



다변량 통계기법과 DW-SIM을 연계한 스크류 공기압축기 성능분석 및 이상감지

Efficiency and Fault Detection of Screw Air Compressors Using Multivariate Analysis and DW-SIM

강병운* · 황동곤**

Byeong-Un Kang* · Dong-kon Hwang**

* Main author, Corresponding author, Chief researcher, WooWon M&E Inc., Republic of Korea (ronalduo7@300302.com)

** Coauthor, Director of Research Institute, WooWon M&E Inc., Republic of Korea (dkh@300302.com)

ABSTRACT

Purpose: This study aims to improve the operational efficiency of screw air compressors, which are widely used in industrial settings, by enabling early detection of potential failures. This approach is intended to minimize unnecessary downtime and enhance overall productivity and workplace safety. **Method:** To achieve this, an anomaly detection algorithm based on Hotelling's T^2 and Q-statistics using principal component analysis (PCA) was developed. Key operating variables—including intake temperature, discharge pressure, discharge temperature, compressed air flow rate, and instantaneous power—were analyzed using multivariate statistical techniques, utilizing approximately one month of operational data collected from an actual industrial site. PCA-based contribution analysis was applied to identify anomalies and diagnose their causes. For performance validation, efficiency analysis using DW-SIM was conducted and the results were compared with the standard calculation formula (KS B 6351) and DW-SIM simulation outputs. **Result:** The analysis results show that the prediction error is less than 5% (RMSE 4.41, MAE 3.55, MAPE 4.76%) and the coefficient of determination (R^2) is 0.75, confirming the reliability and usability of the anomaly detection and performance diagnosis system.

KEYWORD

공기압축기
유지보수
기계학습
성능분석

Air Compressor
Maintenance
Machine Learning
Efficiency

ACCEPTANCE INFO

Received Aug. 1, 2025
Final revision received Sep. 15, 2025
Accepted Sep. 19, 2025

© 2025. KIEAE all rights reserved.

1. 서론

1.1. 연구의 배경 및 목적

산업 및 건축현장에서 발생하는 각종 사고와 생산성 저해 요인을 분석해보면, 일반 생산·사업용 기계 및 설비(예: 동력기계, 건설용 장비, 목재가공기계 등)가 전체의 약 44.3%로 가장 높은 비중을 차지하며 재료, 운반, 등록, 구조물 등 다양한 요소들이 복합적으로 작용함을 알 수 있다. 특히, 사고 당시 기계장치 및 설비 작동과정에서 발생한 비율이 38.3%로 상당히 높으며, 원자재 및 물질의 취급(29.1%), 기계장비의 수리 및 보수(20.2%) 등이 뒤를 잇고 있다[1]. 이는 제조업 특성상 기계와 설비의 안정적 운영이 전체 생산성과 안전에 미치는 영향이 매우 크다는 점을 시사한다. 기존의 설비 보전 방식은 주로 사후(reactive) 또는 일정 주기의 예방적(preventive) 정비에 의존해 왔으나, 이는 불필요한 유지비용 증가, 불가피한 설비 가동 중단 등 비효율을 초래하는 한계가 있다. 따라서 설비종합효율(OEE, Overall Equipment Effectiveness)[2]을 높이기 위한 방안으로, 설비의 고장 및 품질 저하를 사전에 예측하여 적기에 보전활동을 수행하는 '예지정비(Predictive Maintenance)'의 필요성이 크게 대두되어 왔다. 그러므로 본 연구에서는 기계학습 기반의 이상감지 시스템을 통해, 고장 가능성을 조기에 탐지·예측하고, 설비효율을 저

해하는 요인을 분석하여 불필요한 비가동 시간을 최소화하여 전체 생산성 및 작업 안전 수준을 향상시키고자 한다.

1.2. 연구의 방법 및 범위

본 연구의 목적은 산업용 스크류 공기압축기의 이상징후 판단을 위한 알고리즘(예: 주성분 분석 기반 Hotelling's T^2 및 Q-통계량 등)의 개발과 실시간 성능분석 프로그램의 오차율을 검증하는 데 있다. 이를 위해 관련 선행연구를 조사하여 이상징후 판단 알고리즘의 이론적 타당성과 적용 가능성을 확인하며, 실제 산업현장에서 운영 중인 공기압축기 운전 데이터를 수집하고 적용함으로써 프로그램 결과값과 표준(KS B 6351)에 의한 계산값 간의 오차율을 산정한다.

2. 관련 선행연구 고찰

본 절은 기존 연구 사례를 통해 단변량 관리도의 한계를 짚고, 다변량 통계인 PCA- T^2 -SPE-UCL의 이론과 판정 절차(기여도 분석 포함)를 정리하여 본 연구의 방법론적 토대를 분명히 하는 데 목적이 있다.

2.1. 공정 모니터링 기술의 최근 연구 동향

유창규 등(2008)[3]은 단변량 관리도가 변수 간 상관관계를 반영하지 못해 정상·비정상 오판이 발생하고(변수 수가 많을수록 악화), 따라



Fig. 1. Procedure for conducting research

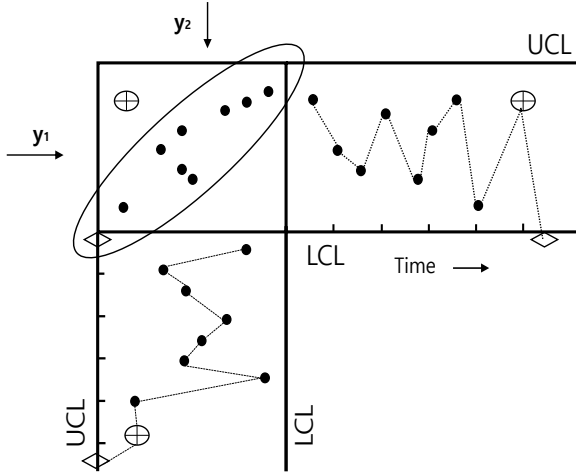


Fig. 2. The problems with using univariate control charts per one variable[4]

서 여러 변수를 한꺼번에 보는 다변량 관리도가 필요하다고 지적한다. 예를 들어 정상인데 비정상으로 오판(⊕)하거나, 반대로 비정상인데 놓침(◇)이 발생할 수 있음을 Fig. 2.로 보여주며, 다변량 감시의 필요성을 명확히 한다.

이어서 이들은 주성분 분석 기반 다변량 통계 공정관리를 표준틀로 제시하는데, 상관이 큰 다차원 데이터를 소수의 주성분 축으로 투영하여 감시하고, 모델 공간의 변화 크기를 보는 Hotelling's T²와 모델이 설명하지 못한 잔차를 보는 Q-통계량(SPE) 두 통계로 상태를 판단한다. 두 통계에는 각각 관리상한선(UCL)이 설정되며, Hotelling's T²는 $F(p, n-p)$ 기반 한계, Q-통계량(SPE)는 Jackson 근사로 한계를 구해 둘 다 관리상한선(UCL) 미만이면 정상으로 판정하는 절차를 분명히 한다. 이해를 돕자면, T²는 선택한 주성분(모델) 공간에서 평균 중심으로부터 얼마나 멀리 있는지, Q(SPE)는 모델이 설명하지 못한 잔차(나머지 축) 쪽으로 얼마나 새어 나갔는지를 확인하는 관리도이다.

