

플랜트 운전이상 조기 감지를 위한 알람 고도화 기법

(주)테크다스 고흥철

I. 서론

최근 센서 기술의 발달, 사물인터넷 기술의 발전, 그리고 센서로부터 나오는 빅데이터들을 빠르게 처리할 수 있는 컴퓨터의 발전으로 플랜트 이상이나 고장을 감지하는 기법이 다양하게 시도되고 있다. 특히 데이터 해석 기술을 이용하여 공장의 운전 데이터 속에 포함된 패턴을 찾고 숙련된 운전자의 경험적 지식을 디지털 전환 기술로 재구성하여 기존의 DCS(Distributed Control System)나 PLC(Programmable Logic Control)에서는 어려웠던 고장이나 공정 이상을 감지한다.

먼저 본 고에서 다루고자 하는 알람 고도화 솔루션과 상태 기반의 유지보수(Condition Based Maintenance) 솔루션을 구분할 필요가 있다. 두 솔루션의 적용 기술은 유사하지만, 전자는 공장 운전을 대상으로 하고, 후자는 설비나 장치의 유지보수를 대상으로 한다. 국내외적으로 고장 예지 솔루션은 많으나, 대부분 유지보수 목적이다. 운전 중 발생하는 이상징후 탐지 솔루션은 많지 않다.

전통적인 DCS나 PLC의 알람 체계에서는 정상과 비정상을 구분하는 경계를 설정하여 플랜트의 고장의 발생을 감지한다. 아래 그림에서 파란색은 정상, 붉은색은 비정상을 나타낸다고 가정할 때, 제어 경계 설정값을 기준으로 정상과 비정상을 판단하면, 제어 경계 우측의 파란색 영역은 정상임에도 비정상으로 판단하는 오류를 범하고, 반대로 제어 경계 좌측의 붉은색 영역은 비정상임에도 정상으로 판단하는 오류를 범한다.

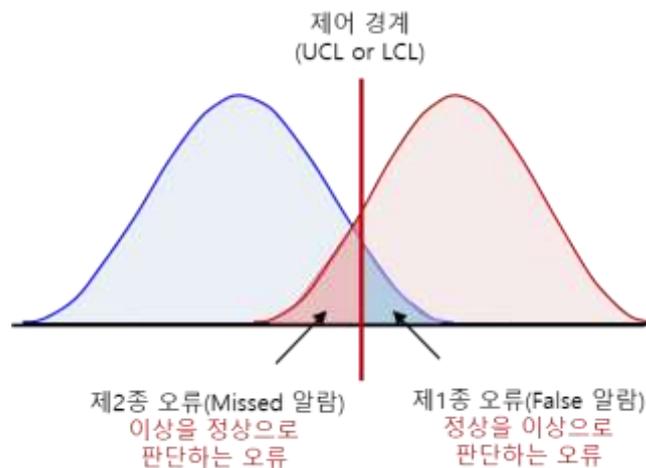


그림 1. 전통적인 공장제어 시스템에서의 제어 경계와 제1/2종 오류

전자, 즉 정상임에도 비정상으로 판단하는 오류를 제1종 오류, 후자, 즉 비정상을 정상으로 판단하는 오류를 제2종 오류라 부른다. 제1종 오류를 줄이려면 제어 경계를 우측으로 이동시켜야 하지만, 이 때 제2종 오류가 반대 급부적으로 증가한다. 반대로 제2종 오류를 줄이려면 제어 경계를 좌측으로 이동시켜야 하지만, 이 때도 제1종 오류가 증가한다. 따라서 제어 경계를 결정하는 방법으로는 오류를 회피할 수 없으며 우리가 할 수 있는 최선의 방법은 제1종 오류와 제2종 오류의 적절한 비율을 고려하는 정도이다.

전통적인 알람 체계를 개선하기 위한 알람 고도화 기법은 궁극적으로 제1종 오류와 제2종 오류를 동시에 줄이는데 목표를 둔다. 이의 개념을 그림으로 나타내면 아래와 같다. 그림 1과 비교하여 제어 경계에 의존하지 않고 근본적으로 제1종 오류와 제2종 오류의 발생 확률을 감소시킨다.

