

C-MAPSS 데이터를 이용한 항공기 엔진의 신경 회로망 기반  
건전성관리윤유리<sup>1</sup> · 김석구<sup>2</sup> · 조성희<sup>3</sup> · 최주호<sup>4,†</sup><sup>1</sup>현대건설기계 신뢰성 평가부<sup>2,3</sup>한국 항공대학교 항공 우주 및 기계공학과<sup>4</sup>한국 항공대학교 항공 우주 및 기계공학부Neural Network based Aircraft Engine Health Management using C-MAPSS  
DataYuri Yun<sup>1</sup>, Seokgoo Kim<sup>2</sup>, Seong Hee Cho<sup>3</sup> and Joo-Ho Choi<sup>4</sup>,<sup>1</sup>Hyundai Construction Equipment, Korea<sup>2,3</sup>Department of Aerospace and Mechanical Engineering, Korea Aerospace University, Korea<sup>4</sup>School of Aerospace and Mechanical Engineering, Korea Aerospace University, Korea

## Abstract

PHM (Prognostics and Health Management) of aircraft engines is applied to predict the remaining useful life before failure or the lifetime limit. There are two methods to establish a predictive model for this: The physics-based method and the data-driven method. The physics-based method is more accurate and requires less data, but its application is limited because there are few models available. In this study, the data-driven method is applied, in which a multi-layer perceptron based neural network algorithms is applied for the life prediction. The neural network is trained using the data sets virtually made by the C-MAPSS code developed by NASA. After training the model, it is applied to the test data sets, in which the confidence interval of the remaining useful life is predicted and validated by the actual value. The performance of proposed method is compared with previous studies, and the favorable accuracy is found.

## 초 록

항공기 엔진의 고장예지 및 건전성 관리(PHM)는 고장 또는 수명한계 도달 전에 잔존 유효 수명을 예측하는 것이다. PHM 기술 중 예측모델을 확립하는 방법은 물리 기반과 데이터 기반 방법이 있다. 물리 기반 방법은 적은 데이터로 정확한 예측이 가능하지만 확립된 손상 물리 모델이 적어서 적용에 한계가 있다. 본 연구는 따라서 데이터 기반 방법을 적용하였으며, 수명 예측을 위해서 신경회로망 알고리즘 중 Multi-layer Perceptron을 이용하였다. 이를 위해 미국 항공우주국(NASA)에서 개발한 C-MAPSS 코드로 생성된 가상 데이터 세트를 이용하여 신경회로망을 학습하였다. 학습된 신경회로망 모델은 테스트 세트에 적용한 후 잔존 유효 수명의 신뢰구간을 예측하고 실제 값을 통해 정확도를 검증하였다. 또한 본 연구에서 제시된 방법을 기존 문헌의 것과도 비교하였고 그 결과 비교적 양호한 정확도를 확인할 수 있었다.

**Key Words** : C-MAPSS dataset(C-MAPSS 데이터세트), Neural network(신경 회로망), PHM(고장 예지 및 건전성 관리)

## 1. 서 론

항공기 엔진은 항공기 안전에 직접적인 연관이 있다. 예상하지 못한 항공기 고장은 대재앙의 사고나 다운타임을 발생시킨다. 이러한 문제점을 방지하기 위하여 엔진의 정확한 잔존 유효 수명을 파악하는 것이 중

요하다. 센서 기술의 빠른 발전과 더불어, 멀티 센서를 이용한 모니터링으로 엔진 건전성 관리가 가능해졌다. 그 결과, 통합적이고 다양한 데이터를 위한 적절한 방법론을 개발하는 것이 잔존 유효 수명 예측의 정확성을 높이기 위하여 중요하다. 이를 위해 수명주기(lifecycle) 동안 성능을 보장하는 신뢰성 최적설계 기술이 최근 30년 동안 개발되었으나 설계단계에서 운용 중 발생하는 모든 상황을 고려하는 것은 불가능하기 때문에 한계가 있다.

따라서 최선의 방안은 운용 중 정비를 철저히 하여 파손을 예방하는 수밖에 없으며 이를 위한 현재의 방식은 주기적 예방정비(PM; preventive maintenance)이지만 잦은 정지 및 교체로 가동률 저하와 높은 비용이 문제가 되고 있다. 또한 이 방식은 Fig. 1과 같이 ‘Fleet Average’에 기반을 두어 정비주기를 설정하는데, 가동조건이 가혹한 경우 예측된 수명보다 빨리 고장에 도달하여 문제가 되거나, 반대로 가동조건이 양호한 경우 늦게 고장에 도달하므로 일찍 정비하는 것이 낭비가 될 수 있다.

최근 이러한 문제 해결을 위해 Fig. 2에서 볼 수 있는 것처럼 고장예지 및 건전성 관리(PHM; prognostics and health management) 기술이 많이 연구되고 있는데 이 기술은 가동 중인 시스템에 대해 핵심 고장모드를 대상으로 실시간으로 건전성(health)을 평가하고 미래의 고장시점을 예측하는 기술이다. 이를 통해 상태기반 정비(CBM; condition based maintenance)가 가능하며, 비용을 절감하면서 더 높은 안전성을 확보할 수 있다. PHM 기술은 센서 데이터로부터 현재의 결합수준을 평가하는 건전성 모니터링(health monitoring), 건전성 데이터를 이용한 예측모델 확립, 확립된 모델을 이용한 미래 고장시점 예측으로 구성된다. 예측모델을 확립하는 방법으로 물리기반 및 데이터기반 두 가지 방법이 존재하는데 [2], 물리기반 방법은 고장 메커니즘을 활용하여 가장 정확한 방법으로 알려져 있으나 확립된 손상물리모델이 많지 않아 그 적용 분야가 좁다. 데이터기반 방법은 손상물리모델 없어도 사용 가능하여 적용범위가 비교적 넓지만 학습을 위해 많은 데이터가 필요하며, 사용조건이 달라지면 다시 학습시켜야 하는 단점이 있다. 물리기반 및 데이터기반 방법의 경우 여러 가지 불확실성(즉

정오차, 제한된 데이터, 모델오류 등)을 효과적으로 고려하기 위해 통계에 기반을 둔 추정/예측 알고리즘이 공통적으로 적용된다.