한편, 실제 공정은 비선형성(변수 관계가 직선이 아닌 곡선·포화 등), 다중 운전모드(제품·부하·조업단계별로 정상 상태가 여러 개), 공정 변화(장비 설정 변경, 원료 교체, 노후화/계절), 센서 이상(바이어스, 드리프트, 노이즈 증가, 고장) 등으로 전통적 선형·단일모델 가정이 흔들리기 쉽다. 유창규 등은 이를 보완하기 위해 커널 주성분 분석(커널 매핑으로 비선형 패턴을 선형처럼 다룸), 혼합(다중)모델(운전 모드별 로컬 주성분 분석으로 거짓 경보 완화), 적응 주성분 분석(시간에 따라 안정적으로 모델 갱신), 센서 이상 진단·보정을 주성분 분석-Hotelling's T²-Q(SPE)-관리상한선(UCL)의 기본 틀 위에서 확장 적용하는 방향으로 정리한다.

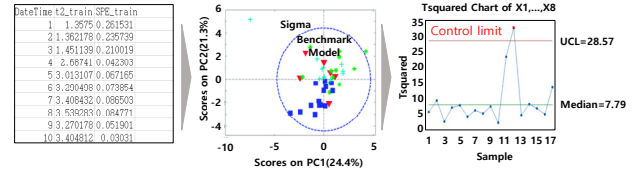


Fig. 3. Procedure for fault detection

2.2. 다변량 통계기법을 활용한 실시간 수질이상 유무 판단 시스템 개발

허태영 등(2015)의 연구[5]는 주성분 분석 기반의 다변량 통계기법을 활용하여, 실시간 데이터 기반의 공정 이상진단 및 경보 시스템 개발 사례를 제시하고 있다. 본문에서는 실시간 수질 데이터(pH, 수온, 알칼리도, 전기전도도, 탁도 등)의 상관관계를 고려한 주성분 분석 모델을 기반으로, 정상/비정상 상태를 구분할 수 있는 Hotelling's T² 통계량을 산출한다. Hotelling's T² 통계량은 정상 데이터로부터 학습된 통계적 모형과 새로운 측정값과의 거리(D_{new} , 마할라노비스 거리 등)를 의미하며 식 (1)과 같이 계산된다.

$$D_{new} = t_{new}^T S^{-1} t_{new} \quad (Eq. 1)$$

여기서, t_{new}^T 는 주성분 공간에서 새 표본의 점수벡터의 전치이며, S 는 훈련(정상)데이터의 점수 공분산행렬로 S^{-1} 는 S 를 표준화하기 위한 역행렬, t_{new} 는 새 표본의 점수벡터이다. 이때 새로 유입된 새 표본이 통계적 정상 범위에서 벗어날 경우 식 (2)에 의해 즉각적으로 이상 상태를 경보한다. 여기서, n 은 전체 표본의 개수, k 는 주성분 모형에서 사용된 주성분 개수(차원의 개수), $F_{(k, n-k, \alpha)}$ 는 자유도 k 와 $n-k$ 를 가진 분포에서 임계확률 α 를 의미한다.

$$D_{UCL} = \frac{k(n-1)}{(n-k)} F_{(k, n-k, \alpha)} \quad (Eq. 2)$$

Fig. 3.는 정상 기간 기준 데이터로 주성분 모델을 학습한 뒤, 주성분 공간에서 Hotelling's T² 관리상한(UCL)을 설정하고, 새 표본을 투영해 T²를 계산한 뒤, 한계 초과 여부로 이상을 판정하는 절차를 보여준다. 또한, 이상 발생 시 기여도 분석 식 (3)을 통해 주요 원인 변수(센서, 공정인자 등)를 파악할 수 있도록 하였다. 이 식은 마할라노비스의 거리에서 새 표본의 점수벡터를 j 번 변수가 만들어내는 점수 변화로 대체한 것이다. $x_{new, j}$ 는 변수 j 값, p_j 는 적재행렬, $(P^T P)^{-1}$ 은 주성분 공간으로의 좌표 변환 보정행렬이다. Fig. 4.는 주성분 점수공간에서 관리상한(파란 타원)을 벗어난 녹색 표본에 대해, 각 변수가 Hotelling's T²를 얼마나 증가시켰는지를 보여주는 기여도(contribution) 예시를 나타낸다. 막대가 클수록 해당 변수가 더 크게 기여하며, 양(+)은 T²를 키우고 음(-)은 줄인다. 모든 기여도의 합은 그 표본의 T²와 일치한다.

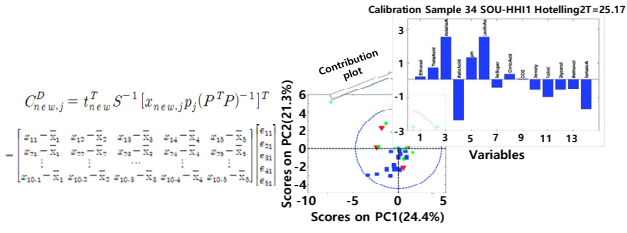


Fig. 4. Contribution analysis: Concept and Formula

$$C_{new,j}^D = t_{new}^T S^{-1} [x_{new,j} p_j (P^T P)^{-1}]^T \quad (\text{Eq. 3})$$

Hotelling's T^2 통계량과 비슷하게 이상 여부 판단을 위해 Q-통계량(SPE)을 활용할 수도 있다. Q-통계량은 주성분 모델에서 재구성(reconstruction)되지 못한 잔차(residual) 크기를 나타낸다. 즉, 관측 데이터가 주성분 모델로 얼마나 잘 설명되는가(설명되지 못하는 부분이 얼마나 큰가)를 나타내는 통계량으로 SPE (Squared Prediction Error)라 하며, 식 (4)와 같이 실제 측정값(x_{new})과 추정값($t_{new}^T P^T$)과의 잔차를 구하고 이를 내적(e_{new}^T)하여 곱하는 방식 식 (5)로 계산할 수 있다.

$$e_{new} = x_{new} - t_{new}^T P^T \quad (\text{Eq. 4})$$

$$Q = e_{new} e_{new}^T \quad (\text{Eq. 5})$$

이상 판정에 있어 95%, 99%, 99.99% 등 다단계 관리상한선을 적용하여, 경보의 민감도/엄격도를 유연하게 조정할 수 있도록 하였으며 실제 현장(정수장) 데이터에 성공적으로 적용해 실질적 효과를 검증하였다. 이 연구는 수질 분야에서 PCA-T²-SPE-UCL 체계를 현장 데이터에 적용해 정상/이상 판정의 실효성과 원인 변수 진단 가능성(기여도)을 동시에 보여준다.

