Recall	0.421		0.947	
F1 score	0.222		0.339	
Balanced accuracy	0.565		0.656	
Log loss	0.699		0.689	

재현율, F1 score, balanced accuracy는 모두 향상되었다. 특히 재현율이 대폭 상승하여 NFF 예측에 아주 좋은 성능을 나타내고 있다. 또한 로그 손실은 기존보다 다소 줄어들어 예측 신뢰도가 약간 개선되었다고 볼 수 있다(Table 9).

지표 해석 결과, 앞의 두 샘플 부품의 모델에 비해 SPB 분석 모델의 전반적인 성능 지표가 낮은 편이나, Threshold를 조정하였을 때 재현율이 대폭 상승하여 NFF 예측 모델로 활용은 가능할 수 있다. 그러나 고장을 NFF로 오답하는 경향이 높으므로 오경보(false alarm) 위험이 큰 거짓 양성률이 너무 많다. 또한 SPB의 NFF 비율이 낮은 편으로, 이는 클래스의 불균형을 초래한다. 그리고 신뢰구간이 좁은 NFF 예측 확률 99.9%에 비해 로그 손실이 69%임을 고려하면 해당 모델은 틀린 예측을 할 가능성이 크다. 즉, 모델의 예측 확률의 신뢰도가 좋지 않다고 역설할 수 있다. 따라서 해당 모델을 실제로 활용하기에는 부적합하다고 판단된다.

2.4 종합 해석

Weibull 분석과 Bayesian network 분석을 종합하면 TCAS computer는 두 분석 모두에서 NFF 가능성이 높게 나타났다. EIVMU는 Weibull 단독 분석에서는 NFF 가능성이 낮게 산출되었으나, Bayesian network에서 특정 항공기 기변과 사용시간 조건이 강한 설명력을 보였고 threshold 조정 후 성능이 개선되어 NFF 가능성이 높은 대상으로 해석된다. 반면 SPB는 Weibull 분석에서 마모고장 특성이 뚜렷하고, Bayesian network에서는 클래스 불균형으로 인한 오답 위험이 높아 실제 고장 가능성이 높은 부품으로 판단된다(Table 10).

본 연구의 가장 큰 시사점은 NFF 문제를 단순 정비 경험이나 사후 검토에만 의존하지 않고, 부품 사용시간과 조건부 변수에 기반한 정량 지표로 관리할 수 있음을 보였다는 점이다. Weibull 분석은 시간 기반 고장 경향을 제시하고, Bayesian network는 특정 항공기 기변, 장탈사유, 반복장탈 여부와 같은 조건 조합의 영향을 해석할 수 있어 두 모델은 상호 보완적이다.

다만 모델의 성능은 데이터 구조와 클래스 균형에 민감하였다. SPB 사례처럼 NFF rate가 낮은 부품에서는 NFF 예측값이 과도하게 산출되거나 false positive가 증가할 수 있다. 따라서 본 모델은 자동 판정 시스템보다는

Table 10. Integrated decision result

항목	Weibull result	Bayesian network result	Final decision
TCAS	NFF probability high	Recall high, NFF likely	NFF likely
EIVMU	Initial failure pattern	Improved after threshold tuning	NFF likely, monitor repair quality
SPB	Aging failure pattern	Class imbalance and false alarm risk	Actual failure likely

정비사가 최종 판단을 수행하는 human-in-the-loop 방식의 보조 지표로 적용하는 것이 적절하다.

본 연구는 NFF 판정을 정비사의 경험에만 의존하던 기존 방식에서 벗어나 데이터 기반 의사결정 체계를 제시하였다는 점에서 실무적 의미를 가진다. 특히, SOS 격리, MRO 재시험 의뢰, 예비품 운영계획 수립 과정에서 우선 검토 지표로 활용될 수 있다.