데이터기반 방법은 신경 회로망(neural network), Gaussian Process Model, Relevance Vector Machine 등의 기계학습(machine learning) 기법을 이용하여 오프라인에서 입력정보 대비 건전성과의 관계를 사전에 훈련한 후 이를 온라인 입력정보에 적용하여 건전성을 추정하고 미래고장시점을 예측하는 것이다. 본 연구에서는 항공기 엔진에 대해 데이터기반 방법을 이용한 수명예측을 수행한다.

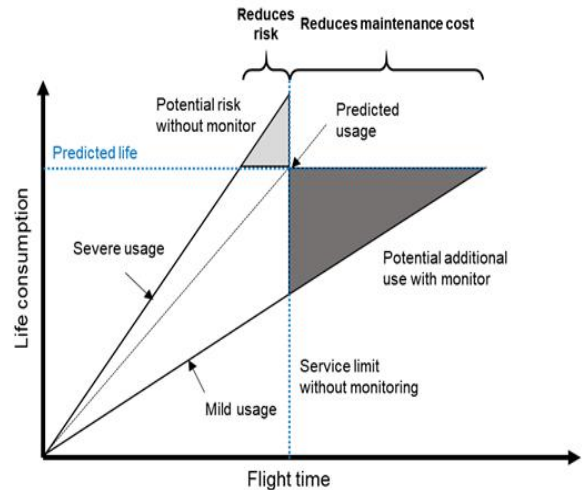


Fig. 1 Assessing Fatigue Life and MTTR based on Fleet Average

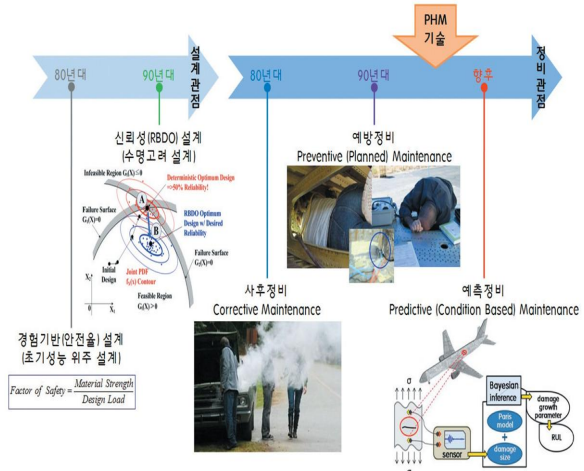


Fig. 2 Trends in Design and Maintenance [1]

PHM의 주요 적용 대상인 항공기 엔진은 높은 안전도 유지를 위해 많은 정비비용이 지출되는 대표적 시스템으로서 Fig. 3에서 보듯 2011년도 전 세계 민항공기의 총운용비 대비 엔진 정비비용이 6%를 차지하고 있으며[3], 이를 낮추기 위해 여러 선진국에서 엔진건전성관리(EHM; engine health management) 기술을 개발하고 있다. EHM은 엔진에 대한 상태 모니터링, 결함탐지, 노화 상태를 추정 및 예측하는 다양한 기능을 수행함으로써 항공기 운영유지비 절감에 큰 역할을 하고 있다. 일반적으로 EHM은 온보드와 오프보드의 두 가지 서브시스템으로 구성되며 온보드에서는 실시간 결함탐지를, 오프보드에서는 엔진성능 열화트렌드 모니터링 및 수명예측을 수행한다 [4, 5].

EHM에서 수명 예측을 위해 가장 많이 사용되는 항공기 엔진의 가상 데이터(C-MAPSS)[6]를 이용하여 데이터 기반 분석법을 적용해 보았다. C-MAPSS를 이용한 데이터 기반 분석 연구는 활발하게 연구되고 있는데, 크게 세 가지 대표적인 방법이 있다. 신경회로망(neural network)과 필터링 방법(filtering method)을 이용한 연구[7, 8], 'health index'를 뽑고 이를 이용하여 잔존 유효 수명을 구하는 연구[9] 그리고 특징인자(feature)의 유사성(similarity)을 이용하여 잔존 유효 수명을 추정하는 방법[10] 등이 활발히 연구되어 왔다. 이 외에도 다양한 방법론이 존재하는데, 이러한 다양한 방법론을 참고문헌[11]에서 분석하고 정리하였다.

본 연구에서는 데이터기반 분석법에서 신경회로망 기법 중에서도 널리 사용되는 MLP(multi-layer perceptron) 이라는 알고리즘을 이용하여 운항 중 모니터링 되는 센서 데이터를 이용하여 수명을 예측하는 문제를 다루었고, 이를 다른 문헌과 비교하는 과정을 거쳤다.