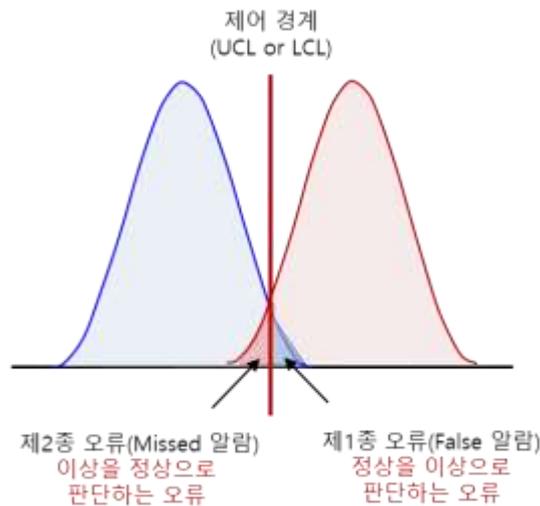


그림 2. 알람 고도화 기법을 통한 제1/2종 오류 감소

II. 알람 고도화 기법

1) Shewhart 알람 체계

DCS나 PLC의 경보체계는 센서 값이 경계치를 벗어날 때 경보를 발행한다. 창시자의 이름을 따서 Shewhart 관리도(Control Chart)라 부른다. 현대화된 대부분의 DCS나 PLC가 이 경보체계를 채택하지만, 1920년대에 고안된 매우 오래된 방식이다.

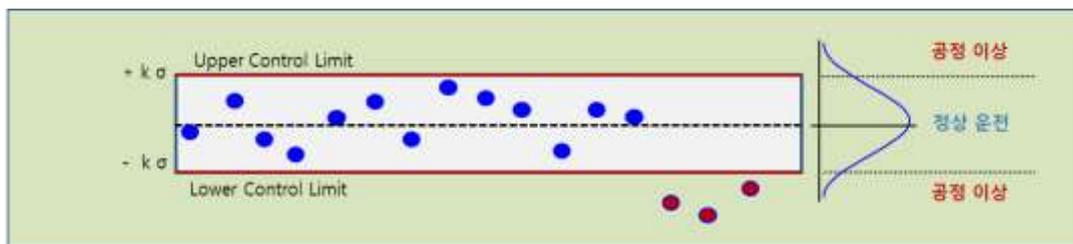


그림 3. Shewhart 관리도

Shewhart 제어도에서 경계 설정값 $k\sigma$ 의 k 는 센서 측정값의 표준편차 σ 의 배수를 의미한다. k 를 크게 하여 정상 구간을 늘리면 정상을 이상으로 판단하는 제1종 오류를 줄일 수 있지만, 이상을 정상으로 판단하는 제2종 오류가 증가한다. 반대로 k 를 작게 하여 정상 구간을 줄이면 이상을 정상으로 판단하는 제2종 오류를 줄이는 대신에 정상을 이상으로 판단하는 제1종 오류가 증가한다. 따라서 최적의 k 를 결정하여 오류를 합리화(Rationalization)하는 것이 필요하다. 제어 경계의 크기에 따른 확률은 그림 4와 같다.

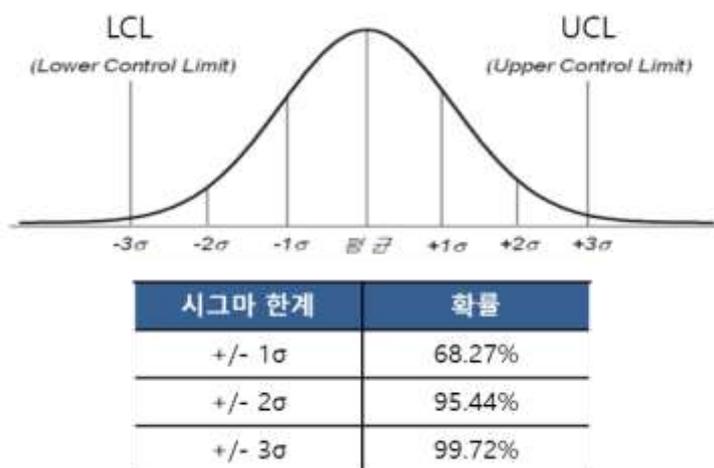


그림 4. 제어 경계에 따른 포함 확률

Shewhart 관리도에서 경보 경계는 영국의 EEMUA¹가 제시한 표 1의 가이드라인을 참고하여 결정할 수 있다. EEMUA 가이드라인은 운전자의 업무 부하를 고려하여 제시한 것으로, k 를 통해 경보 빈도수를 예측하여 이를 적용할 수 있다. 운전자가 관리하는 Tag 수와 그림 4의 확률을 고려하여 경계를 설정한다.