2.3. Multivariate PCA-based Composite Criteria Evaluation Method for Anomaly Detection in Manufacturing Data

이혁수 등(2024)의 연구[6]는 제조 데이터의 이상 탐지를 위해 PCA 기반 복합 판단기준(Composite Criteria) 방법을 제안하였다. 고차원 공정변수를 PCA로 저차원 주성분 공간에 사상한 뒤, 상관구조를 고려하는 Hotelling's T^2 와 시계열 자기상관(Box-Pierce) 각 통계의 관리상한선(UCL)을 카이제곱분포 χ^2 로 근사하였다.

$$T^2 \text{ UCL} = \frac{(n-1)p}{n-p} F_{\alpha}(p, n-p) \sim \chi_{\alpha}^2(p) \quad (\text{Eq. 6})$$

$$BP \text{ UCL} = \frac{(n^2-1)p}{n(n-p)} F_{\alpha}(p, n-p) \sim \chi_{\alpha}^2(p) \quad (\text{Eq. 7})$$

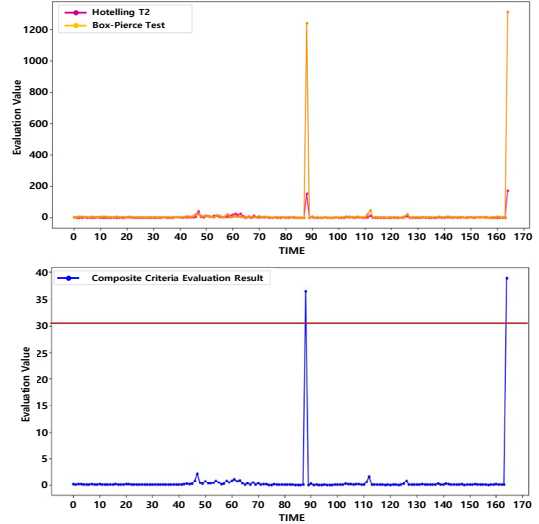


Fig. 5. Composite criteria evaluation visualization results

여기에서 n 은 전체표본의 수, p 는 차원의 개수, $F_{\alpha}(p, n-p)$ 는 자유도 p 와 $n-p$ 를 갖는 F분포를 의미하며 두식은 모두 주성분공간에서의 관리상한선으로 사용된다. 식 (8)은 Hotelling's T^2 와 시계열 자기상관 통계량을 각자의 관리상한선으로 정규화하여 하나의 합성기준값으로 통합한다. 여기에서 α_1 은 T^2 통계량 α_2 은 시계열 자기상관 통계량, β_1 은 T^2 관리상한선, β_2 는 시계열 자기상관 관리상한선이다.

$$\text{Composite Criteria Evaluation Value} = \frac{\alpha_1}{\beta_1} + \frac{\alpha_2}{\beta_2} \quad (\text{Eq. 8})$$

이러한 방법을 통해 화학제품 생산 데이터셋 10개로 유효성을 검증한 결과, 데이터셋 관점 정확도 70%, 개별 데이터 관점 정확도 99.82%를 보고하였다. Fig. 5.는 Hotelling's T^2 와 자기상관통계의 각각 통계량(상)과 합성기준값(하)의 관리상한선을 나타낸 것이다. (상)은 스케일이 달라서 같은 축에서 비교가 어려운 반면 (하)는 두 신호를 한 줄로 합쳐 임계를 초과하는 부분을 명확히 판단할 수 있다. 이 연구는 다변량 상관성과 시계열성을 동시에 고려해 과다 경보를 줄이면서도 실시간 이상탐지에 적용 가능하다는 점과 Hotelling's T^2 기반의 대안적 이상감지 절차도 제시하였다.

2.4. 시간 영역 통계 기반 웨이퍼 이송 로봇의 고장 진단

김해진 등(2024)의 연구[7]는 반도체 공정의 핵심 설비인 웨이퍼 이송 로봇을 대상으로, 시계열통계 기반 주성분 분석의 차원축소와 Hotelling's T^2 를 결합한 고장 진단 알고리즘을 제안하였다. 로봇 프레임에 가속도센서를 부착해 진동을 측정하고, 서보드라이버의 토크를 100개의 실험 세트별로 15초간 수집하여 분석하였다. 이후 정규화 및 주성분 분석으로 저차원 공간을 구성하고, 정상 군집 중심 대비 유클리드 거리(변수간 상관관계를 배제)와 Hotelling's T^2 를 계산하여 관리상한(UCL)으로 이상을 판정하는 구조를 제시하였다.

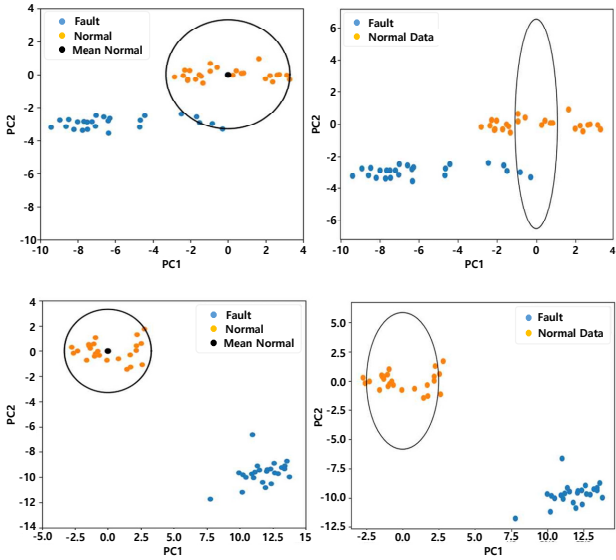


Fig. 6. PCA graph of acceleration (top) and torque (bottom) data

Fig. 6.은 가속도(상)와 토크(하) 각각의 주성분 공간에서의 이상감지 그래프이다.

본 논문은 속도 조건(200/300/400/500mm/s)을 구분해 정상/고장을 판정하면 정확도가 크게 향상됨을 보였다. 가속도 신호에서 유클리드 거리는 속도 분리 전 59%였으나 속도별 분리 후(88/100/100/96%)로 개선되었고, Hotelling's T^2 는 분리 전 60%, 분리 후(100/100/100/98%)로 향상되었다. 반면 토크 신호는 유클리드 거리 조건에서 100%의 높은 정확도를 보였으며, T^2 기준에서도 대부분 100%(500mm/s에서 98%)를 달성했다. 따라서 동일 가중의 유클리드 거리보다 공분산을 반영하는 Hotelling's T^2 가 이상감지에 더 효과적이며, 신호 채널 측면에서는 가속도보다 토크가 평균적으로 더 높은 진단력을 가진다.

2.5. 소결

이상, 주성분 분석(PCA)을 기반으로 Hotelling's T^2 , Q-통계량(SPE)을 활용한 데이터 기반 이상감지의 유효성을 다양한 분야에서 입증하고 있다. 다만, 기존 문헌이 '다변량 통계 = 이상 판정'에 머무는데 비해 본 연구는 이를 바탕으로 대상 영역을 기계설비(스크류 공기압축기) 분야로 확장 및 통계적 이상 판정을 기여도 분석과 DW-SIM 기반 성능 추정에 연계하여, 이상-원인-개선안-운영반영으로 이어지는 피드백 프레임워크를 제시하고자 한다.