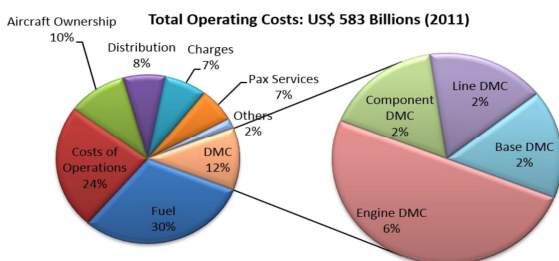


Fig. 3 Repartition of Operating Costs for Airline Companies [3]

## 2. 항공기 엔진의 데이터 기반 PHM 적용

### 2.1 C-MAPSS (항공기 엔진의 가상 데이터)

데이터 기반 PHM을 엔진 운용데이터에 적용하기 위하여 C-MAPSS 데이터를 이용하였다. 이는 NASA의 Prognostics Center of Excellence(PCoE)에서 제공하는 터보 팬 엔진 시뮬레이션 모델로부터 얻은 'Run to fail' 데이터 세트이다. C-MAPSS는 연료 공급 및 13개의 건전성 파라미터(health parameter)로 이루어진 총 14가지의 입력 값(건전성 파라미터)을 통해 고장 및 열화의 영향을 고려한 항공기 가스 터빈 엔진의 출력 데이터를 얻을 수 있다. 고장 및 열화의 대상은 항공기 엔진의 주요 부품인 팬(fan), 저압 압축기(LPC; low pressure compressor), 고압 압축기(HPC; high pressure compressor), 저압 터빈(LPT; low pressure turbine), 고압 터빈(HPT; high pressure turbine)이다.

C-MAPSS에는 Table 1과 같이 'Train set'와 'Test set'로 구성된다. 'Train set'는 시스템의 전체 수명에 대한 결함 성장 및 열화를 확인할 수 있어 알고리즘 모델의 학습을 위하여 활용된다. 'test set'는 시스템이 수명을 다하기 전 임의의 순간까지의 데이터를 담고 있으며 마지막 사이클에서의 잔존 유효 수명을 확인할 수 있기 때문에 알고리즘 검증에 위하여 활용되는 데이터이다. 따라서 'train set'를 이용하여 모델을 구축 후 그 모델을 이용하여 'test set'의 정확한 잔존 유효 수명을 예측하는 것이 C-MAPSS 데이터를 이용한 EHM 구현의 목적이라고 할 수 있다. C-MAPSS는 운용 설정 파라미터의 조합과 결함 조건에 따라 총 4가지(FD001~FD004)의 데이터 세트로 구성되어 있다. 운용 설정 파라미터는 총 3가지로 구성되어 있는데, 마하 수, 고도, TRA(throttle resolver angle)이다. 이 세 가지 운용 설정 파라미터를 조합하여 FD002 및 FD004에 대한 총 6가지의 운용 조건(Operating conditions)이 만들어져있다. 결함 조건(Fault conditions)의 경우, 항공기 엔진의 주요 부품인 팬(fan) 또는 고압 압축기에 대한 결함이 인가되었다. FD001 및 FD002 데이터세트는 고압 압축기의 고장이 인가된 경우이며 그 외의 FD003 및 FD004 데

이터세트는 팬과 고압 압축기에 고장을 인가한 데이터가 혼합되어 있다. 본 연구에서는 4개의 데이터 세트 중에서 6개의 운용 조건과 1개의 결함 조건이 적용된 FD002에 MLP(multi-layer perceptron) 기법을 적용하여 잔존 유효 수명(RUL; remaining useful life) 예측을 진행했다.

C-MAPSS 데이터의 구조는 Fig. 4와 같은데 1열에 있는 ‘unit’ 열은 엔진의 종류를 나타내며 ‘train set’와 ‘test set’에는 각각 260개, 259개의 엔진 데이터가 존재한다. 2열은 사이클을 나타내며, 3~5열은 앞서 설명한 것처럼 운용 설정이다. 6~26열에는 해당 엔진 및 시간에 따른 센서 데이터를 의미한다.

Table 1 C-MAPSS Dataset

| Dataset              | FD001 | FD002 | FD003 | FD004 |
|----------------------|-------|-------|-------|-------|
| Train                | 100   | 260   | 100   | 249   |
| Test                 | 100   | 259   | 100   | 248   |
| Operating conditions | 1     | 6     | 1     | 6     |
| Fault conditions     | 1     | 1     | 2     | 2     |

| unit |      | time    |         |      |        | operating settings |        |       |       |        | sensors |  |  |  |  |  |
|------|------|---------|---------|------|--------|--------------------|--------|-------|-------|--------|---------|--|--|--|--|--|
| Var1 | Var2 | Var3    | Var4    | Var5 | Var6   | Var7               | Var8   | Var24 | Var25 | Var26  |         |  |  |  |  |  |
| 1    | 1    | -0.0007 | -0.0004 | 100  | 518.67 | 641.82             | 1589.7 | 100   | 39.05 | 23.419 |         |  |  |  |  |  |
| 1    | 2    | 0.0019  | -0.0003 | 100  | 518.67 | 642.15             | 1591.8 | 100   | 39    | 23.424 |         |  |  |  |  |  |
| 1    | 3    | -0.0043 | 0.0003  | 100  | 518.67 | 642.35             | 1598   | 100   | 38.95 | 23.344 |         |  |  |  |  |  |
| 1    | 4    | 0.0007  | 0       | 100  | 518.67 | 642.35             | 1592.8 | 100   | 38.88 | 23.374 |         |  |  |  |  |  |
| 1    | 5    | -0.0019 | -0.0002 | 100  | 518.67 | 642.37             | 1592.8 | 100   | 38.9  | 23.404 |         |  |  |  |  |  |
| 1    | 6    | -0.0043 | -0.0001 | 100  | 518.67 | 642.1              | 1594.5 | 100   | 38.98 | 23.367 |         |  |  |  |  |  |
| 1    | 7    | 0.001   | 0.0001  | 100  | 518.67 | 642.48             | 1592.3 | 100   | 39.1  | 23.377 |         |  |  |  |  |  |
| 1    | 8    | -0.0034 | 0.0003  | 100  | 518.67 | 642.56             | 1593   | 100   | 38.97 | 23.311 |         |  |  |  |  |  |
| :    | :    | :       | :       | :    | :      | :                  | :      | :     | :     | :      |         |  |  |  |  |  |
| :    | :    | :       | :       | :    | :      | :                  | :      | :     | :     | :      |         |  |  |  |  |  |
| :    | :    | :       | :       | :    | :      | :                  | :      | :     | :     | :      |         |  |  |  |  |  |