표 1. EEMUA 경보 설정 가이드라인

경보 수준	설정 기준
Emergency Priority	매우 드물게
High Priority	Shift 당 5회 이하
Medium Priority	시간 당 2회 이하
Low Priority	시간 당 10회 이하

¹ Engineering Equipment & Materials Users Association

2) CUSUM 기법

CUSUM(Cumulative SUM) 기법은 단변량 통계기법의 하나로 설정값(SP)과 측정값(PV)의 차이를 누적 계산하여 누적 합계가 일정 범위를 벗어나면 알람을 발행한다. Shewhart 기법이 측정값의 변화가 클 때 변별력이 우수한 반면에 CUSUM 기법은 측정값이 설정값으로부터 편향되어 나타날 때 변별력이 우수하다. CUSUM 통계량은 다음 식과 같다.

$$S_m = \sum_{i=1}^m (x_i - \mu_0)$$

여기서 m은 샘플의 개수를 의미하며, \bar{x} 는 단위 시료의 평균을 의미한다. μ_0 는 시료 전체의 평균을 의미한다. 그림 5의 위 그림과 같이 평균의 변화가 발생하는 경우 식별이 용이하며, 밸브 개도율이나 제어기의 출력이 포화되어 있거나 시설에 변형될 때 나타나는 현상이다.

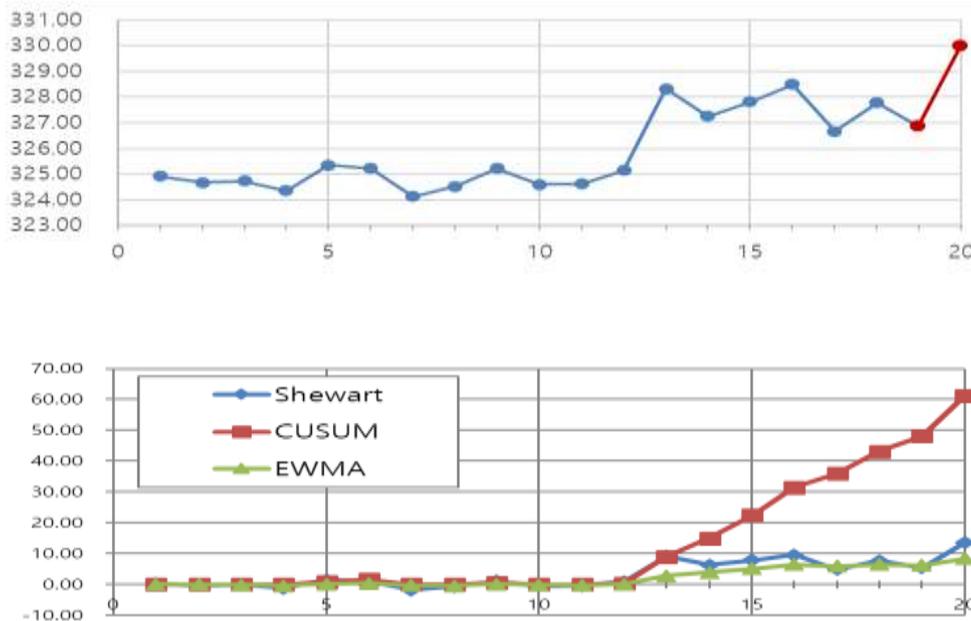


그림 5. CUSUM 변별력; (위) PV 트렌드, (아래) 붉은색 실선 CUSUM 트렌드

3) EWMA 기법

EWMA(Exponential Weighted Mean Average) 기법은 단변량 통계기법의 하나로 과거와 현재의 측정값을 분할 반영하고, 일정 범위를 벗어나면 알람을 발행한다. 지수함수 형태의 특성을 가져서 붙인 이름이다. Sensor Drift, Fouling, Deactivation, 누설처럼 연속적으로 값이 커지거나 작아지는 변화에 대해 변별력이 우수하다. 그리고 노이즈가 커서 상하 경계를 넓게 필요로 하는 경우에 노이즈를 제거하고 좁은 경계에서 관측값을 관리하고자 할 때 필요하다.

