

다변량 데이터 압축법을 이용한 회분식 PVC 중합공정의 실시간 감시

이영학 · 한종훈[†] · 이종구*

공정산업의 지능자동화 연구센터/포항공과대학교 화학공학과

*(주)LG화학 연구소

(1998년 10월 16일 접수, 1999년 2월 19일 채택)

Real-Time Monitoring for a Batch PVC Polymerization Process Based on Multivariate Data Compression Methods

Young-Hak Lee, Chonghun Han[†] and Jong Ku Lee*

Automation Research Center, Department of Chemical Engineering,

Pohang University of Science and Technology

*LG Chemical Ltd., R & D Center

(Received 16 October 1998; accepted 19 February 1999)

요약

PVC(Polyvinyl Chloride)는 세계 4대 범용 수지 중의 하나로 생활용품이나 의료용품 등까지 다양한 용도로 널리 사용되고 있는 소재이다. PVC 공정은 발열반응에 의해 생성된 열을 제거하는 문제와 강한 비선형성이 존재하기 때문에 불안정한 조업이 자주 발생한다. 이러한 문제는 대부분 조업자의 경험에 의존하고 있다. 최근 들어 분산 제어 시스템(DCS)[◎]나 공정 정보 시스템(PIS)[◎] 도입되면서 데이터에 기반한 모니터링기술이 크게 발전해왔다. 수집된 방대한 양의 데이터는 MPCA(multiway principal component analysis)를 이용해 축소공간으로 투영함으로써 보다 간략한 데이터 표현이 가능하며, 기존의 SPC(statistical process control) 방법과 결합하여 효과적인 모니터링을 수행한다. 또한 이러한 방법은 감지된 이상에 대한 각 변수의 기여도(contribution)를 통해 수월한 진단을 가능하게 한다. 본 논문에서는 실제 산업체의 PVC 반응기의 이상 감시 모델을 구성하기 위해 MPCA 모델을 도입하였고, 통계적인 척도를 이용해 수집된 정상 회분 데이터로부터 통계적 제어 한계선을 설정하였다. 실시간으로 도입되는 데이터를 통해 감시 모델과 비교함으로써 현재 공정의 상태를 파악할 수 있게 된다. 또한 이러한 방법을 통해 감지된 이상에 대한 contribution 값들을 이용해 이상의 원인을 분석하는 방법을 제안하였고 산업체 PVC 중합반응기에 적용하여 공정 이상을 효과적으로 밝혀낼 수 있었다.

Abstract—PVC(Polyvinyl Chloride) has been widely used to the daily commodities and medical appliances as one of ordinary four resins. PVC processes have unstable problems due to heat removal mechanism from exothermic reaction and strong nonlinear behaviors. As a result, most of batch process operations are mostly dependent upon the experience of the operators. As distributed control system(DCS) and plant information system(PIS) have been installed in industry, monitoring skills have been rapidly developed. The enormous amount of data being collected are summarized using one of data compression methods, MPCA(multiway PCA), which is capable of compressing the information down into low dimensional spaces. Process performance can be monitored by using statistical process control(SPC) charts constructed from MPCA model. It can be used to analyze assignable causes using contribution of the variables on the abnormal events. To develop the process monitoring model for industrial PVC reactor, in this study, MPCA model is introduced, and its control limits are determined from the in-control batch data, which are compared with real-time process measurement data to know the status of process. Furthermore analytical diagnosis methodology based on the contribution of out-of-control state are proposed and tested on PVC reactor. The proposed analysis tools show the efficient monitoring diagnostic results.

Key words : PVC Polymerization Reactor, Multiway PCA, Statistical Process Control, Process Monitoring

1. 서 론

오늘날 치열한 경쟁시장에서 존속하기 위해 산업체에서는 수요자

*E-mail : chan@postech.ac.kr

들의 요구에 신속히 대응하고 제품의 신뢰성을 유지하면서 현 수준의 생산력을 계속해서 향상시키려는 노력이 이루어져 왔다. 이에 발맞춰 생산성을 저해하는 문제들을 조기에 감지하는 기술이 핫 이슈로 대두되어 왔다. 공정의 이상 현상들은 생산품의 질의 저하, 생산량의 감소, 복구작업의 증가, 심지어는 공장의 shutdown에 이르기까

지 크고 작은 경제적 타격을 입힌다. 이러한 공정 이상을 조기에 감지하고 진단하는 기술은 최근 들어 더욱 활발하게 발전했다.

공정의 성능을 감시하는 기술로는 공장의 실제 센서값과 수식모델의 예측값의 비교를 통한 방법, 정상 조업 조건에서 행해진 공정과 현재의 공정상태를 단변량의 통계공정제어(SPC) 차트들을 이용한 방법 등이 있으며, 이러한 방법들은 초기 감시기술로 이용되어 왔다. 그러나 수식모델은 정확한 세부 모델들이 요구되고 파라미터들이 불확실 하며 실시간으로 개선되어야 한다는 단점이 있으며, Shewhart 차트, Cusum Plot, EWMA 차트 등의 SPC 차트들을 이용한 방법은 서로 강하게 상호작용이 있는 변수들의 감시에는 그릇된 결과를 준다.

최근 화학공장에 분산제어 씨스템, 공정정보 씨스템 등이 지속적으로 도입됨에 따라 방대한 양의 조업데이터를 축적하고 활용할 수 있게 되었다. 이러한 데이터와 단변량 통계방법을 이용해 SPC 방법의 단점을 개선하고 공정의 상태나 품질을 감시하는 방법이 Kresta 등에 의해 90년대 초반에 발표됐다[6]. 이 방법의 장점은 강한 상호관계를 가지는 변수들을 다룰 수 있는 것은 물론, 측정오차의 제거, missing data의 처리, 그리고 특히 진단할 수 있는 능력을 가진다는 것 등이 있다[8].

이러한 발전은 회분식 공정으로 전파되었다. 회분식 공정은 다품종 소량생산 공정이고 고부가가치 제품생산에 적합하기 때문에 산업체에서 중요한 역할을 담당해 왔다. 회분식 공정은 다음의 단계로 운전이 진행되는데 특정 원료의 회분용기로 주입되고, 온도, 압력, 교반, 유량 등의 공정 변수가 시간에 따라 변하므로 이를 제어하면서 공정이 진행되며, 일정 전환율에서 용기로부터 제품을 뽑아내는 단계로 운전이 진행된다. 일반적으로 이런 과정 중에 공정이 혼들리는 요인은 평균추이를 벗어나는 변수들의 움직임, 원료 주입상의 에러, 불순물로 인해 생기는 외란 등으로 알려져 있다. 성공적인 회분조업이라 함은 일정한 원료의 혼합 및 일정 시간의 매질들의 주입과 변수들의 특정추이가 재현성을 갖도록 조업하는 것이다. 유량, 온도나 압력 profiles은 servo-controllers에 의해 구현되고, 시간에 따른 각각의 조업은 PLC(Programmable Logic Controller) 같은 도구를 이용해 실제 현장에서 구현된다.

회분식 공정의 감시는 안전 조업이나 고질의 제품생산을 보장하게 된다. 그러나 제품의 질을 측정하는 온라인 센서가 부족하고 비선형성이 존재하며 공정상에 존재하는 모든 화학반응, 혼합, 열전달 현상들을 수식화하기가 어렵다. 이런 어려움에도 불구하고 연구가 활발히 이루어졌는데 수식적 접근과 지식베이스기반 접근방법이 그것이다. 이 또한 연속 공정상에서 일어났던 문제들과 동일한 문제점들이 있었고 이에 대한 대안으로 발표된 것 또한 경험적 모델에 기반한 단변량 통계 투영 방법들이다. 회분식 공정의 단변량 통계 투영 방법은 데이터를 압축하여 저 차원공간을 형성하고 이에 데이터를 투영하여 정보를 도출하는 방법이며 이 공간은 성공적인 회분들의 공정 변수에 대한 시간들을 요약하여 가지고 있다. 이 방법을 이용하면 기존의 방법들보다 더 많은 진단정보를 제공하게 된다. 이런 방법을 이용한 온라인 감시는 에너지를 절약시키고, 원료물질 및 공장의 효율을 증가시키는 효과를 가져오며 실제 회분식 공정의 감시에 널리 이용되고 있다.