Fig. 4 C-MAPSS Data Structure

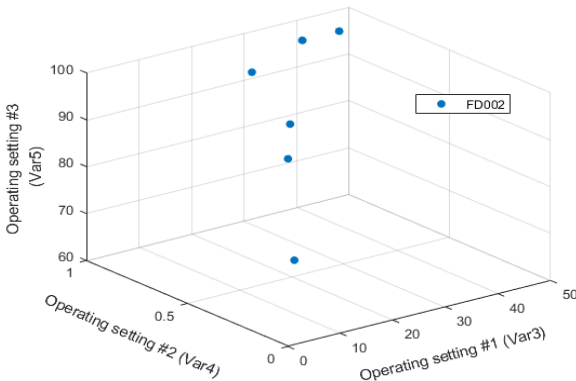


Fig. 5 Six Operating Conditions for All Engine Units of FD002

## 2.2 데이터 전처리

인공신경망을 이용한 데이터 분석을 위해서는 센서 데이터를 이용하여 여러 과정을 거쳐야 한다. 이를 데이터의 전처리 과정이라고 부르며, 데이터의 군집화, 센서 선정, 센서의 정규화의 순서로 진행된다. Figure. 5에서는 두 번째 데이터 세트 FD002의 엔진 데이터에 대하여 각각의 운용 설정(Operating settings) 값을 보여주고 있는 것인데 모든 점들이 6개의 군집으로 분류된다는 것을 확인할 수 있다. 이는 Table 1에 명시된 것처럼 FD002는 6개의 운용 조건의 조합으로 시뮬레이션 된 데이터 세트임을 Fig. 5에서 확인할 수 있는 것이다. 이 조합을 구하기 위해서는 데이터 군집화(clustering)라는 분석 방식을 사용하게 되는데, 데이터 군집화란, 주어진 데이터의 특징을 고려하여 여러 그룹으로 나누는 방법이다. 데이터를 군집화하기 위한 다양한 알고리즘이 존재하지만, 그 중에서도 k-means 알고리즘을 사용하여 군집화를 하였다. 그룹의 개수(k)를 정하고 그 그룹의 중심(centroid)과 그룹 내의 데이터들과의 거리를 제곱합 함수로 정하여 이 함수 값(V)을 최소화 하는 방향으로 반복 계산한다. 이는 Eq. 1의 식으로 표현 할 수 있다. Equation 1에서 i번째 클러스터의 중심을  $\mu_i$ 로 두면, 클러스터에 속하는 점의 집합을  $S_i$ 라고 할 때, 최적의 그룹 중심 값을 찾고 중심 값으로부터 데이터의 분산을 최소화 하는 해를 구한다.

$$V = \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (1)$$

클러스터 중심  $\mu_i$ 의 경우 Eq. 2의 식을 통해 클러스터에 있는 데이터들의 중심 값으로 반복(t)계산으로 재설정해 준다. 이때 k개의 그룹에 대한 초기 중심은 k-means++ 알고리즘으로 할당된다. 이 알고리즘에서 첫 번째 중심( $i=1$ )은 무작위로 선택된 데이터로 설정되고, 이후로는 이전 중심( $i-1$ )과의 거리가 클수록 높은 확률로 선택되도록 가중치를 가하여 하나의 데이터를 선택하는 방식으로 그 중심( $i=2, \dots, k$ )이 설정된다.

$$\mu_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j \quad (2)$$

Figure. 5에서 볼 수 있는 것처럼, 총 6개의 그룹으로 나뉘지기 때문에, k를 6개로 두고, k-mean 알고리즘을 적용한다. 이를 통해 전체 데이터셋을 ‘regime 1’에서 ‘regime 6’으로 라벨링한 후 데이터를 이용하여 각각의 인공 신경망 모델을 만든다.

군집화가 되어 있지 않은 상태에서 센서 값의 그래프로 나타내보면 센서 값의 경향이 전혀 보여 지지 않는다. 예를 들어 Fig. 6과 같이 260개의 엔진 중 첫 번째 엔진의 2번 센서를 그래프로 그려보았다. 이는 위에서 사용한 군집화를 이용하여 해결할 수 있는데, 군집화 알고리즘을 통해 데이터를 총 6개의 집단으로 라벨링 한 후 각 센서 값을 그려보면 Fig. 7과 같이 나타내진다. Figure. 7은 두 번째 엔진에 대한 1번 센서 데이터를 그린 그래프로 6개의 ‘regime’ 별로 데이터를 구분했을 때 센서 값들의 열화 트렌드를 시각적으로 확인할 수 있다. RUL 예측을 효율적으로 진행하기 위해 총 21가지의 센서 데이터 중 일부만을 선별하는 과정을 거치게 된다. 이를 센서 선택(sensor selection)이라고 부르는데, 이 과정을 거치는 이유는 열화와의 상관도를 계산하고 이를 통해 좀 더 상관성이 있는 센서들을 이용하여 인공 신경망 알고리즘의 학습 속도를 높이고 좀 더 잘 학습되게 하기 위해서이다. 이때, 각 ‘regime’ 별로 센서 선택을 하게 되면, ‘regime’마다 조금씩 다른 센서가 선택된다. Equation 3과 같은 Spearman 상관계수 식을 통해 그 값이 0.4를 넘는 센서를 Table 2에 정리하였다. 상관 계수 결과와 선별된 데이터와 그래프는 각각 Table 2 및 Fig. 8에서 확인할 수 있다.