3. 연구방법

3.1. 데이터기반 센서이상예측 및 성능분석 구조

용적형압축기는 산업현장에서 가장 오랫동안 사용되던 압축기로서 일정한 공간 내 체적의 감소를 통해 압력을 증가시키는 압축방식을 지닌다. 대표적으로 피스톤식 또는 스크류식으로 구분할 수 있으며 피스톤에 비해 효율이 높고 변속운전이 가능한 장점으로 대부분의 현장에서 스크류 압축기를 쉽게 찾아볼 수 있다. 본문에서는 이러

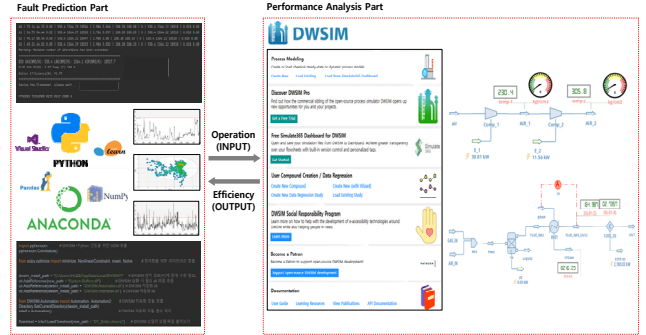


Fig. 7. Composition of the smart maintenance program

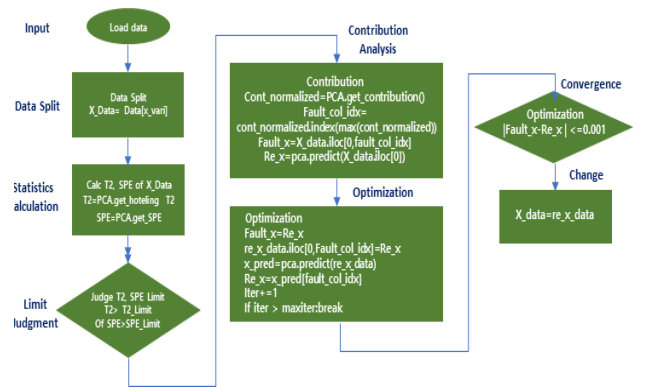


Fig. 8. Fault detection algorithm

한 스크류압축기의 운전데이터를 수집하여 학습시키고 새로운 데이터의 정상여부를 판단하는 구조와 해당 데이터를 공기압축기 모델에 연동하여 성능을 도출하는 프로그램을 설명하도록 한다.

당 프로그램은 Fig. 7.처럼 크게 이상감지 파트(좌측)와 효율분석 파트(우측)로 구분할 수 있고 이상감지는 Python, 효율분석은 DW-SIM을 사용하였다. Python은 오픈소스 프로그래밍 언어로 데이터 분석, 통계처리, 머신러닝, 인공지능, 웹 개발, 자동화, 공학 계산 등 거의 모든 분야에서 활용되고 있으며 DW-SIM은 오픈소스 화학공학 공정 시뮬레이터로 다양한 단위조작(반응기, 열교환기, 분리탑 등), 유체흐름, 열역학 계산 등 시스템 전체를 쉽게 모델링할 수 있다.

먼저 데이터기반 이상감지 파트의 구조는 Fig. 8.의 절차로 진행된다. 데이터를 학습/평가로 분리한 뒤 주성분 분석으로 Hotelling's T^2 및 Q-통계량을 계산해 관리 상한을 초과하면 이상으로 판정한다. 이상이면 변수별 기여도를 계산해 영향이 가장 큰 변수를 선택하고, 그 변수의 관측값(Fault_x)과 재구성값(Re_x)의 차이가 최소화 되도록 반복 최적화를 수행한다(예: $|Fault_x - Re_x| \leq \epsilon$). 수렴하면 보정된 값으로 X를 갱신한다.

다음 공기압축기 운전효율을 파악하기 위해서는 용적형 압축기의 시험 및 검사방법(KS B 6351)에 따라 식 (9), 식 (10)으로 계산이 가능하다. 여기서, i 는 압축 후 높아진 공기온도를 낮추는 중간냉각기의 수, k 는 공기의 비열비(C_p/C_v), P_s 는 압축기 입구에서의 절대압력, Q_s 는 흡입상태로 환산한 공기량, P_d 는 압축기 출구에서의 절대압력이다. 이렇게 구해진 이론단열공기동력 L_{ad} 과 압축기 축동력

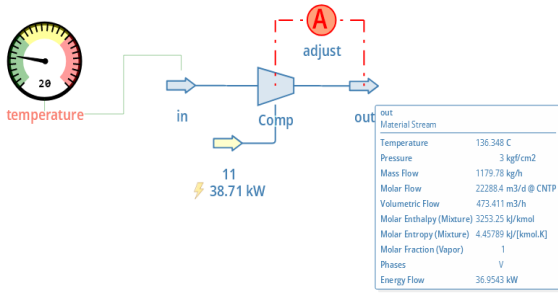


Fig. 9. Air compressor model implemented with dw-sim

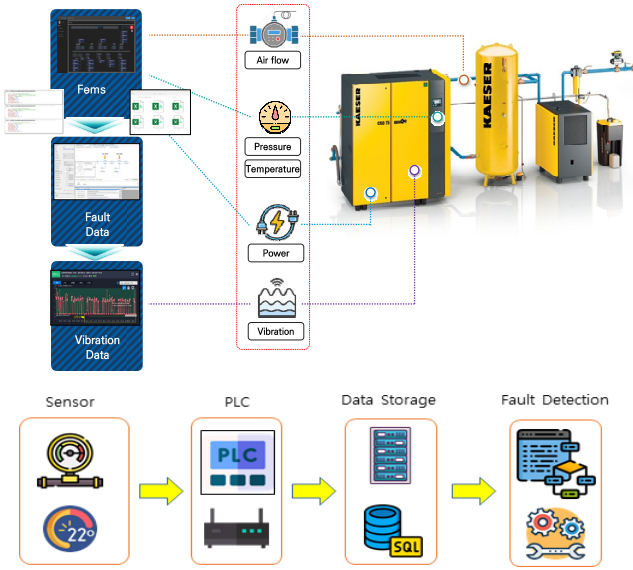


Fig. 10. Communication infrastructure for data Acquisition

I_s 의 비로 효율 η_{ad} 을 정의한다.

$$I_{ad} = \frac{(i+1)k}{k-1} \cdot \frac{P_s Q_s}{0.06} \left[\left(\frac{P_d}{P_s} \right)^{\frac{k-1}{(i+1)k}} - 1 \right] \quad (\text{Eq. 9})$$

$$\eta_{ad} = \frac{I_{ad}}{I_s} \times 100 \quad (\text{Eq. 10})$$

또 다른 방법으로 Fig. 9.과 같이 입구 조건(온도·압력·유량)과 출구 조건(온도·압력)을 입력으로 하는 모델을 구성해, 효율과 소비전력을 계산할 수 있다. Adjust(A) 블록은 제어변수(예: 온도·압력)에 맞도록 목표 조건(효율)을 자동 조정한다. 시뮬레이션 결과로 토출 상태, 에너지, 등엔트로피·폴리트ropic 효율, 전력(에서 38.7 kW)을 확인할 수 있다.

3.2. 공기압축기 운전데이터 취득 환경

2023년 시화국가산단 내 금속 도금단지에는 전동기 용량 75kW 공기압축기 2대와 37kW 공기압축기 1대가 설치되었다. 후단에는 압축공기 공급의 안정성과 수요 변동 완충을 위해 리시버 탱크 1대를 두었으며, 수분 제거를 목적으로 냉동식 드라이어 1대와 흡착식

Table 1. Air compressor specification

Division	Unit	No. 1	No. 2	No. 3
Type	-	Screw	Screw	Screw
Discharge flow	Nm ³ /min	11.84	6.68	11.84
Discharge pressure	kg/cm ²	7.5	7.0	7.5
Power	kW	75+1.5	37	75+1.5
Consumption rate	kWh/Nm ³	0.106	0.092	0.106
Rotation speed	rpm	3600	3600	3600
Power supply	V	380V/60Hz	380V/60Hz	380V/60Hz
Manufacturer	-	Kaeser	Kaeser	Kaeser
Date	yy.mm	2022.11	2020.06	2022.11

에어드라이어 1대를 각각 설치하였다. 생산된 압축공기는 도금 처리 제품의 건조와 도금조 내부 기화된 약품의 인체 흡입 방지를 위한 블로잉 용도로 활용하고 있다.