회분식 공정 감시를 위한 단변량 통계방법의 역사를 간단히 살펴보면, MacGregor와 Nomikos[4, 11, 12]는 측정된 공정 변수들을 이용한 회분식 공정 감시를 위해 Multiway Principal Component Analysis(MPCA)와 Multiway Projection to Latent Structures(MPLS)에 기반한 방법들을 제시하였다. 품질변수가 유용할 때 회분의 진행을 감시하면서 최종 품질을 예측하기 위해 MPLS를 이용할 수 있다. 공정변수와 품질변수, 초기 조건들을 포함한 회분 set-up 데이터를 포함하여 감시를 한 Multiblock Multiway Projection 방법도 발표됐다[4]. MPCA 모델은 또한 회분식 공정 데이터를 분석하여 프로세싱상의 주요 혼

들림의 원인을 확인하는데 이를 이용해 그 혼들림을 줄임으로써 일정하고 고품질인 제품을 생산하는데 일조하게 된다[5]. 회분식 공정 데이터는 일반적으로 각 회분마다의 외란과 물리적인 제약 조건들로 인해 일정한 전환율에 대해서 동일한 양의 데이터를 갖지 않는다. Kassidas 등[3]은 Dynamic Time Warping(DTW)을 이용해 측정값의 개수가 같지 않은 두 데이터를 동조화(synchronizing) 시킴으로써 회분식 공정의 분석, 감시 및 진단의 전처리 작업을 좀 더 수월하게 하는 방법을 발표했다.

이러한 통계적인 공정감시 방법은 각 측정값들에 따른 변수의 contribution을 이용하여 진단을 좀 더 수월하게 할 수 있는 장점을 가지고 있지만, contribution 값들에 대한 또 한 번의 해석을 요한다. 이러한 일련의 과정을 자동화하려는 연구가 기존의 진단방법과 결합하여 활발하게 일고 있는데 Venkatasubramanian과 Vedam[13]은 연속공정에서 PCA를 적용하고 발생한 이상의 contribution 값들의 해석을 signed digraphs를 이용한 이상진단기법과 결합하여 공정이상감시 및 진단을 자동화한 논문을 발표하였다.

본 논문에서는 실제 산업체의 PVC 반응기의 이상 감시 모델을 구성하기 위해 MPCA 모델을 도입하였고, 통계적인 척도를 이용해 수집된 정상 회분 데이터로부터 통계적 제어 한계선을 설정하였다. 실시간으로 도입되는 데이터는 이와 비교함으로써 현재 공정의 상태가 정상인지 아닌지를 파악할 수 있게 된다. 또한 이러한 scheme을 통해 감지된 이상의 contribution 값을 이용해 이상의 원인을 분석하는 방법을 제안하였다.

2. 대상공정 : PVC 종합공정

대상공정은 PVC(Polyvinyl Chloride) resin을 제조하는 회분식 종합 반응기이다. 일반적으로 PVC는 혼탁 중합(suspension polymerization) 공정이며 물을 매질로, VCM(Vinyl Chloride Monomer)과 유용성 개시제 용액을 혼탁안정제와 교반력에 의해 지름이 약 100 μm인 입자로 만들어 개개의 입자가 한 개의 괴상중합(VCM과 개시제만의 중합) 반응계로서 독립적으로 중합을 진행시키는 회분식 공정을 일컫는다. VCM 중합 반응은 자유 라디칼 중합에 의한 발열 반응이다. 물은 VCM을 분산시키는 매체일 뿐만 아니라 중합도중에 발생하는 중합열을 외부로 전달하는 역할을 한다. 전체 공정은 중합 반응기(Fig. 1)와 stripping column, 탈수 및 건조부, VCM 회수부로 나뉜다. 중합 반응기에서 반응이 종료되면 slurry(PVC resin 입자들이 물에 분산되어 있는 것)는 blow down tank를 거쳐 미반응 VCM을 회수하고 stripping column으로 이동한다. 수지에 미량 남아있는 VCM을 제거하고 탈수 및 건조시킨다. PVC 혼탁 중합 반응기의 조업 싸이클은 다음과 같다[1].

(1) 중합기에 일부의 물, 개시제 및 분산제 등의 부원료를 넣고 특정한 온도까지 가열한 후 물과 VCM을 함께 중합기로 투입한다.

(2) 원료를 넣은 후에 교반을 시작하고 중합기 jacket으로 스팀이나 특정온도의 냉각수를 넣어 가열 또는 냉각하여 일정한 온도를 유지하며 중합을 진행한다.

(3) 중합도중 중발되는 VCM을 환류 냉각기에서 냉각시켜 잠열(latent heat)을 제거한다. 반응 초기과 후반에 교반기 냉각수(sealing water)를 공급하고, 반응도중 감소하는 물을 보충하고 반응기 온도를 제거하기 위해 급냉수(quenching water)를 보급한다.

(4) 중합이 전환율에 의해 끝났음을 확인한 후 반응기로부터 slurry를 뽑아내고 cleaning 및 antifouling 조작을 행한다.

위 공정은 스팀, 냉각수, 급냉수의 공급량의 3개의 조작변수를 갖

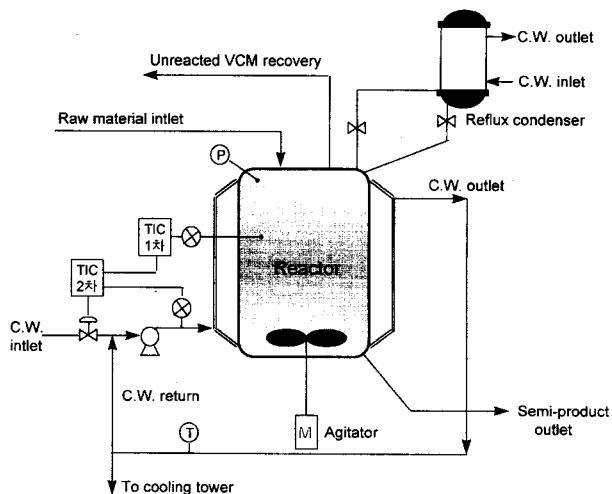


Fig. 1. PVC polymerization reactor.

는다. 스텝은 반응이 일어나기 전에 반응 온도를 최소의 시간에 원하는 수준까지 올려서 반응물질을 가열시켜주고, 냉각수는 중합열을 제거해 주며, 금냉수는 반응열을 제거시켜줌과 동시에 단량체(monomer)와 고분자와의 밀도차에 의한 부피감소를 보충해 줘서 반응물질의 부피를 일정하게 유지시켜 준다. 또한 스텝에 의한 가열작업 후에 반응 온도와 공정의 외란을 제거해 주기 위해 캐스케이드제어가 이용된다.