$$r_s = \rho_{rg_X, rg_Y} = \frac{cov(rg_X, rg_Y)}{\sigma_{rg_X} \sigma_{rg_Y}} \quad (3)$$

이렇게 선정된 센서 데이터는 범주를 통일시켜 스케일에 따른 영향을 없애기 위해 전체의 평균과 표준편차를 이용하여 Eq. 4를 통해 정규화가 이루어졌다.

$$x \text{ norm}_i = \frac{x_i - \mu_i}{\sigma_i} \quad (4)$$

센서 데이터에는 노이즈가 포함되어 있기 때문에 정확한 수명 예측이 힘들 수 있다. 따라서 Eq. 5와 같은 가우시안 커널 필터를 이용하여 정규화가 이루어진 센서 데이터에 대하여 노이즈를 제거하였다. Equation 5에서 매끄러운(smoothing) 정도를 조절하는 커널 폭

(kernel width,  $\rho$ ) 값은 교차검증(cross validation)을 거쳐 적절한 값을 선정하였다. Figure 9에는 정규화가 진행된 데이터를 점선으로 표시하였고 이후 가우시안 커널필터를 이용해 노이즈가 제거된 것은 실선으로 표시하였다.

$$z(t) = \frac{\sum_{i=1}^E K(t, t_i) z_i}{\sum_{i=1}^E K(t, t_i)}, \quad (5)$$

$$\text{where } K_G(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\rho^2}\right)$$

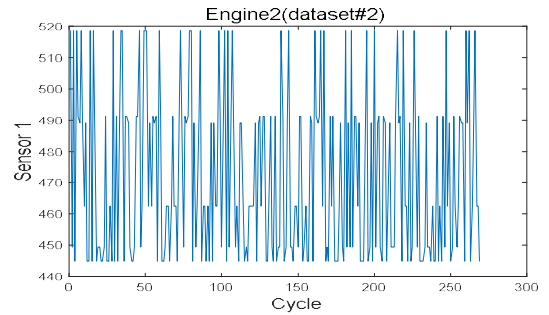


Fig. 6 Sensor 2 of Engine 1 without K-mean Clustering Algorithm

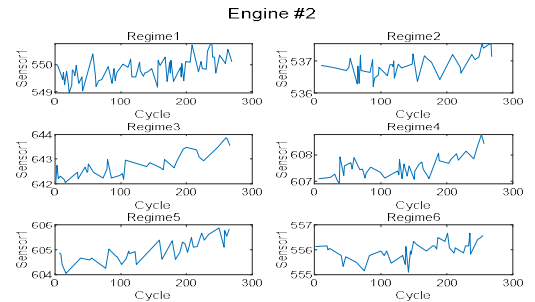


Fig. 7 Sensor 1 Data from Engine #2 Plot for All Operating Regimes

Table 2 Spearman Correlation Calculation Results

|          | Sensor number                          |
|----------|--|
| Regime 1 | #2,3,4,9,11,15,17                      |
| Regime 2 | #2,3,4,9,11,15,17                      |
| Regime 3 | #2,3,4,6,7,8,9,11,12,13,15,16,17,20,21 |
| Regime 4 | #2,3,4,7,8,9,11,12,13,15,17,20,21      |
| Regime 5 | #2,3,4,9,11,15,17                      |
| Regime 6 | #2,3,4,6,7,8,9,11,12,13,15,17,20,21    |