$$Z_t = \lambda x_t + (1-\lambda)Z_{t-1}$$

여기서 Z_t 가 출력값이며, t 는 현재 시간을 $t-1$ 은 Scanning Time 이전 시간을 의미한다. 파라미터 λ 의 크기를 조절하여 현재 측정값의 반영 비율을 조절한다. λ 가 1이면 현재 측정값을 그대로 출력하며, 0이면 현재 측정값을 무시한다. 보통 0.3 ~ 0.8 사이의 값을 갖는다. 그림 6의 위 그림과 같이 값이 서서히 커질 때 변별력이 우수하다.

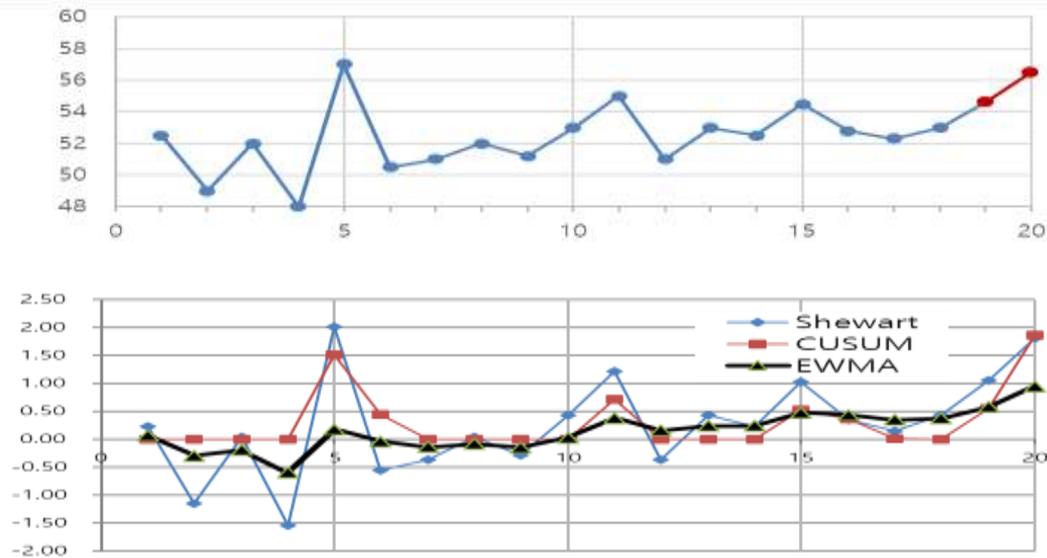


그림 6. EWMA 변별력; (위) PV 트렌드, (아래) 붉은색 EWMA 트렌드

4) FPM 기법

FPM(First Principle Model)은 이론적인 수식을 이용하여 측정값을 예측하며, 예측값과 측정값의 차이를 통해 공정이상을 감지한다. 예측값과 측정값의 차이가 일정 규모 이상(2 시그마, 또는 3 시그마)으로 커지면 알람을 발행한다. 단변량 통계기법과 다변량 통계기법이 모두가 변별력이 부족할 때 FPM 기법이 유용할 수 있다. 가장 간단한 사례는 용기의 액위(Level) 고착이다.