3. 0| 론

3-1. Multiway Principal Component Analysis 모델링

화학, 심리 측정학, 그리고 image analysis와 같은 분야들에서는 실험 연구용 데이터가 3차원 배열의 형태를 취하는 경우가 있다. 이는 회분식 반응기의 감시문제에서도 마찬가지이다. 회분식 공정들을 감시하기 위해 사용되는 데이터는 성질을 이해하기 위해 한 batch 동안 J개의 변수와 K개의 sampling time을 갖는 데이터가 측정된다. 할 때 이러한 유형의 데이터를 N회의 조업동안 수집하면 3차원 데이터 행렬인 $\underline{X}(N \times K \times J)$ 의 형태를 갖는다. 이 데이터 행렬은 통계적 혹은 기하학적 해석을 수월하게 하기 위해서 2차원화(unfolding)하는데, 이는 세 방향에서 두 가지씩 모두 6개의 2차원 행렬을 형성할 수 있다. 각 6개의 행렬에 PCA를 적용하면 다른 종류의 variability를 설명하게 된다. 이 중에서 본 연구의 목적에 적합한 unfolding을 찾으면 $BATCHES \times [VARIABLES * SAMPLE TIMES] \times [BATCHES * SAMPLE TIMES] \times VARIABLES$ 이다. Fig. 2는 두 가지 unfolding의 개략적인 형태이며, [A]는 회분들 사이의 유사성 분석이나 그룹화를 통한 공정개선, 통계적 공정 감시에 유용하며, [B]는 공정 변수간에 상관관계와 각 회분시간에 따른 조업이력에 관한 정보를 제공한다.

본 연구에서는 데이터의 비선형성을 선형화하고 변수와 시간간의 상관관계를 동시에 고려하기 위해 [A]의 unfolding을 채택하여 MPCA 모델링을 하였고, MPCA는 초기변수를 축소공간으로 투영하여 데이터를 압축하는 PCA처럼 펼쳐진 데이터의 변수들과 시간에 대한 변화를 요약함으로써 \underline{X} 에서의 batch들 사이의 variability를 분석하였다. 즉, 3차원 배열인 \underline{X} 를 Fig. 3에서처럼 score vector(t)와 loading matrix(P)들로 구성된 일련의 PC들에 residual matrix(E)가 더해진 형태로 분해시켜 준다. 이러한 분해를 수행함으로써 데이터는 Signal($\sum_{r=1}^R t_r \otimes P_r$) 부분과 Noise(E) 부분으로 나눈다. 이 과정은 식 (1)

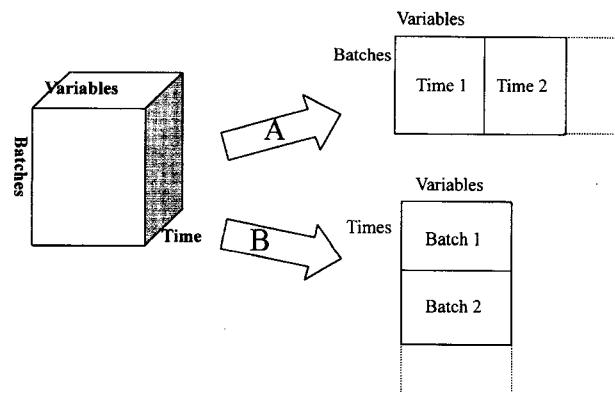


Fig. 2. Unfolding of process data matrix.

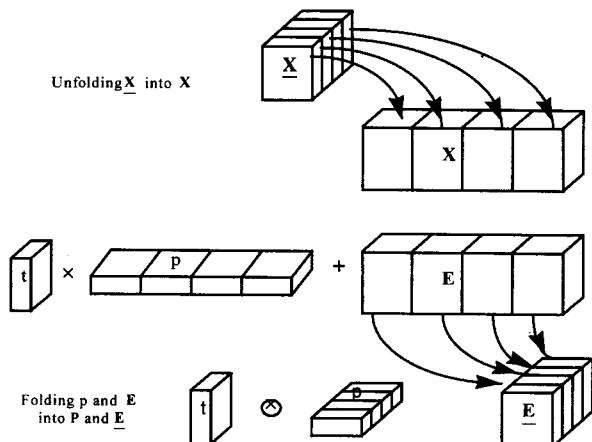


Fig. 3. MPCA modeling procedures.

로 표현하였고 Fig. 3에 나타내었다.

$$\underline{X} = \sum_{r=1}^R t_r \otimes P_r + \underline{E} \quad (1)$$

여기서 t_r 는 임의의 회분의 r번째 PC 값에 대한 t-score 값에 해당하고, loading 행렬(P_r)은 변수와 시간에 대한 정보를 가지고 있다. 따라서 t-score 벡터의 i번째 값은 전체 데이터의 i번째 회분과 관련이 있고, loading 행렬의 요소들은 그 회분의 t-scores를 주는 가중치이다. 그러나 Fig. 1에서의 [A] 경우는 PC의 파라미터들이 너무 많아 covariance structure내의 변수간만의 상관관계를 파악하기가 쉽지가 않다. 이런 어려움은 [B]를 이용한 모델을 통해 해결할 수 있다.

통계적인 모델이 구성되면 임의의 PC에 대해 어느 변수가 주요한 영향을 끼쳤는지, 어느 시간대에 어떤 PC가 지배적인지, 회분에 따른 PC의 영향은 어떤지를 각 PC의 Explained Percent Variance($EPV_{n,j}$)를 측정함으로써 알 수 있고, 이 정보는 공정 감시 및 진단을 위한 지식 베이스로 활용된다. 변수에 대한 EPV의 정의는 식 (2)와 같다. 임의의 j번수의 초기값 분산 합에 대한 모델값 분산의 합의 비로써 정의된다. 즉 j번째 변수의 전체 분산 중에 모델에 의해 설명되는 분산의 비이다. 이 수식은 각 회분과 시간에 대해서도 동일한 폴로 나타낸다.

$$EPV(j) = \frac{\sum_{n=1}^N \sum_{k=1}^K (\hat{X}(n, k; J=j) - \bar{X}(j))^2}{\sum_{n=1}^N \sum_{k=1}^K (X(n, k; J=j) - \bar{X}(j))^2} \quad (2)$$

3-2. 회분식 공정의 실시간 감시

모델 데이터로부터 MPCA 모델링 과정과 모델 해석 도구(EPV)에

대해 설명하였고 이제 임의의 한 회분에 대해 각 시간에 따른 실시간 감시 과정을 서술하고자 한다. 외부로부터의 한 회분 $[X_{new}(K \times J)]$ 에 대해 score 값과 residual은 다음 식을 이용해 구한다.

$$t_r = X_{new} \circ P_r \quad (3)$$

$$E = X_{new} - \sum_{r=1}^R t_r \otimes P_r \quad (4)$$

식 (3)과 (4)의 연산에서 score와 residual 값을 구하려면 X_{new} 는 완전한 데이터이어야 하지만 실시간으로 들어오는 데이터는 항상 완전할 수가 없다. 이를 해결하기 위해서는 미래값을 예측할 수 있는 적당한 예측 방법을 이용하여야 한다.

회분식 공정의 감시에 사용하는 미래값의 예측 방법에는 세 가지가 제안되어 왔다. 첫째, k시간 이후부터 조업 종료시까지의 미래값들이 모델로 구성한 데이터들의 평균값과 완벽하게 일치한다고 가정하여 mean-centering된 데이터에 대하여 아직 측정되지 않고 있는 미래값들을 0으로 대체하는 방법이 있다. 두 번째 공정에 발생하는 외란이 조업 내내 지속적으로 존재하게 되면 모델의 mean trajectory로부터 벗어나는 정도가 유사하다는 점에 착안하여 아직 측정되지 않은 미래값들에도 현재의 편차가 그대로 지속된다고 가정하고 예측하는 방법이다. 세 번째는 PCA가 missing value를 해결할 수 있다는 점을 이용하여 아직 측정되지 않은 공정 변수를 missing value로 취급하여 예측하는 방법이 있다. 위의 세 가지 방법들은 각각 장점과 단점을 가지고 있어 공정에 따라 적절한 방법을 선택하여야 한다. 회분식 공정의 제어는 회분 전 시간동안 정해진 rate의 tracking(set-rate tracking)을 요구하므로 servo-controller가 이용되는데, 일단 mean trajectory와 차이가 생기면 그 시점에서 servo-controller가 일정한 off-set을 가지고 작동하므로 일반적으로 두 번째의 방법이 좋은 예측결과를 준다. 본 연구의 simulation은 모두 두 번째 예측방법을 통해 이루어졌다.