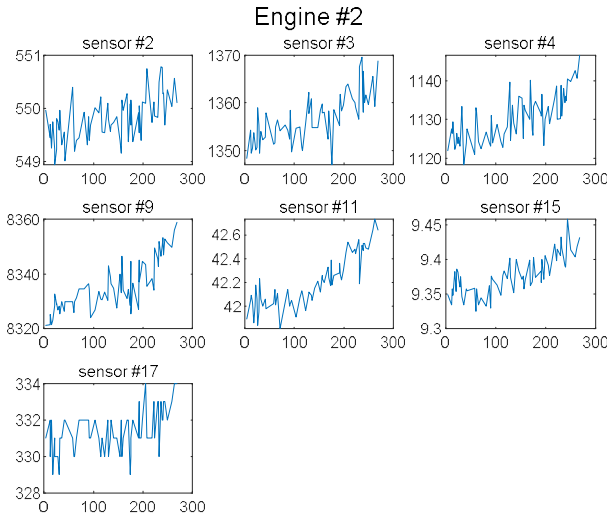


Fig. 8 Selected Sensor Data from Engine #2 for Operating Regime 1

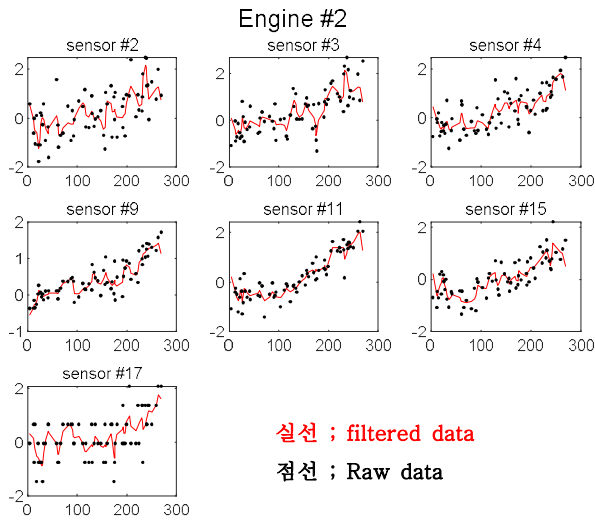


Fig. 9 Normalized Sensor Data from Engine #2 for Operating Regime 1



Fig. 10 4-fold Cross Validation

### 2.3 인공 신경망 학습 및 검증

앞서 언급했던 Table 1처럼 C-MAPSS 데이터세트는 ‘train set’와 ‘test set’로 구성되어 있다. 인공 신경망 모델을 생성하기 위하여 ‘train set’를 활용하는데, ‘train set’에서 최적의 인공 신경망 모델을 생성하기 위해 ‘train set’를 ‘train’과 ‘validation’으로 나누어 학습 시키게 된다. ‘train’을 75%, ‘validation’을 25%로 나누고 인공 신경망 모델을 만들게 되는데, ‘train’을 통해 학습된 모델을 ‘validation’을 이용해 검증한다. 이때, 교차 검증을 위해 k-fold cross validation이라는 방법을 활용하였다. k는 사용자가 임의로 설정할 수 있는데, 본 연구에서는 Fig. 10과 같이 k를 4로 설정하였고 총 데이터를 4개로 나눈 후, 첫 번째 검증에서는 맨 처음 25%를 ‘validation’으로 두고 나머지를 ‘train’으로 두었다. 두 번째 검증부터는 그 다음 25%를 각각 ‘validation’으로 두고 반복하여 검증을 시행한다. 4번의 검증에서 평균값을 이용하여 최적의 ‘hyper parameter’를 찾는다.

모델을 최적화하기 위해서는 노드 수, ‘epoch’, ‘training function’을 이용, ‘hyper parameter’를 변경하여 최적의 학습 결과 값을 찾아내야 한다. Table 3과 같이 ‘hidden layer’ 수를 1로 고정하고 노드 수는 5씩 증가 시키면서 ‘epoch’과 MATLAB에서 제공하는 함수인 ‘training function’을 변경하였는데, 모델 최적화를 위해 사용된 것은 R제곱이다. R제곱이란 통계학에서 결정계수를 뜻하는 말로, 주어진 자료로부터 추정된 모델과 자료와의 적합도를 계산하는 방법이다. 결정 계수의 값은 0에서 1사이이며, 상관관계가 높을수록 1에 가까운 값을 갖는다. R제곱의 식은 Eq. 6, Eq. 7 및 Eq. 8을 이용하여 계산하였다.  $SS_E$ 는 추정치  $\hat{y}_i$ 와 주어진 자료  $y_i$ 사이의 오차 제곱 합이다.  $SS_T$ 는  $y_i$ 의 평균값인  $\bar{y}$ 와 추정치  $\hat{y}_i$ 의 오차의 제곱 합에  $SS_E$ 를 더한 값으로 평균값  $\bar{y}$ 과 주어진 자료의 오차의 제곱합으로도 나타 낼 수 있다. 각각의 ‘hyper parameter’ 조합에 대한 R제곱을 15회씩 반복 계산하였으며 R제곱 평균과 표준편차 값을 통해 해당 조합의 모델 적합도를 평가하였다.

$$R^2 = 1 - \frac{SS_E}{SS_T} \quad (6)$$

$$SS_T = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2 \quad (7)$$

$$= \sum (y_i - \bar{y})^2$$

$$SS_E = \sum (y_i - \hat{y}_i)^2 \quad (8)$$

본 연구에서는 인공 신경망을 이용한 출력(output) 값을 해당 엔진에 대한 잔존 유효 수명의 추정치로 사용하였고, 실제 잔존 유효 수명을 'target' 자료로 사용하였다. R제곱을 이용하여 최고의 'hyper parameter' 조합을 찾아내었고 이는 Table 4의 결과 값과 같다.

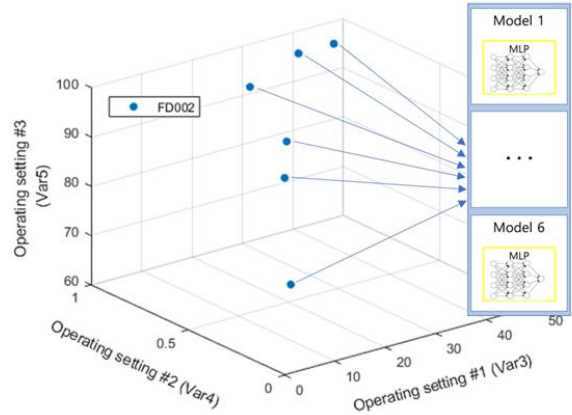
Table 4와 같이 'regime'별 최적화된 결과를 활용하여 노드 수, 'epoch', 'training function'을 고정시킨 후 'train set' 100%를 이용하여 MLP의 구조를 생성하고 인공 신경망 모델을 학습시켰다. Figure 11과 같이 MLP는 총 6개의 모델이 만들어 졌는데, 이는 각 'regime' 별로 하나씩 인공 신경망 모델을 생성했기 때문이다. 각각의 'regime'별로 'target'값을 잔존 유효 수명으로 두고 입력 값을 앞에서 진행한 전처리 과정을 거친 센서 값으로 두어 모델을 생성하고 학습시켰다. 인공 신경망을 이용하여 나온 잔존 유효 수명의 추정치와 실제 잔존 유효 수명을 Fig. 12에 비교하였다. Figure 12에 명시된 'Engine 50', 'Engine 100', 'Engine 200' 및 'Engine 260'은 'train set'의 엔진 'unit' 번호를 의미한다.