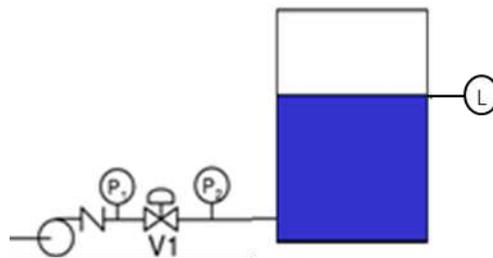


그림 7. 펌프 및 용기의 액위 센서

DCS 나 PLC 알람 체계에서는 위의 예에서 용기의 액위 측정이 고착되더라도 알람 경계 범위 내에 존재하면 정상으로 판단한다. 단순 물질수지 식의 FPM 을 이용하여 액위 고착을 감지한다. 펌프가 가동하면 액위는 지속적으로 증가하는데, 만약에 액위 측정기가 고착되면 FPM 계산값과 액위 측정값의 차이가 커지게 되어 액위 센서의 고착을 감지한다. 그림 8 이 대표적인 액위 측정기 고착 사례이다.

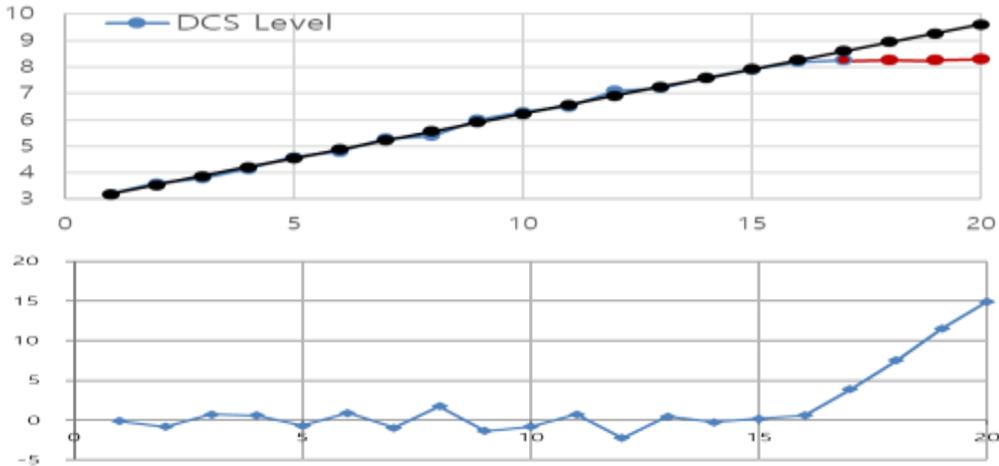


그림 8. FPM 변별력; (위) 붉은색 PV 트렌드와 검정색 FPM 트렌드, (아래) PV 와 FPM 차이 트렌드

5) PCA 기법

PCA(Principal Component Analysis)는 다변량 통계기법이다. 과거 운전 데이터에 대해 머신러닝을 실시하여 변수 간의 패턴을 찾고, 실제 관측 데이터가 학습했던 패턴과 일치하는지 여부를 통해 공정이상을 감지한다. 이를 개념적으로 표현하면 아래 그림과 같다. 그림처럼 Δt 시간만큼 조기에 공정이상을 감지한다.

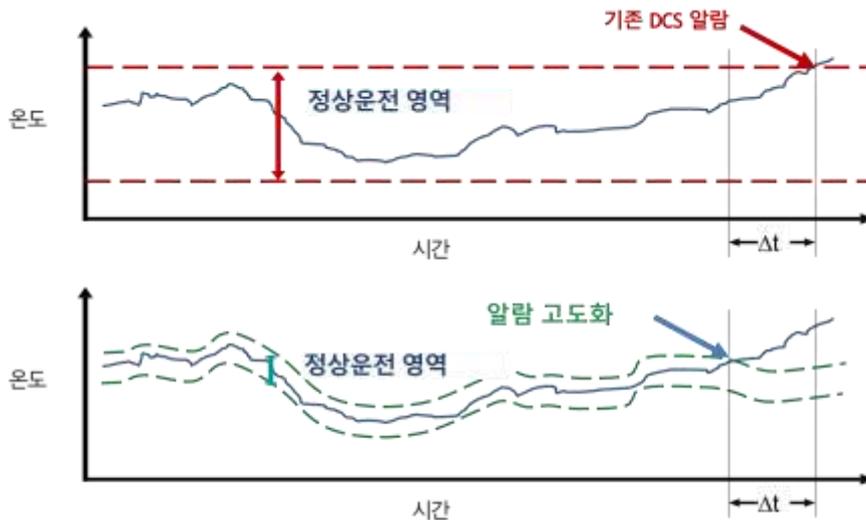


그림 9. 알람 고도화 기법의 효과

PCA에서는 Hotelling T^2 와 SPE(Squared Prediction Error)로 관측값의 정상 여부를 감시한다. Hotelling T^2 는 관측값이 좌표의 중심에서 벗어난 정도를 나타내는 통계량이다. 변수의 개수가 작을 때 효과적이므로 관측값을 소수의 주성분으로 표현하는 PCA에서 적합하다. Hotelling T^2 는 다음 식으로 계산한다.