III. 결 론

본 연구는 항공기 LRU 부품의 NFF 현상을 예측하기 위해 Weibull 분포 기반 생존 분석과 베이지안 네트워크 모델을 결합한 방법론을 제시하였다. K항공사의 실제 정비 데이터를 기반으로 TCAS Computer, EIVMU, SPB 3개 부품을 분석한 결과, TCAS computer와 EIVMU는 NFF 가능성이 높고, SPB는 실제 고장일 가능성이 높음을 예측하였으며 이는 실제 NFF 비율과 부합하였다.

Weibull 모델은 형상 모수를 통해 고장 패턴을 직관적으로 해석하고 NFF 적합성을 판정하는 데 효과적이었다. 베이지안 네트워크 모델은 다변수 조건부 확률로 NFF 발생 원인을 세부적으로 분석할 수 있으며, threshold 최적화를 통해 재현율 중심의 성능 개선이 가능하다. 두 모델을 결합한 Hybrid 접근법은 단순 통계 분석 대비 더 포괄적인 NFF 예측이 가능하다.

본 연구 모델을 실무에 적용하면 SOS 격리 대상 부품의 불필요한 정비 행위 및 부품 장탈을 사전에 방지할 수 있다. 단독 자동화 판단보다는 'NFF 예측 지표' 기반의 알람 시스템으로 활용하여 최종 판단은 정비사가 수행하는 hybrid human-AI 시스템이 적합하다.

향후 연구에서는 QAR 등 실시간 센서 데이터, MTBUR/MTBF 신뢰성 지표를 추가 변수로 활용하고, random forest, gradient boosting, DNN 등 머신러닝 기법과의 융합을 통해 모델의 일반화 능력을 향상시킬 예정이다.

References

1. James, I., Lombard, D., Willis, I., and Goble, J., "Investigating no fault found in the aerospace industry", Reliability and Maintainability Symposium, 2003, pp.441-446.
2. Khan, S., Phillips, P., Jennions, I., and Hockley, C., "No fault found events in maintenance engineering part 1: Current trends, implications and organizational practices", Reliability Engineering & System Safety, 123, 2014, pp.183-195.
3. Khan, S., Phillips, P., Jennions, I., and Hockley, C., "No fault found events in maintenance engineering part 2: Root causes, technical developments and future research", Reliability Engineering & System Safety, 123, 2014, pp.196-208.
4. Khan, S., Phillips, P., Jennions, I., and Hockley, C., "No fault found, retest ok, cannot duplicate or fault not found? towards a standardised taxonomy", 1st Annual Symposium on Tackling No Fault Found in Maintenance Engineering, Through-life Engineering Services, Cranfield University, 2013, pp.246-253.
5. ARINC Report 672, "Guidelines for the Reduction of No Fault Found (NFF)", ARINC, 2008.
6. Hansa, C., and Hribernik, K., "Nff special session: Potentials of applying methods, tools, processes and knowledge from testing in product development to the NFF problem", 2014.
7. Kim, M., and Ji, W. K., "The usefulness of hard time task for weapon system in

- considering shape parameter of Weibull life time distribution and maintenance cost", *Journal of the Korea Academia-Industrial Cooperation Society*, 17(1), 2016, pp.274-283.
8. Suciati, F. K., and Suwondo, E., "Evaluation and improvement of unscheduled removal component reliability calculation method with no fault found (NFF) shop finding results", *IOP Conference Series: Materials Science and Engineering*, 645, 2019, pp.012007.
9. Qi, H., "No-Fault-Found and Intermittent Failures in Electronic Products", 2008.
10. Park, Y. J., Fan, S. K. S., and Hsu, C. Y., "A review on fault detection and process diagnostics in industrial processes", 2020.
11. Singh, V. P., Ganguly, K., and Samad, T. A., "No Fault Found: A Systematic Literature Review and Future Research Agenda", 2021.
12. Polymenis, A., "Modelling Reliability: A Practical Application from Industry", 2017, pp.543-548.