**Table 3** Parameters for 4-fold Cross Validation

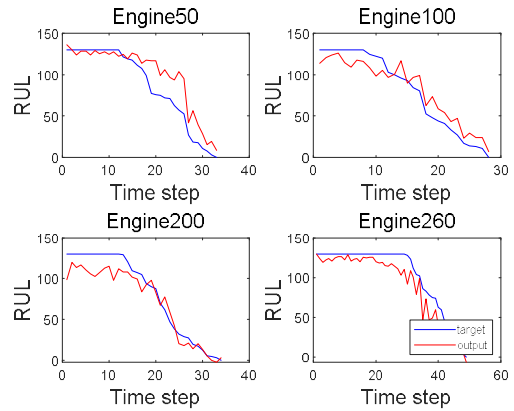
| Parameter         | Values                     |
|-------------------|----------------------------|
| Hidden nodes      | 5, 10, 15, 20, 25          |
| Hidden layers     | 1                          |
| Epochs            | 200, 400, 1000             |
| Training function | trainlm, trainbr, traingcp |

**Table 4** Best Hyper Parameter Using R-squared Calculation Results

|         | Hidden nodes | Epochs | Training Fcn | R <sup>2</sup> mean | R <sup>2</sup> std |
|---------|--------------|--------|--------------|---------------------|--------------------|
| Regime1 | 5            | 200    | traingcp     | 0.529               | 0.01               |
| Regime2 | 5            | 200    | traingcp     | 0.538               | 0.02               |
| Regime3 | 5            | 200    | traingcp     | 0.544               | 0.01               |
| Regime4 | 5            | 200    | traingcp     | 0.535               | 0.02               |
| Regime5 | 5            | 200    | traingcp     | 0.533               | 0.02               |
| Regime6 | 25           | 400    | trainbr      | 0.561               | 0.01               |



**Fig. 11** Six Neural Network Models for Each Operating Regime



**Fig. 12** RUL Trend Using the Trained Neural Network in Regime 6

## 2.4 인공신경망을 이용한 잔존 유효 수명 예측

앞서 설명한 것처럼 'train set'를 이용하여 생성한 인공신경망 모델을 'test set'에 적용하여 'test set'에서 주어진 259개의 엔진의 잔존 유효 수명을 예측한다. 'test set'의 경우 시스템이 종료되기 전에 특정 시점에서 임의로 관측이 중단된 데이터이기 때문에 그 관측 중단 시점에서부터 고장까지의 잔존 유효 수명을 정확하게 맞추는 것이 이 데이터 세트의 목적이다.

인공 신경망은 확률 통계를 기반으로 반복적인 계산을 이용하여 최솟값을 찾아가는 방식으로 학습하기 때문에 답이 일정하게 나올 수 없다. 답이 매번 구할 때마다 다르고 값들 간의 극소 값과 극대 값의 차이가 크다. 어떤 특정 값을 제외하거나 단순히 평균값을 구

하는 대신에 얻어진 모든 값을 활용하면서 하나의 답을 도출하기 위하여 히스토그램 개념을 도입하였다. 인공 신경망을 통해 200회 반복 계산하여 각 엔진별 잔존 유효 수명 값을 200개씩 얻을 수 있다. 잔존 유효 수명 값이 200개인 이유는 마지막 사이클(관측 중단 시점)에서의 추정치만을 취하였기 때문이다. 이렇게 총 259개의 엔진에 대하여 200개씩 얻어진 잔존 유효 수명 값을 누적하여 Fig. 13과 같은 확률 분포의 히스토그램 그래프를 그릴 수 있다. 일반적으로 잔존 유효 수명이 분포로 제시될 경우 보수적으로 95% 신뢰 하한 값을 최종 추정치로 정하므로 본 논문에서도 이 값을 채택하였다. 추가적으로, 본 논문에서 사용한 데이터셋 FD002의 경우, 사이클마다의 'regime'이 다르기 때문에 총 6개의 인공 신경망 모델 중에서 해당 'regime'에 맞는 모델을 사용하게 된다. 그러므로 Fig. 13의 결과는 마지막 사이클의 'regime'에 맞는 인공 신경망 모델을 통해 잔존 유효 수명을 예측한 것이다.

예측한 잔존 유효 수명과 정답이라고 알려진 잔존 유효 수명 값을 비교해 보고 결과를 확인 할 수 있다. 다음과 같은 Eq. 9 [12]을 이용하여 잔존 유효 수명 예측 점수를 계산할 수 있다. 이 식은 PHM08 데이터 챌린지에도 사용된 식으로 정답 확인 비교를 위해 사용된다. 정답보다 더 사용할 수 있다고 예측한 결과일수록 패널티가 더 크고 일찍 예측한 결과일수록 패널티가 더 작다. 다른 논문과의 결과 값을 비교한 것은 Table 5와 같다.