$$T^2 = \sum_{k=1}^K \left(\frac{t_k}{s_k} \right)^2$$

여기서 K는 주성분의 개수를, s_k 는 k번째 주성분의 표준편차를 의미한다. 관측값의 T^2 가 학습 데이터의 95%(또는 99%)보다 크면 공정이상으로 판단한다. 피드 변화와 같이 학습에 사용했던 측정값과 다른 값이 입력될 때 이상을 감지한다 주성분 좌표 t_1 과 t_2 를 이용하여 설명하면 아래 그림의 좌측은 관측값이 경계 범위 안쪽에 있으므로 정상이고, 우측은 관측값이 경계 범위 밖에 있으므로 공정이상으로 판단한다.

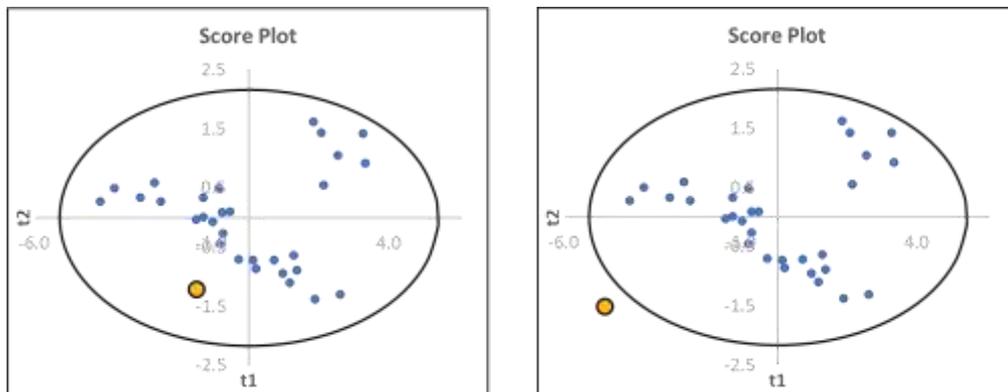


그림 10. Hotelling의 T^2 개념도

Hotelling의 T^2 와 한 쌍을 이루는 SPE는 PCA 모델의 잔차($\mathbf{E} = \mathbf{X} - \mathbf{TP}^T$)의 제곱의 합을 나타내는 통계량이다. 관측값의 SPE가 학습 데이터의 95%(또는 99%)보다 크면 공정이상으로 판단한다. SPE는 관측값을 주성분에 수직으로 투영한 값 \mathbf{TP}^T 과 \mathbf{X} 의 차이의 제곱으로 아래 식으로 정의한다.

$$SPE = \sum_{n=1}^N (x_n - \hat{x}_n)^2$$

여기서 N은 X변수의 개수를, x_n 는 n번째 변수를, \hat{x}_n 는 n번째 변수의 예측모델(\mathbf{TP}^T)로 계산한 값을 의미한다. SPE를 그림으로 설명하면 아래 그림과 같고, PCA 모델과 관측값의 차이를 의미한다. 학습 데이터의 패턴과 관측값의 패턴의 차이라 클수록, 또는 노이즈가 클수록 큰 값을 갖는다.

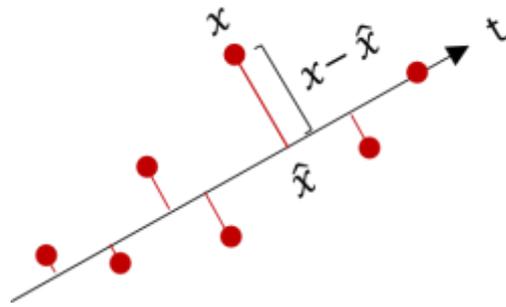


그림 11. SPE 개념도

공장 데이터는 공선성(Collinearity)이 높으므로 서로 독립적인 잠재변수를 도출하여 이용하는 PCA가 알람 고도화 기법으로 장점이 있다. 아래 그림은 SPE의 간단한 사례로 두 변수의 패턴, 즉 증가와 감소의 패턴이 기존의 것과 다를 경우, 즉 하나는 증가하고 다른 하나는 감소하여 패턴을 따르지 않을 때 공정이상을 예측한다.