새로이 시작되는 회분을 실시간으로 감시하는 과정을 정리하면, ① 우선 시간 k까지의 각 변수의 측정값들을 취하고, 모델 데이터로부터 k번째 시간까지에 대응하는 평균과 표준편차로 감시코자 하는 벡터를 centering과 scaling시킨다. ② k번째 시간 이후에는 위에서 제시한 세 가지 예측법 중에 하나를 택하여 회분이 끝나는 시간까지의 값들을 예측한다. ③ (3), (4)식을 이용하여 score 값과 SPE 값을 다음의 식들을 이용해 구하고 k값이 회분 끝 시간이 될 때까지 이 과정을 반복한다.

$$\begin{aligned} t_{r,k} &= X_{new,k} \circ P_r \\ E &= X_{new,k} - \sum_{r=1}^R t_{r,k} \otimes P_r \\ SPE_k &= \sum_{j=1}^J E(k,j)^2 \end{aligned} \quad (5)$$

모델은 유효하지만 도입되는 데이터의 변화가 커질 때 $t_{r,k}$ 는 제어 한계선 밖으로 이동하고, $t_{r,k}$ 는 제어 한계선내에 존재하는데도 불구하고 $E(k,j)$ 값의 변화가 커짐으로써 여러 회분에 대해서 계속하여 제어 한계선을 벗어날 때는 더 이상 모델이 유효하지 않으므로, 새로운 모델을 만들어야 한다. 일반적으로 이상의 정확한 이해는 각 PC의 t-score 차트를 이용해 가능하고, 공정 운전의 변화에 대한 정보는 SPE 차트와 T^2 (Hotelling's statistics) 제어 차트를 통해 제공되며, 이들은 서로 상호 보완적인 관계로써 공정 운전의 추이를 보여준다. 특히 $T^2(D)$ 는 PC공간에서 정상 조업의 중심과 X_{new} 의 t-score 값과의 거리의 척도이며 (6)식에 의해 구해진다.

$$D = t_k' S^{-1} t_k \frac{I(I-R)}{R(I^2-1)} \quad (6)$$

R은 MPCA 모델의 PC수, I는 모델 데이터 회분의 수, S는 모델

데이터의 t-scores의 추정된 공분산 행렬이다.

일단 SPE 차트와 D 제어 차트를 통해 이상이 실시간으로 감지되면 t-score 감시 차트를 통해 구체적인 이상의 원인을 파악할 개략적인 토대를 각 PC에 대해 지배적인 변수들이 존재한다는 개념을 바탕으로 수립한다. 이상이 감지된 시간대에서 이상 거동을 보인 변수의 상대적인 확률적 크기에 대한 정보는 변수의 contribution 값을 이용한다. 미래값 예측을 거친 X에 대해 임의의 k번째 시간에서 각 변수의 contribution은 식 (7)에 의해 구해진다. 식 (7)에서 $C_{i,k}$ 는 kj 번째 unfolding 행렬의 변수값들과 r번째 PC의 loading 벡터 요소들을 j번째 변수에 대해 곱한 후, 모든 시간대에 따라 더한 값이다.

$$C_{i,k} = \sum_{k=1}^K (X_{kj} \times |p_r|_{kj}) \quad (7)$$

결국 이런 정보는 현장의 운전원으로 하여금 좀더 수월한 진단을 수행하게 한다. 공정의 조업 cycle, 장치의 특성, 제어루프, 에너지와 물질 수지식 등의 지식베이스와 이러한 contribution 값을 통해서 추론된 정보를 바탕으로 진단이 수행된다.

3-3. 제어 한계선의 결정

새로운 batch의 감시는 과거에 성공적으로 조업한 회분들의 PC 공간에서 통계적인 분포와 새로운 batch의 조업 상황을 비교함으로써 이루어진다.

t-scores 감시를 위한 제어 한계선은 모델 데이터의 회분들로부터 임의의 한 시간에 대해 한 제어 한계값이 정해진다. k번째까지의 모델 데이터로부터 k번째 제어 한계값을 구하고자 할 때 일단 회분의 나머지 시간대의 변수값들을 예측한다. 예측하여 완전히 구성된 데이터로부터 t-scores를 구한다. 이 값들은 normal distribution을 갖으며 참분산을 모르므로 t-분포를 따른다. 유의 수준 α 에서 한계값은 다음식에 의해 계산된다. 이 식에서 $t_{n-1, \alpha/2}$ 는 n-1의 자유도와 $\alpha/2$ 의 유의 수준을 갖는 t 값이고, n은 측정값 수(여기서는 회분 수), s_{ref} 는 주어진 시간 k에서의 t-score 값들의 추정된 표준 편차이다.

$$\pm t_{n-1, \alpha/2} s_{ref} \left(1 + \frac{1}{n} \right)^{1/2} \quad (8)$$

식(6)의 Hotelling's Statistic(D)은 F(R, I-R)-분포를 갖으며 테스트 가설은 어떤 회분이 정상인지 비정상인지를 구분하는 것이다.

SPE 차트의 한계선은 residual 행렬의 k 번째 시간에 해당하는 부분 행렬로부터 구한다. $E(k, j)$ 는 multinormal distribution $N(0, \Sigma)$ 이며 SPE는 가중치가 부여된 chi-squared 분포(χ^2_m)이고, 가중치(g)와 자유도(h)는 Σ 의 고유값의 합수이다. 유의 수준 α 와 시간 k에서 SPE에 대한 제어한계값들은 식 (9)에 의해 구한다. 여기서 m과 v는 SPE의 추정 평균과 추정 분산이며 $\chi^2_{m, v, \alpha}$ 는 α 의 유의 수준과 $2m/v$ 의 자유도를 갖는 chi-squared variable의 critical value이다.

$$SPE_\alpha = (\nu/2m)\chi^2_{m, v, \alpha} \quad (9)$$

Contribution 차트의 제어 한계선은 모델 데이터의 회분들로부터 임의의 시간에 대해 각각의 제어 한계값이 정해진다. k번째까지의 모델 데이터로부터 k번째 제어 한계값을 구하고자 할 때 일단 회분의 나머지 시간대의 변수값들을 예측한다. 예측하여 완전히 구성된 데이터와 loading 계수들을 각 변수(unfolding 행렬의 변수)끼리 곱하고, 시간에 따른 각 공정변수를 합함으로써 contribution 값을 구한다. 이 값들은 normal distribution을 갖으며 참분산을 모르므로 t-분포를 따른다. 유의 수준에서 한계값은 다음식에서처럼 t-score의 제어한계선과 비슷한 식으로 계산된다. S_k 는 주어진 시간 k에서 변수

j 에 대한 contribution 값들의 추정된 표준 편차이다.

$$\pm t_{n-1, \alpha/2} S_{kj} \left(1 + \frac{1}{n} \right)^{1/2} \quad (10)$$

이 제어한계선은 회분의 감시가 진행되는 동안 각 시간에 따라 변수별로 설정되며 어느 시간, 어느 변수에서 상대적으로 이상에 큰 기여를 했는지에 대한 기준을 제공한다. 정상 제어 상태라 할지라도 다른 변수에 대해 상대적으로 큰 기여도를 갖는 변수가 존재하기 때문에 이상에 대한 변수의 기여도는 절대적인 기준인 이 제어한계선을 통해 확인한다.