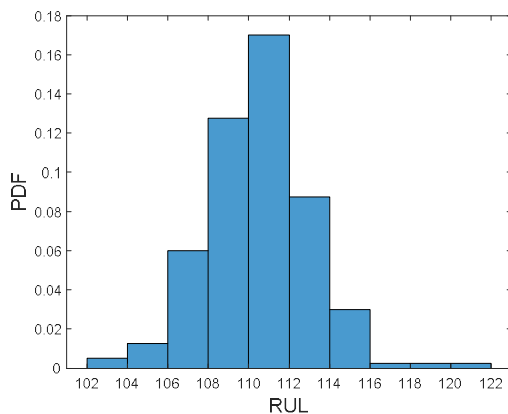


Fig. 13 Distribution of RUL Estimation from Engine #259

$$S = \begin{cases} \sum_{i=1}^n e^{-\left(\frac{d_i}{a_1}\right)} - 1 & \text{for } d_i < 0 \\ \sum_{i=1}^n e^{\left(\frac{d_i}{a_2}\right)} - 1 & \text{for } d_i \geq 0 \end{cases} \quad (9)$$

where,

$S$  is the computed score,

$n$  is the number of Engines,

$$d_i = \hat{t}_{RUL} - t_{RUL}$$

$$= (\text{Estimated RUL} - \text{True RUL}),$$

$$a_1 = 13, a_2 = 10$$

Table 5 Performance Comparison with Other Papers for FD002

| Dataset 2      |                        |
|----------------|------------------------|
| Paper #        | Score S of each method |
| 1 [13]         | 7800000                |
| 2 [14]         | 18255                  |
| 3 [15]         | 13018                  |
| Present method | 18402                  |

### 3. 결론

본 연구에서는 인공신경망을 적용하여 항공기 엔진의 수명 예측 프로세스를 구축하였다. NASA에서 배포한 가상의 항공기 엔진 열화 데이터를 이용하여 각 운용조건에 맞는 6개의 인공 신경망 모델을 생성하고 이를 이용하여 잔존 유효 수명을 예측해 보았다. 각 엔진별 수명을 예측한 결과 상당부분에서 정답에 근접한 예측을 수행함을 확인하였고 근래에 나온 논문들과 비교해 보았을 때, 비슷한 수준이거나 더 좋은 결과를 갖는다는 것을 알 수 있었다. 항공기 엔진의 사용 중 고장발생 전에 적절한 조치를 취할 수 있는 가능성을 제시하였다. 물론, 이 데이터는 가상의 상황에서 임의로 만들어진 가상 데이터이기 때문에 실제 항공기 엔진의 이 알고리즘을 적용하기 위해서는 검증과 개선이 필요하다. 이는 향후 실제 항공기 엔진 데이터 확보를 통해 검증 및 개선이 이루어질 수 있다.

후 기



이 논문은 2016년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (2016R1A2B4015241).

이 출판물은 2018년도 한국항공대학교 교비지원 연구구비에 의하여 지원된 연구의 결과임(2018-01-008).

## References

- [1] J.-H. Choe, "Introduction to Prognostics and Health Management technology," *J. KSME*, vol. 53, no. 7, pp. 26-34, 2013.
- [2] J. Choi, "Research Trends and Application Cases for Prognostics and Health Management," *J. Aerosp. Syst. Eng.*, vol. 8, no. 4, pp. 7-17, 2014.
- [3] B. Lamoureux, J.-R. Massé, and N. Mechbal, "Improving aircraft engines prognostics and health management via anticipated model-based validation of health indicators," *Progn. J.*, vol. 2, no. 1, pp. 18-38, 2014.
- [4] T. Kobayashi and D. L. Simon, "Integration of on-line and off-line diagnostic algorithms for aircraft engine health management," *J. Eng. Gas Turbines Power*, vol. 129, no. 4, pp. 986-993, 2007.
- [5] L. C. Jaw, "Recent advancements in aircraft Engine Health Management (EHM) technologies and recommendations for the next step," *Proc. ASME Turbo Expo*, vol. 1, no. April, pp. 683-695, 2005.
- [6] A. Saxena and K. Goebel, "PHM08 Challenge Data Set," NASA Ames Prognostics Data Repository, Moffett Field, CA, 2008.
- [7] F. O. Heimes, "Recurrent neural networks for remaining useful life estimation," in 2008 International Conference on Prognostics and Health Management, 2008.
- [8] S. Zheng, K. Ristovski, A. Farahat, and C. Gupta, "Long Short-Term Memory Network for Remaining Useful Life estimation," 2017 IEEE Int. Conf. Progn. Heal. Manag. ICPHM 2017, pp. 88-95, 2017.
- [9] J. B. Coble, "Merging data sources to predict remaining useful life--an automated method to identify prognostic parameters," pp. 1-223, 2010.
- [10] T. Wang, "Trajectory Similarity Based Prediction for Remaining Useful Life Estimation," University of Cincinnati. 2010.
- [11] E. Ramasso and A. Saxena, "Review and analysis of algorithmic approaches developed for prognostics on CMAPSS dataset," in PHM 2014 - Proceedings of the Annual Conference of the Prognostics and Health Management Society 2014, 2014, pp. 612-622.
- [12] A. Saxena, K. Goebel, D. Simon, and N. Eklund, "Damage propagation modeling for aircraft engine run-to-failure simulation," 2008 Int. Conf. Progn. Heal. Manag. PHM 2008, 2008.
- [13] G. S. Babu, P. Zhao, and X. L. Li, "Deep convolutional neural network based regression approach for estimation of remaining useful life," *Int. Conf. database Syst. Adv. Appl.*, pp. 214-228, 2016.
- [14] X. Li, Q. Ding, and J. Q. Sun, "Remaining useful life estimation in prognostics using deep convolution neural networks," *Reliab. Eng. Syst. Saf.*, vol. 172, no. December 2017, pp. 1-11, 2018.
- [15] C. Zhang, P. Lim, A. K. Qin, and K. C. Tan, "Multiobjective Deep Belief Networks Ensemble for Remaining Useful Life Estimation in Prognostics," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 28, no. 10, pp. 2306-2318, 2017.