Fig. 10.은 공기압축기 운전데이터 취득을 위한 통신시스템 구성을 나타낸 것이다. 계측 장비로는 압축공기 유량계, 압축공기 압력과 온도 보정을 위한 압력계 및 온도계, 운전 전력 모니터링을 위한 전력량계, 그리고 진동 및 밸런싱 평가를 위한 가속도계를 설치하였다. 각 센서에서 수집된 신호는 RS-485 케이블과 MODBUS RTU 프로토콜을 통해 실시간으로 PLC에 전달되며, 이후 PLC와 도금단지 로컬서버는 UTP 케이블로 연결되어 1초 간격으로 데이터를 FEMS (Facility Energy Management System) 데이터베이스(DB)에 저장하고 체계적으로 관리된다. 본 연구에서 개발한 프로그램은 FEMS 서버에 구축된 데이터베이스로부터 필요한 운전데이터를 호출하여 분석 및 모니터링에 활용할 수 있도록 구성되었다.

데이터 취득 과정 중 압축기 흡입측에서 직접 공기량을 계측하는 것은 설치 여건과 측정 불확실성으로 인해 어려운 경우가 많다. 이에 본 연구에서는 압축기 출구측에 플로우미터를 설치하고, 계측된 유량을 흡입조건으로 환산하여 적용하였다. 압축공기의 유량은 출구측에서 계측할 수 있으나, 성능평가 및 비교를 위해서는 이를 흡입조건 기준으로 환산하는 것이 필수적이다. 이는 압축기의 본질적 성능 지표가 흡입유량에 의해 정의되기 때문이며, ISO 1217 및 KS B 6351 등 국제적으로 통용되는 규격에서도 흡입조건 기준의 체적유량으로 규정하고 있다. 이를 바탕으로 출구측에서 계측한 노멀상태 공기량을 압력·온도 보정을 통해 흡입조건 기준의 체적유량으로 환산하여 적용하였다. 식 (11)에서 Q_n 은 노멀상태의 공기량을, Q_t 는 흡입조건으로 환산된 공기량을 의미하고 t 는 측정 시 흡입공기의 온도, P_t 는 흡입공기의 절대압력이다.

$$Q_t = Q_n \times \frac{273 + t}{273} \times \frac{760}{P_t} \quad (\text{Eq. 11})$$

측정의 정확도를 위해 유체의 온도, 밀도, 조성, 점도 등의 영향이 적고 유량측정범위가 넓어 다양한 환경에 적용이 가능한 Table 2.의 볼텍스 플로우미터를 Fig. 11.과 같이 설치하였다.

직관의 거리를 충분히 확보해야 정확한 측정이 가능하기 때문에

Table 2. Vortex flowmeter specification

Division	Unit	No. 1	No. 2	No. 3
Type	-	Wafer	Wafer	Wafer
Size	mm	40	25	40
Max. Flow	m ³ /h	290	130	290
Min. Flow	m ³ /h	13	8	13
Pressure	MPa	1.0	1.0	1.0
Accuracy	%	±1.0	±1.0	±1.0
Power supply	V	DC 12~45	DC 12~45	DC 12~45
Manufacturer	-	Oval	Oval	Oval
Date	yy.mm	24.04	24.04	23.12

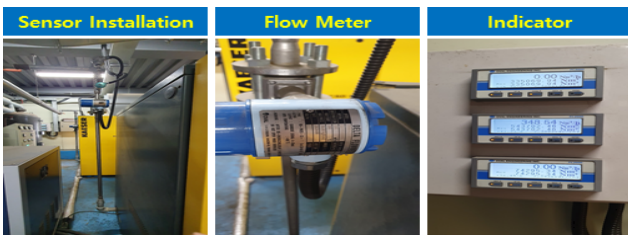


Fig. 11. Air compressor system instrument installation status

상류에 10D, 하류에 5D 이상을 확보하여 설치하였고 온압보정을 위해 RTD 온도센서와 압력트랜스미터도 추가적으로 설치하였다. 유체의 유속분포는 직관(직선 배관)의 중심에서 가장 빠르고, 벽 쪽에서는 느려지는 포물선 형태(fully developed profile)가 일반적인데 배관의 굴곡(엘보), 밸브, 펌프, 분기 등 장애물이 있으면, 유속이 고르게 퍼지지 않고 휘어지거나, 소용돌이가 생겨 불균일해진다. 이러한 이유로 유량계 전방에 유체의 흐름을 정돈시켜 분포를 고르게 해주는 플로우컨디셔너도 설치하였다.

4. 이상감지 및 성능분석

4.1. 공기압축기 운전현황 조사

공기압축기 운전데이터는 2025년 4월 28일부터 5월 29일까지 약 1개월간, 1초 간격으로 연속 취득하였다. 수집된 주요 항목은 순시 전력, 토출압력, 토출온도, 흡입온도, 압축공기량 등 총 5개 변수로 구성된다. 취득된 데이터는 항목별로 약 2,592,000건(1개월×30일×24시간×60분×60초)이며, 이를 통해 운전특성을 분석하였다.

도급단지에서 보유한 3대의 공기압축기 중 변속형 압축기 1대가 가동 중이며, 그 운전 특성은 Fig. 12.에 나타난 바와 같이 단속 운전 형태를 보였다. 압축기 동력은 부하 운전 시 급격히 상승한 후, 무부하 또는 정지 상태에서는 약 0~10kW 수준까지 하강하는 주기가 반복적으로 관찰되었다. 이러한 운전 패턴은 저장탱크(리시버)의 압력이 목표치에 도달하면 압축기가 정지(무부하)하고, 압력이 하강하면 재기동하는 제어 방식과 일치한다. 한편, 이와 같은 빈번한 ON/OFF 현상은 일반적으로 압축기 용량이 과대하거나 리시버 용량이 부족하여 압축공기 저장 능력이 낮을 때, 또는 압력 설정 폭이 과도하게 좁을 때 발생한다. 이러한 현상은 Fig. 13. 압축공기 사용량 분석을 통해서도 확인할 수 있다. 압축공기 생산 패턴을 분석한