3-4. 이상의 확인 및 진단

이상 감시 차트들의 통계적 의미를 볼 때 각 차트들은 0.05와 0.01의 유의 수준에 의해 이상과 정상의 판단을 하게 되는데 0.05 유의 수준이라 하면 100개의 도입된 측정값에 대해 5개 정도는 정상인데 이상이라고 판단할 가능성을 전제한 기준이다. 따라서 어떤 공정이 제어영역 밖으로 벗어났음을 확인하는 규칙으로는 일반적으로 다음과 같은 세 가지가 있다.

- ① 95 % 신뢰도 구간(i.e. 0.05 유의 수준; warning limits)을 벗어나는 값이 두 포인트 이상이 발생했을 때
- ② 99 % 신뢰도 구간(i.e. 0.01 유의 수준; action limits)을 벗어나는 값이 한 포인트 이상이 발생했을 때
- ③ 일곱 포인트가 연속으로 증가하거나 감소했을 때

위의 세 가지 기준에 의해 이상이라고 판단한 후 진단하려면 이상이라고 판단한 점(축소 공간 상의 score 값)을 원래의 공정 변수로 환원시켜 줘야 하는데 이때 사용하는 것이 바로 각 변수의 contribution이다. Contribution의 값을 통해 어떤 변수들이 제어한계선을 벗어나게 하는 주요 원인으로 작용했는지를 알 수 있고, 이상이 발생한 시점부터 이상이 전파된 경위를 공정의 비정상 상태 중의 하나로 간주하고 fault library에 저장한다. 물론 fault library에는 조업자의 지식이나 경험을 바탕으로 각 fault에 대한 적절한 조치에 대한 정보는 포함되어 있어야 한다.

4. 결과 및 고찰

4-1. 데이터의 수집 및 전처리

데이터는 현장에 구축된 PIS로부터 11개의 공정 변수에 대해 360

Table 1. Process measurements

No. Measurement	No. Measurement
v1 Reactor temp.	v6 Quenching water heat
v2 Jacket inlet temp.	v7 Jacket cooling water heat
v3 Jacket outlet temp.	v8 Agitator heat
v4 Reactor pressure	v9 Agitator sealing water heat
v5 Conversion	v10 Reflux condenser cooling water heat
	v11 Agitator ampere

개의 측정값(sampling time은 1.5분)들을 얻었다. 측정 변수들은 Table 1에 나타내었다. 이 측정값 중 batch 조업 초반(원료의 투입 및 가열기간)과 후반(blow down)의 데이터는 불규칙한 센서값의 분포로 인해 제거하고 중합 반응이 일어나는 시간동안 측정된 241개의 데이터를 사용하였다. 중합반응 동안의 측정값들은 모든 회분에 대해 일정한 반응의 정도를 나타내기 때문에 각 회분의 synchronizing은 필요가 없다[3]. Fig. 4는 이렇게 얻어진 21개 batch에 대한 반응기의 온도, 압력, agitator ampere를 나타냈다. 이 데이터는 모두 조업이 끝난 후 실험실에서 분석을 통하여 최종 제품의 물성을 만족시킨다. 18번 batch의 경우는 최종제품의 물성을 만족시킴에도 불구하고 온도와 압력의 경향이 다른 정상 batch들과는 달름을 알 수 있다. 이 회분을 outlier라 판단하여 모델 데이터에서 제거시킬 수도 있지만 이 회분 또한 조업자들의 조작이 반영된 것으로 모델에 포함시켰다.

4-2. 모델 구성 및 모델 데이터와 비정상 회분 데이터의 분석

위에서 설명한 회분식 반응기에 대해 공정감시 씨스템을 MPC 모델링 기법을 이용하여 구축했다. 모델링에 이용한 데이터는 21개의 정상 회분이며, 각 회분은 11개의 변수(열)와 241번의 측정값(행)으로 이루어진 데이터 행렬이다. 3차원 배열의 데이터는 비선형성을 제거하고 PCA 알고리즘을 사용하기 위해 각 시간에 따라 unfolding(Fig. 2의 [A])시켜서 각 행렬의 열들을 centering, scaling 한다. 이 행렬에 NIPALS(Nonlinear Iterative Partial Least Square) 알고리즘을 적용하여 3개의 PC(고유벡터값을 통한 cross validation)를 구했으며, 각 PC가 모델을 설명하는 정도(R^2X : 현재의 PC에 의해 설명된 X의 sum of squares의 비)와 고유값(Eigenvalue: 각 PC의 산포도에 대한 척도)은 Table 2에 정리하였다. 여기에서 4번째 이상의 PC들은 공정의 중요한 정보를 나타내기 보다는 공정 잡음 등 random variation을 나타내게 되어 감시 모델 구성에서 제외된다.

구성한 모델의 데이터들은 표준 회분들로 구성되어 있으므로 그 자체에 많은 정보가 포함되어 있으며, 이 정보들은 비정상 회분 조

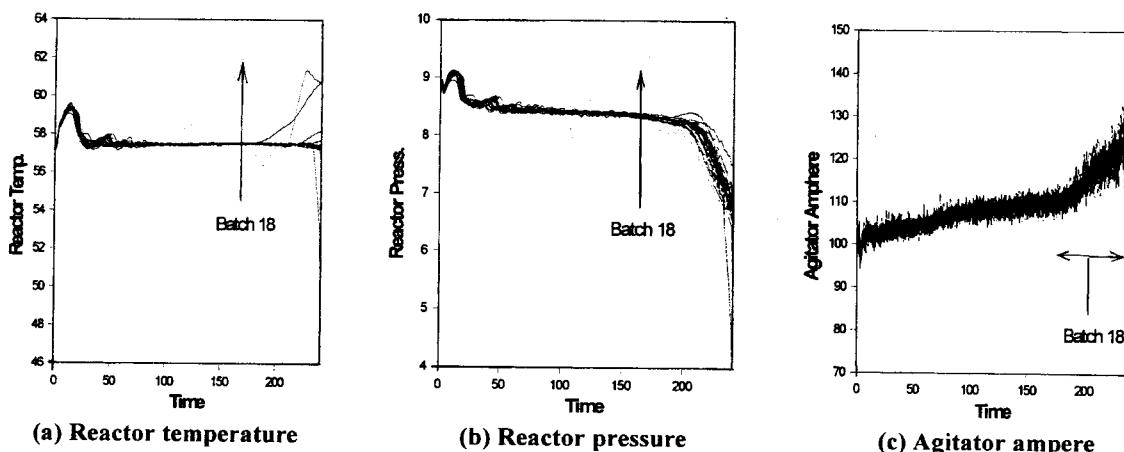


Fig. 4. The trajectories of principal process measurements.