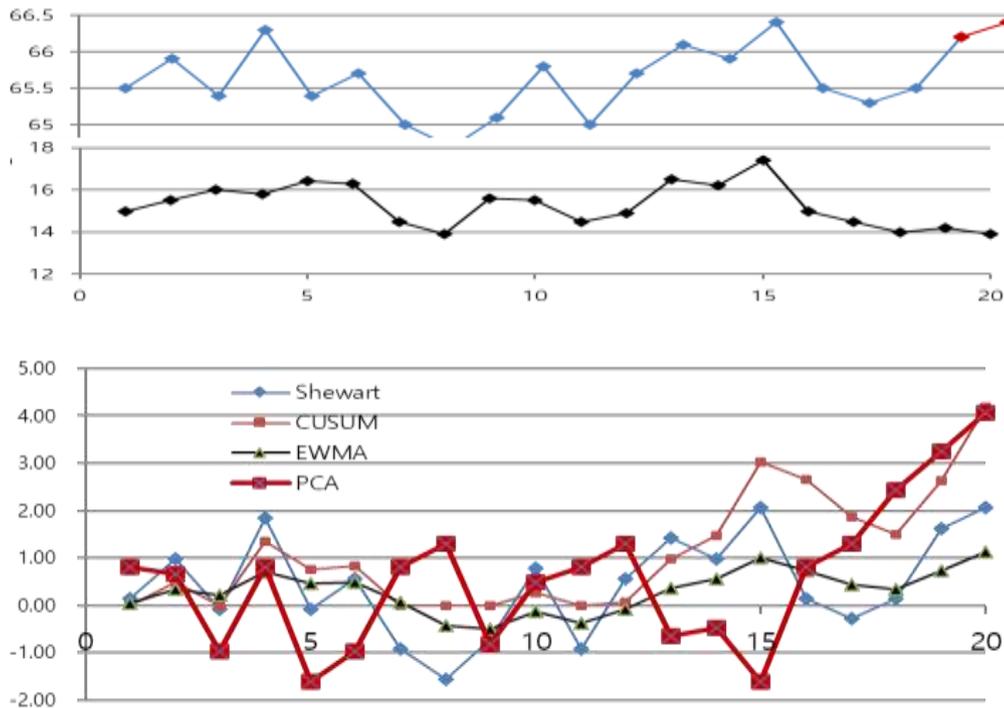


그림 12. PCA 변별력; (위) 두 변수의 PV 트렌드, (아래) 붉은색 PCA 트렌드

III. 맺음말

여러 종류의 고장예지 시스템이 도입되고 있으나 제대로 사용하지 못하고 있는 경우가 많다. 마치 유행처럼 고가의 외국 제품을 사전에 충분한 분석 없이 도입하는 경향도 있다. 제대로 된 시스템을 도입하기 위해서는 다음을 고려해야 한다.

첫째, 도입 목적을 명확히 해야 한다. 기존의 알람 체계를 개선하는 것인 목적인지 아니

면 회전기계와 같이 특수 설비의 고장예지가 목적인지 명확히 해야 한다.

둘째, 처음부터 고도화 기법을 바로 적용하는 것보다 단변량 기법에서부터 순차적으로 다변량 통계기법, 별도의 고도화 기법으로 확장하는 것이 좋다.

셋째, 알람 체계나 고장예지의 기법은 현장 지식을 얼마만큼 구현했느냐가 중요하다. 도메인 전문가의 참여가 절대적으로 중요하다.

알람 고도화 기법의 경보 체계는 일과성 과제가 아니다. 설비의 전주기에 걸쳐 지속적으로 보완되어 공장 조업에 필수적인 지적 자산으로 발전하는 시스템이다.