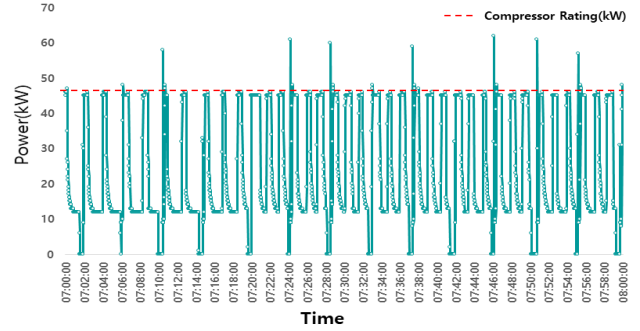


Fig. 12. Air compressor power consumption pattern monitoring

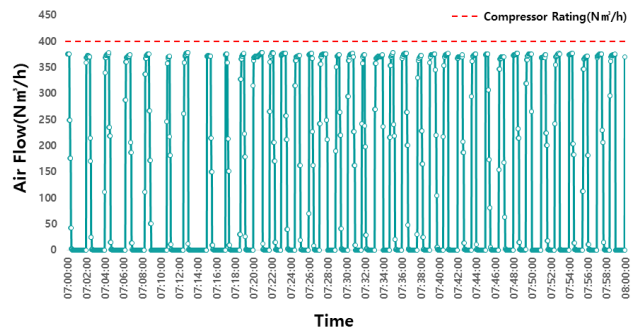


Fig. 13. Compressed air production pattern monitoring

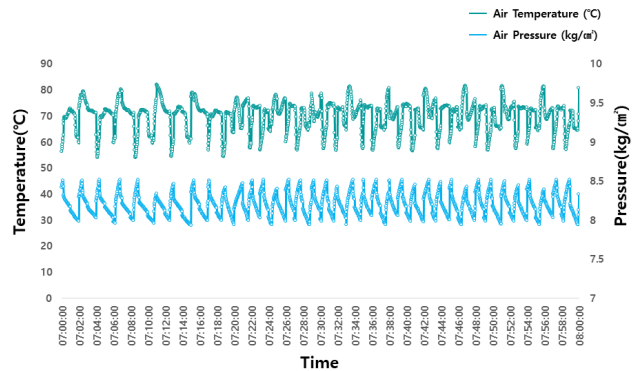


Fig. 14. Exhaust air temperature and pressure monitoring

결과, 약 1~2분 단위의 짧은 주기로 ON/OFF 사이클이 반복되는 현상이 확인되었다. 이는 도금 처리시설의 공기 수요 변동에 압축기가 즉각적으로 반응함에 따라 사이클링 운전 빈도가 불필요하게 증가한 것으로 해석된다. 특히 압축기 모터가 반복적으로 ON/OFF를 수행할 경우, 모터와 전기계통에 과도한 기계적·전기적 스트레스와 무부하 운전 손실로 인해 압축기 효율이 감소하는 문제를 야기할 수 있다.

대부분의 산업 현장에서는 공기압축기의 생산 압력을 약 6.0~7.5 kg/cm²g 범위로 관리한다. 그러나 본 사례(Fig. 14.)에서는 7.0~8.5 kg/cm²g($\Delta P=1.5\text{kg/cm}^2$)로 상대적으로 높게 설정되어 있어 에너지 소비가 증가하는 것으로 판단된다. 또한, 공기압축기의 출구 온도는 이론적 단열 압축 시 200°C 이상에 도달할 수 있으나, 실제 운전은 오일과 표면손실로 인해 등온에 가까운 폴리트로픽 과정으로 본 사례에서는 평균 약 73.6°C 수준으로 유지되고 있다.

4.2. 공기압축기 이상감지 및 성능분석

공기압축기 운전 현황을 분석한 결과, 명확한 휴지 구간 없이 24 시간 연속 운전되며 주·야간 모두에서 변동 폭(최대-최소 범위)이 크게 나타났고, 간헐적으로 완전 정지 구간도 확인되었다. 운영 패

Table 3. Test data for fault detection

Division	Pressure (kg/cm ²)	Power (kW)	Dis.Temp. (°C)	Suc.Temp. (°C)	Flow (Nm ³ /h)
CASE 1	8.305	45	77.1	18.2	366.67
CASE 2	8.391	45	80	18	368.08
CASE 3	8.308	46	83.4	17.4	11.59
CASE 4	8.498	46	61	16.8	372.96
CASE 5	2	47	69.9	16.6	330
CASE 6	8.169	47	69.1	18.1	112.86
CASE 7	3	45	30	18.1	375.95
CASE 8	8.28	45	81.6	17.2	3.81
CASE 9	8.469	46	60	16.9	373.45
CASE 10	8.494	46	82.4	17.2	369.71
CASE 11	8.464	46	65.6	14.9	377.22
CASE 12	8.383	45	63	15.1	376.24
CASE 13	8.237	22	71.5	14.7	206.36
CASE 14	8.199	43	70.7	14.7	9.99
CASE 15	8.096	43	72.8	15.2	29.93
CASE 16	8.45	46	20	14.8	377.27
CASE 17	8.323	44	69	14.6	260.83
CASE 18	8.204	46	70.6	14.6	46.58
CASE 19	8.432	46	65.8	14.8	377.64
CASE 20	8.254	25	70.9	14.6	221.43
CASE 21	8.357	45	63.4	20.9	366.72
CASE 22	8.431	45	67.7	5	368.66
CASE 23	8.419	46	72.5	20.6	368.95
CASE 24	8.282	46	76.8	20.5	162.35
CASE 25	8.27	46	76.2	20.4	127.34
CASE 26	8.25	44	75.6	20.5	7.73
CASE 27	8.22	44	80.3	21.6	109.8
CASE 28	8.357	45	85.5	21.5	362.33
CASE 29	8.289	45	62.7	21.4	361.41
CASE 30	8.424	46	72.5	20.7	351.41
CASE 31	8.291	45	62.7	5	365.69
CASE 32	8.365	45	66.4	17	368.09
CASE 33	8.456	45	69.4	16.9	369.04
CASE 34	8.426	46	69.9	16.7	370.94
CASE 35	8.303	45	73.2	16.7	236.79
CASE 36	8.241	23	74.4	16.7	208.2
CASE 37	8.17	45	72.7	16.7	1.18
CASE 38	8.514	46	68.9	16.7	369.95
CASE 39	8.259	45	63.4	17	358.32
CASE 40	8.239	23	74.1	16.7	214.12
CASE 41	8.44	47	60.9	18	231.78
CASE 42	8.365	45	63.5	17.9	215.07
CASE 43	8.517	46	74.1	19.2	365.25
CASE 44	8.501	46	59.2	18.2	202.21
CASE 45	8.31	45	67.7	18	142.4
CASE 46	8.513	45	74.7	19.3	364.9
CASE 47	7.98	14	63	19.9	292.04
CASE 48	8.318	27	66.4	18	327.2
CASE 49	8.294	48	68	17.9	106.45
CASE 50	8.178	48	68.9	19.7	296.77

턴이 일정하다는 정상성(stationarity) 가정하에 총 약 2,592,000개의 시계열 관측치 중 약 20,000개를 단순 무작위로 추출하여 학습 데이터로 사용하였다.

이후 이상 감지 평가를 위해 흡입온도, 토출온도, 토출압력, 압축 공기 유량, 순시전력의 5개 변수에 대해 임의의 Table. 3. 테스트 데이터 50개를 생성하고, 이를 기 학습 모델의 주성분 공간(Principal Component Space)에 투영하였다. Fig. 15.에서와 같이 일부 데이터 점이 학습 분포의 중심(평균 벡터)으로부터 유의하게 이격되는 양상이 관찰되었다.