Table 2. Model overview

No. of PC	$R^2 X$	Eigenvalue
1	0.365	7.669
2	0.581	4.537
3	0.710	2.707

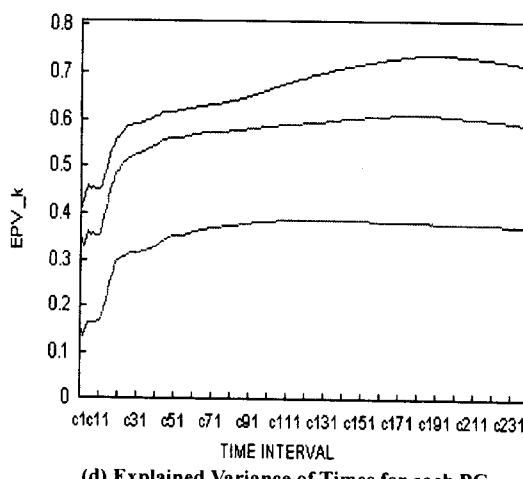
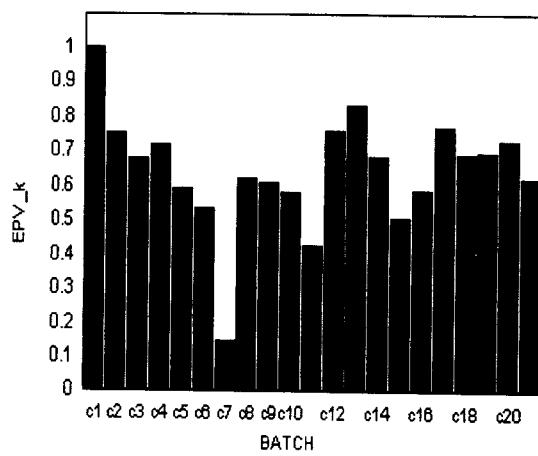
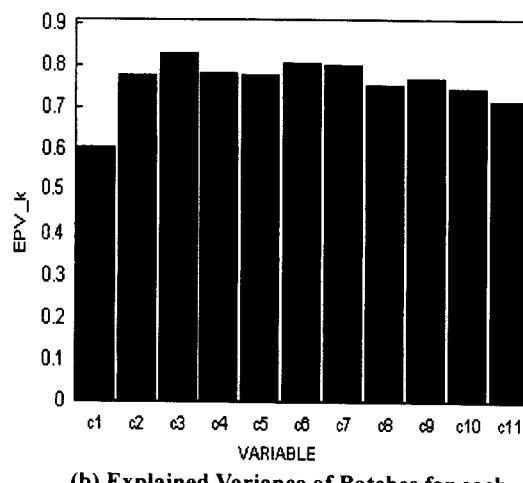
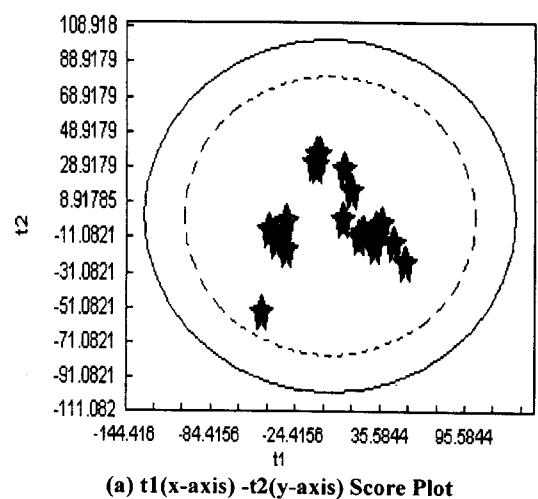
업 데이터로부터의 정보와 함께 이상 감시 및 진단에 유용한 정보로 활용된다. 여기에서 비정상 회분 조업 데이터란 미리 설정되어 있는 최종 제품의 품질 기준을 만족시키지 못하는 조업이 이루어졌을 때 수집된 데이터들을 말한다. 모델용 데이터의 Unfolding 행렬들(Fig. 2의 [A], [B])과 비정상 회분 데이터를 통해서 얻을 수 있는 정보는 각 t-score 감시 차트 중에 어느 차트가 이상을 감지할 확률이 가장 높은지와 어떤 변수가 불안정한지 등이다. 또한 회분의 각 시간에 따라 감시를 수행할 때, 시간에 따른 모델의 정확성을 파악함으로써 SPE를 통한 감시가 유용한 시간대를 알 수 있다.

먼저 모델에 비정상 회분을 모든 PC에 대해 투영해봄으로써 어느 PC에서 주로 이상이 발생할 것인지를 유추한다. 어느 PC가 이상이 감지될 확률이 제일 높은지를 확인한 후, 그 PC에 어떤 변수들이 가장 큰 영향을 미치는지 파악하여 미리 이상 감지에 대한 기본적인 골격을 갖춘다.

모델에 비정상 회분을 투영한 결과는 Fig. 6에 나타낸다. 처음 21개의 회분은 모델에 이용된 것이고 22-29의 회분은 조업이 끝났을

때 비정상의 품질을 보이는 회분들을 나타낸다. (a), (b)에서 보는 것처럼 PC1과 PC2에서는 이상 회분들의 감지력이 떨어지고 반면에 PC3-(c)에서 이상에 대한 감지확률이 제일 높음을 알 수 있다. 이를 토대로 Fig. 5(c)를 통해 어느 변수가 세 번째 PC를 가장 잘 설명하고 있는지를 알 수 있다. (c)에서는 반응기 온도, 급냉수의 열량 등이 주요 변수임을 알 수 있다. 또한 주요 변수들은 Fig. 2의 [B]를 이용한 PCA 모델링을 이용한 loading coefficients을 통해서도 알 수 있다. Fig. 7은 이를 나타낸다. 이와 관련하여 이상이 발생하고 있음을 확인함으로써, 실시간 감시 및 진단이 좀더 수월하게 이루어진다. Fig. 5는, 변수(b), 회분(c), 시간(d)들이 PC에 따라 어느 정도 변동량을 설명하고 있는지(EPV Plots)를 보여준다. (a)는 데이터 대부분의 분산을 설명하는 PC1과 PC2에 대해 회분들의 t-score 값을 plotting한 figure이며, 회분들 사이의 유사성이나 clustering의 정보를 제공한다.

각 시간대에 따라 모델이 설명하는 정도를 설명하는 Time's EPV Plot[Fig. 5-(d)]을 보면 모델이 다른 조업 시간대와 비교했을 때 조업 초기의 변화들이 거의 설명하지 못하고 있음을 알 수 있다. 이 시간 영역은 PVC 반응기 내부의 가열이 끝난 직후 회분 조업이 시작되는 초기로서 일반적으로 전체 회분조업을 결정짓는 중요한 부분이라 할 수 있다. 따라서 이 시간 영역에서의 변화들이 제대로 설명되지 못하고 있다는 것은 조업 초기부에 MPCA 모델의 covariance structure를 벗어나는 조업 조건이 자주 발생한다는 것을 의미하며 결과적으로 공정 이상의 확률이 커질 것이라는 공정지식과 일치한다고 할

**Fig. 5. Model analysis charts.**

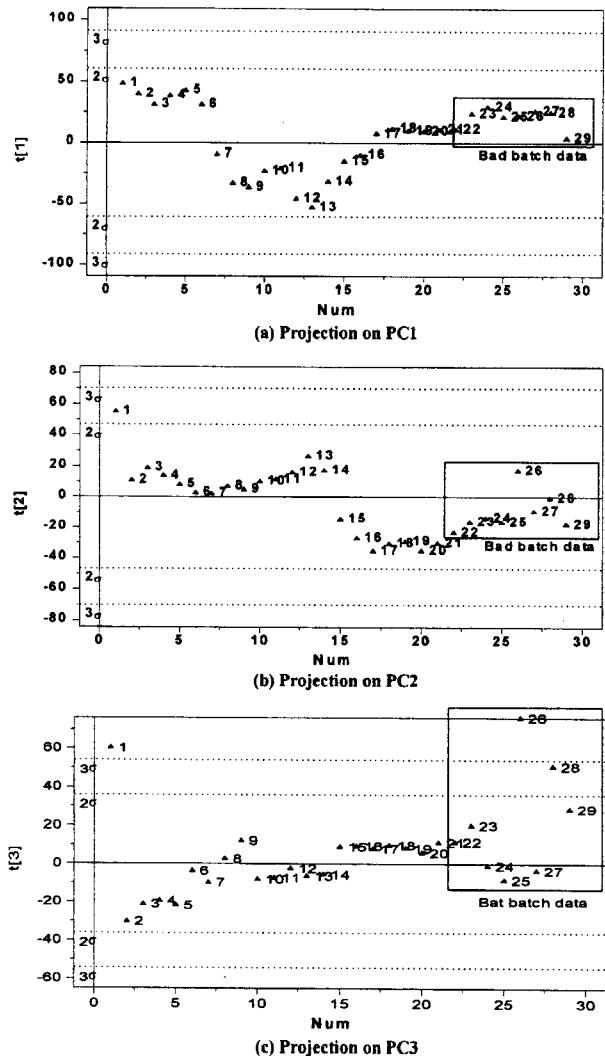


Fig. 6. Projection results on the model of eight bad batches.

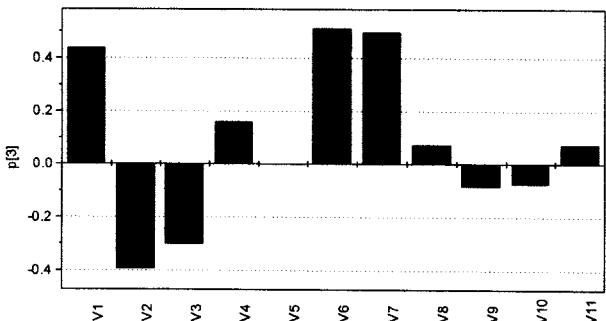


Fig. 7. Loading coefficients of 3rd principal component for the data matrix unfolded into the time direction.

수 있다. 이러한 점에서 조업 초반부의 감시는 t-score나 D 제어 차트보다는 SPE 차트를 통하여 더 명확하게 나타날 것이다.

4-3. 공정 감시

Fig. 6에서 보인 것처럼 PC3의 경우는 95 % 신뢰도 구간을 넘는 variance의 회분들이 나타난다. 감시차트 중에서도 세 번째 PC의 score 값의 변동에 더욱 더 주의를 요한다. 모델을 통한 정보 추출이

끝나면 다변량 SPC 감시 차트들을 이용해 실제 감시가 이루어진다.

k시간까지 얻어진 데이터를 바탕으로 앞으로 이 회분이 정상적인 조업으로 끝날 수 있는지를 판단하는 작업은 다음의 두 단계로 이루어진다. 첫째, 적절한 미래값 예측방법을 이용해 한 회분이 끝나는 시간까지 각 센서값들을 예측한다. 둘째, 예측된 데이터를 이미 구성된 모델에 투영하여 그 모델에 의해 구해진 한계값들과 비교해 현재 조업 중 어떤 시간에 이상이 발생했는지를, 또한 이 회분이 끝까지 정상적인 상태를 유지할 것인지를 판단한다. 감시 정보는 축소 공간에서 얻은 두 가지 정보를 통해 파악하게 된다. 첫째는 모델에 기반한 이상 감시를 할 수 있는 t-score 값, 각 t-score들의 효과를 모두 포함한 Hotelling's Statistic(D)이다. 두 번째로는 모델의 신빙성 (Goodness-Of-Fit) 및 모델에서 설명하지 못하는 정보에 대한 감시효과를 주는 SPE(Squared Prediction Error) 값이다. 즉, 특정 시간에서 모델의 이상 유무는 SPE 값-모델이 설명하지 못하는 노이즈나 모델에서 예상하지 못한 이상들을 감지할 수 있는 모니터링 차트로써 t-score 값으로 잡지 못하는 이상들도 이 값의 추이를 통해서 감지할 수 있다. 그래서, 이 회분을 재운전해야 할 지는 t-score 값으로 결정하게 된다. 이상의 감지는 세 가지 경우가 있을 수 있는데 t-score 값이 제어한계선을 넘었을 때 SPE 값 또한 넘었을 때 그리고 넘지 않았을 때의 두 가지와 t-score 값이 제어한계선을 넘지 않았는데 SPE 값만 제어한계선을 넘었을 때이다. 두 값이 다 넘었을 때는 확실히 이상이라고 판단할 수 있지만, t-score 값만 넘었을 때는 모델이 설명해 주는 variation 내에서 이상이 발생한 것이고, 즉 구성한 모델이 유효함을 나타내는 것이고 만일 SPE 값만 넘겼다면 이는 모델이 설명해 주지 못하는 조업변화가 있거나 noise part에서 이상이 발생했음을 의미한다. 이런 경우에는 모델 재구성을 요한다.

이제 실제 모니터링 차트들을 보고서 이상을 어떻게 감지하고 진단하는지 알아보기 위해 두 개의 비정상 회분에 대한 차트들을 Fig. 8과 11에 나타내었다.

Fig. 8에서 확인된 out-of-control의 t-score 값들에 대한 각 변수의 기여도는 Fig. 9에 임의의 시간에서 각 변수에 따라 나타나었으며 정상 회분들을 통해 미리 구해진 contribution 제어한계선을 이용하여 이상 원인 제공을 한 변수를 시간에 따라 선정할 수 있다.

Fig. 8에서 이상을 확인하는 과정을 살펴보면 2-4절의 규칙 중 첫 번째 이상(37th time) 확인 포인트는 세 번째 규칙인 일곱 포인트가 연속으로 감소하였을 때이다. 그림에서 보이는 것처럼 31번째 측정값에서부터 계속해서 감소하고 있음을 확인하였다. 일단 이러한 이상이 감지되면 기준으로 정해 놓은 한계선을 넘어갈 가능성이 많은 것이다. 결국 64번째 측정값에서부터 두 포인트가 연속으로 95 % 한계선을 벗어남으로써(두 번째 규칙) 한 번 더 이상을 확인할 수 있었다. 67번째 포인트를 지나면서 99 % 한계선을 벗어남으로써 어떠한 action을 취해주어야 함을 알 수 있다.

Fig. 8과 9를 이용해서 적절한 진단을 내리는 과정은 다음과 같다.

(1) 각 t-score 값들 중 이상에 민감하게 반응하는 것을 선택한다. Fig. 6에서 예측했던 것처럼 세 번째 t-score 차트가 가장 먼저 제어한계선을 넘어섰다. 다음, 두 번째 score 차트가 벗어났음을 보고서 일단 세 번째 PC의 전파에 의해 두 번째 PC가 영향을 받았음을 알 수 있다. 그러나 바로 두 번째 차트에서 감지된 이상들은 두 번째 PC와 밀접한 변수들이 적절하게 잘 조작되어 바로 제어한계선 안으로 들어가지만, 세 번째 PC 감시차트에서는 계속해서 이상조업이 감지되고 있다.

(2) 결국 Fig. 8의 감시 차트들 중 t3-score와 그것의 contribution을 나타내는 Fig. 9를 비교해 봄으로써 이상을 일으킨 근본적인 변수를 선정할 수 있고, 시간에 따른 contribution을 검토해 봄으로써 비정상