Fig. 16.에서 두 통계량 모두에 대해 95% 및 99% 관리상한선(임계값)을 기준으로 이상치 여부를 판별할 수 있다. 관리상한선을 초과하는 데이터들은 통계적으로 정상 범위에서 벗어난 값으로 간주되며, 실제로 두 그래프 모두에서 여러 샘플이 99% 관리상한선을 반복적으로 초과하는 현상이 확인되었다.

Hotelling's T²와 Q-통계량 검정을 통해 기준에 학습된 데이터기준으로 카이제곱분포의 유의수준 0.01 이상의 범위(T² 11.345, SPE 1.546)에 포함되면 이상값으로 판정되며 Table. 4.처럼 이상치 발생 시 1로 비정상값 발생여부, Fault tag, 재계산값(예측치), 센서부위

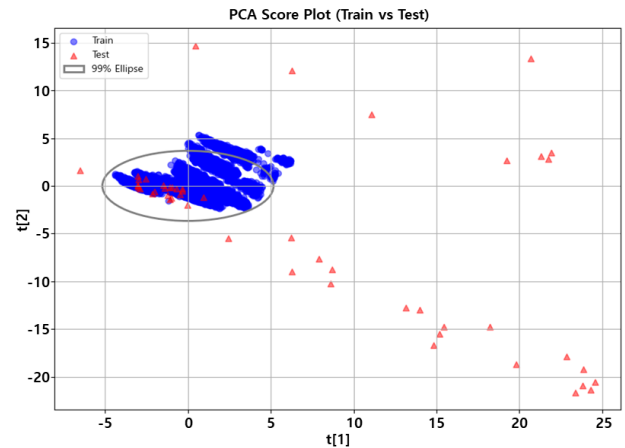


Fig. 15. Data distribution in principal component space

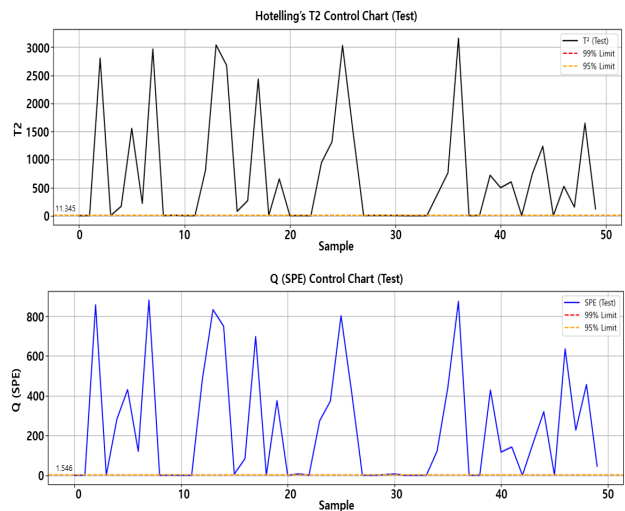


Fig. 16. Hotelling T² chart and Q-statistic (SPE) chart analysis results

의 잔여수명을 출력한다.

Fig. 17.처럼 정상범위를 벗어난 주성분값은 기여도 분석을 통해 원인을 추적할 수 있는데 케이스별로 5개 변수 중 이상치의 가장 큰 원인을 나타내고 알고리즘에 따라 최적값으로 대체함으로써 이상감

Table 4. Fault detection results

Division	Sensor Fault	Fault Tag	Fault Value	Optimi. Value	Remain. Month
CASE 1	0	-	-	-	-
CASE 2	0	-	-	-	-
CASE 3	1	Flow	11.59	384.27	0.21
CASE 4	0	-	-	-	-
CASE 5	1	Dis.Press.	2.0	9.59	0.46
CASE 6	1	Flow	112.86	375.95	0.30
CASE 7	1	Dis.Press.	3.0	7.92	0.71
CASE 8	1	Flow	3.81	381.86	0.21
CASE 9	0	-	-	-	-
CASE 10	0	-	-	-	-
CASE 11	0	-	-	-	-
CASE 12	0	-	-	-	-
CASE 13	1	Power	22	42.20	0.51
CASE 14	1	Flow	9.99	377.09	0.21
CASE 15	1	Flow	29.93	378.28	0.22
CASE 16	1	Dis.Temp.	20.0	65.60	1.65
CASE 17	1	Flow	260.83	376.96	0.68
CASE 18	1	Flow	46.58	382.83	0.23
CASE 19	0	-	-	-	-
CASE 20	1	Power	25	42.49	0.59
CASE 21	0	-	-	-	-
CASE 22	1	Suc.Temp.	5	18.49	3.68
CASE 23	0	-	-	-	-
CASE 24	1	Flow	162.35	372.86	0.38
CASE 25	1	Flow	127.34	372.79	0.32
CASE 26	1	Flow	7.73	368.57	0.22
CASE 27	1	Flow	109.8	369.52	0.30
CASE 28	0	-	-	-	-
CASE 29	0	-	-	-	-
CASE 30	1	Flow	351.41	368.23	4.75
CASE 31	1	Suc.Temp.	5	18.25	3.74
CASE 32	0	-	-	-	-
CASE 33	0	-	-	-	-
CASE 34	0	-	-	-	-
CASE 35	1	Flow	236.79	377.06	0.57
CASE 36	1	Power	23	42.09	0.54
CASE 37	1	Flow	1.18	377.87	0.21
CASE 38	0	-	-	-	-
CASE 39	0	-	-	-	-
CASE 40	1	Power	23	42.19	0.54
CASE 41	1	Flow	231.78	368.22	0.58
CASE 42	1	Flow	215.07	367.13	0.52
CASE 43	0	-	-	-	-
CASE 44	1	Flow	202.21	364.20	0.49
CASE 45	1	Flow	142.4	370.26	0.35
CASE 46	0	-	-	-	-
CASE 47	1	Power	14	42.79	0.36
CASE 48	1	Power	27	44.24	0.60
CASE 49	1	Flow	106.45	376.45	0.29
CASE 50	1	Flow	296.77	373.94	1.03

지 모델이 완성된다. 특히, 주성분 기반의 기여도 분석은 여러 변수의 상관관계를 함께 고려함으로써, 실제로 가장 중요한 이상 변수를 정확히 식별하고, 모델의 신뢰성을 높이는 데 크게 기여한다.

4.3. 공기압축기 성능분석 오차율 검증

다음 성능분석 오차율 검증을 위해 Fig. 18. 절차에 따라 이상감지용 샘플값을 성능분석 프로그램 및 KS B 6351 표준수식에 반영하고 산출결과를 비교하였다. 스탠다드결과와 DW-SIM결과오차발생 범위를 다각도로 검증하기 위해 평균제곱근오차, 평균절대오차, 평균절대오차비율, 결정계수와 같은 정확도 지표를 활용하여 다면적으로 평가하였다.

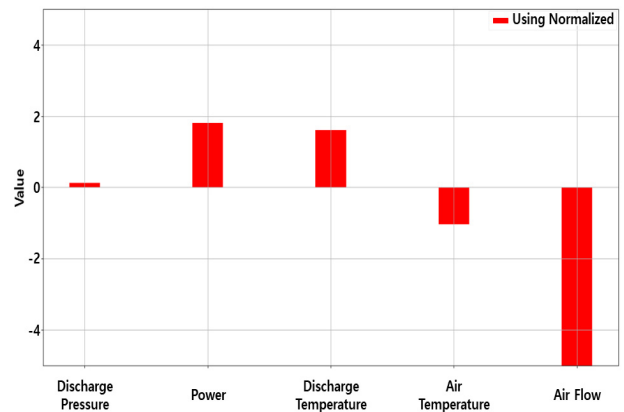


Fig. 17. Contribution analysis for outliers

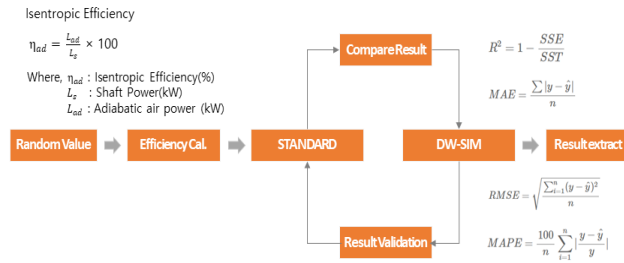


Fig. 18. The process for validating error rates

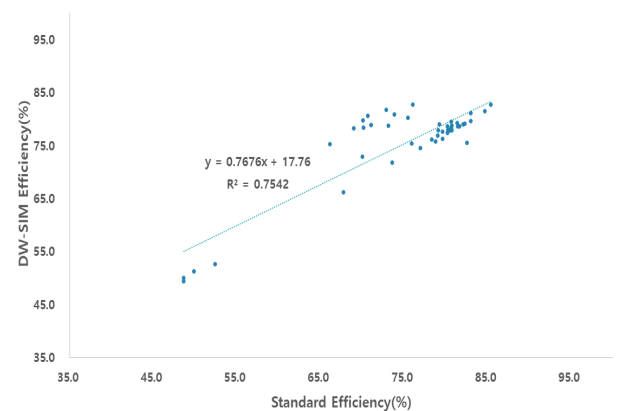


Fig. 19. Correlation between standard and DW-SIM efficiency analysis results

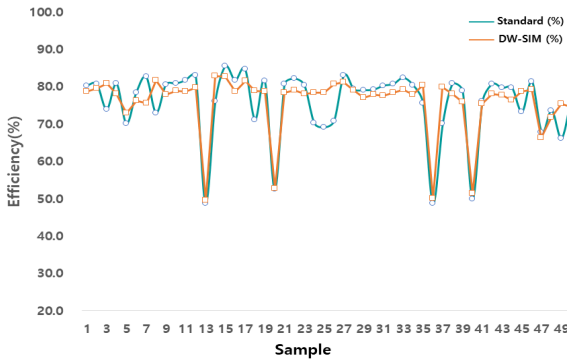


Fig. 20. Comparison of standard and DW-SIM efficiency analysis results

Fig. 19.에서 성능분석모델의 결정계수(R^2)는 약 0.754로 평가되었다. 이는 DW-SIM 기반 예측모델이 실제 효율값의 변동성을 약 75.4% 수준으로 설명할 수 있음을 의미하며, 예측값이 실제값과 양호한 수준의 상관성을 보임을 나타내지만, 여전히 잔차가 존재하므로 입력 데이터 품질 및 변수의 보정 (센서교정, 이상치 처리)등을 통해 정확도의 추가적 향상이 요구된다.

Fig. 20.은 스탠다드결과와 DW-SIM의 효율 산출값을 비교한 결과이다. 두 방법 간 오차 지표는 RMSE=4.41, MAE=3.55로 산출되어 절대 오차 규모가 크지 않았다. 또한 MAPE=4.76%로, DW-SIM 결과는 스탠다드 결과 대비 평균 5% 미만의 상대 오차를 보여 비교적 높은 정확도를 가지고 있음을 확인할 수 있었다.

5. 결론

본 연구는 산업현장에서 널리 사용되는 스크류 공기압축기의 효율과 이상징후의 조기 탐지를 위해, 기계학습 기반 예지정비 시스템과 주성분 분석을 적용한 이상징후 탐지 알고리즘을 개발하였다. 1개월 이상 축적된 실제 운전 데이터를 바탕으로, 주요 운전 변수(흡입 온도, 토출압력, 토출온도, 유량, 전력 등)를 주성분으로 축약하고, 주성분공간에서의 Hotelling's T^2 통계량과 Q-통계량을 활용하여 이상 여부를 진단하였다. 두 통계량을 통해 이상 데이터를 효과적으로 판별하고, 기여도 분석을 통해 원인 진단도 수행하였다.

또한, 표준 규격(KS B 6351)과 비교·평가하여 R^2 , RMSE, MAE, MAPE 등 다양한 지표에서 비교적 양호한 예측 정확도를 확인하였다. 본 시스템은 설비의 신뢰성 향상과 운영 효율성 증대에 기여할 수 있다고 판단되며, 장기 데이터를 바탕으로 향후 다양한 산업 환경에 적용가능한 실시간 진단 시스템으로 확장할 예정이다.

Acknowledgement

본 논문은 한국에너지기술평가원의 지원을 받아 수행한 연구과제 결과의 일부이다(과제번호: 2410005867).

References

- [1] 정병용, 신연봉, 중소기업의 재해 특성 및 안전관리 정책에 관한 연구, 산업공학, 제12권 제4호, 1999, pp. 595-601. // (B.Y. Jeong, Y.B. Shin, Accident characteristics and safety management policy in the small manufacturing industry, IE Interfaces, 12(4), 1999, pp.595-601.)
- [2] 백종엽, 강운진, 강경식, 중소제조업체의 설비효율향상 방안에 대한 연구, 대한안전경영과학회 춘계학술대회 논문집, 2010, pp.219-2300. // (J.Y. Baek, Y.J. Kang, G.S. Kang, Improving overall equipment effectiveness (OEE) in Korean small and medium manufacturing industries, Proceedings of the Safety Management and Science Conference, 2010, pp.219-230.)
- [3] 유창규, 최상욱, 이인범, 공정 모니터링 기술의 최근 연구 동향, 화학공학 제46권 제2호, 2008, pp.233-247. // (C.G. Yu, S.U. Choi, I.B. Lee, Recent research trends of process monitoring technology: State-of-the art, Korean Chemical Engineering Research, 46(2), 2008, pp.233-247.)
- [4] Theodora Kourti, John F. MacGregor, Process analysis monitoring and diagnosis using multivariate projection methods, Chemometrics and Intelligent Laboratory Systems, 28(1), 1995, pp.3-21.
- [5] 허태영 외 3인, 다변량 통계기법을 활용한 실시간 수질이상 유무 판단 시스템 개발, 대한환경공학회지 제37권 제3호, 2015, pp.137-144. // (T.Y. Heo et al., Development of real-time water quality abnormality warning system for using multivariate statistical method, Journal of Korean Society of Environmental Engineers, 37(3), 2015, pp.137-144.)
- [6] H.S. Lee, Y.K. Jo, J.P. Jeong, Multivariate PCA-based composite criteria evaluation method for anomaly detection in manufacturing data, 2024 26th International Conference on Advanced Communications Technology, 2024, pp.107-115.
- [7] 김혜진 외 3인, 시간 영역 통계 기반 웨이퍼 이송 로봇의 고장 진단, 문화기술의 융합 제10권 제4호, 2024, pp.663-668. // (H.J. Kim et al., Fault diagnosis of wafer transfer robot based on time domain statistics, The Journal of the Convergence on Culture Technology, 10(4), 2024, pp.663-668.)