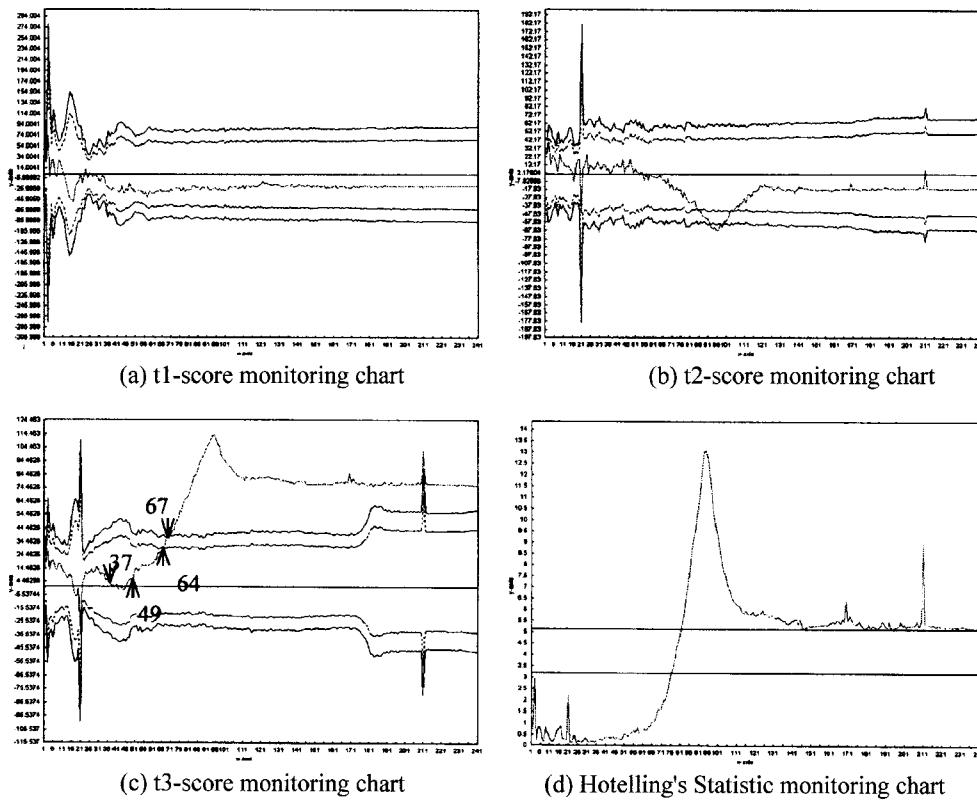


Fig. 8. The monitoring charts for abnormal batch(I).

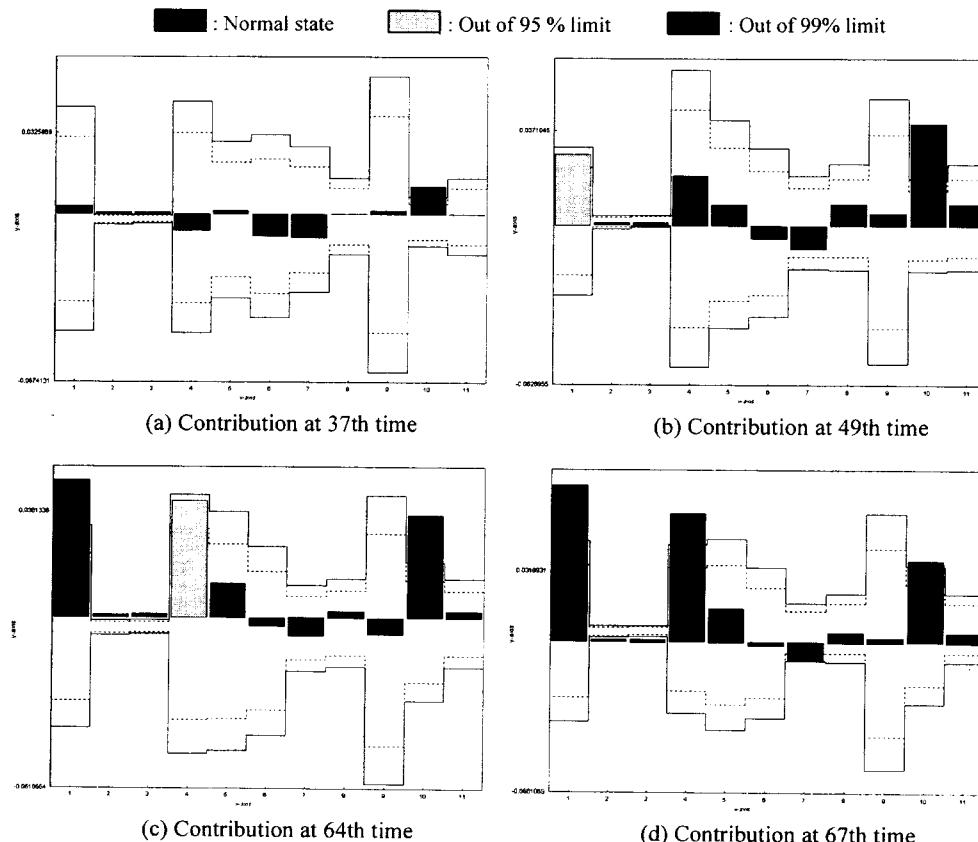


Fig. 9. Contribution charts for abnormal batch(I).

회분의 trend까지 파악할 수 있다. Fig. 9에서 옅은 회색(중간 띠)은 유의 수준 0.05를 절은 회색(마지막 띠)은 유의 수준 0.01을 넘었을 때 나타나는 명암이다. 이상의 trend는 정성적인 정보(명암의 변화)로써 nominal 변수로 저장하고 하나의 fault library로 정해진다. Fig. 9에서 보면 37번째 시간에서 환류 응축기의 냉각수 열량이 이상, 즉 큰 variation을 나타냄으로 인해 반응기 온도와 압력에 영향을 미쳤으며(b, c) 이는 공정 전체에 영향을 끼치게 된다. 즉, 반응기의 열을 제거 시켜 주는 장치는 가장 중요한 것이 jacket이고 급냉수와 교반기 sealing water, 응축기가 제열을 도와준다. 응축기의 load가 갑자

기 줄어들었다면 jacket의 냉각수 유량과 급냉수의 유량 제어를 위해 제어기가 활발한 작동을 하게 되고 stability 영역을 벗어나면 공정은 굉장히 불안정하게 된다.

(3) 근본적인 변수와 관계되어 있는 장치들이 잘 발생하는 failure를 점검함으로써 이상에 대한 corrective action을 취한다. 응축기의 경우는 fouling을 피하는 것이 급선무라 하겠다. 왜냐하면 응축기는 cleaning이 쉽지 않을 뿐더러 반응기의 효율을 높이는데 큰 역할을 하므로 cleaning 동안 응축기 사용이 자유롭지 않으면 큰 경제적 손실을 가져올 수 있다. 이러한 진단은 fault library가 완전히 구축되면

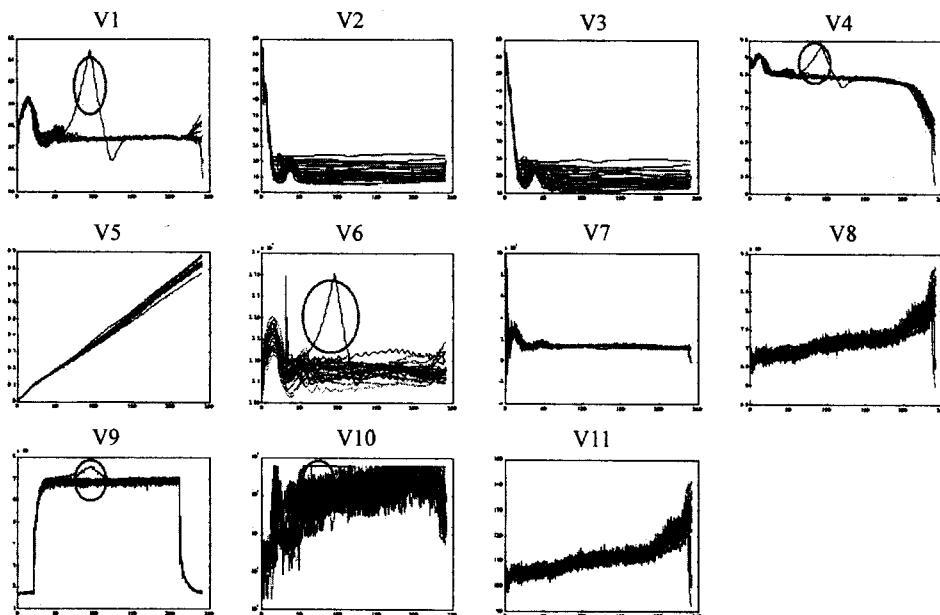


Fig. 10. The trajectories of eleven variables as shown in Table 1 for abnormal batch(I).

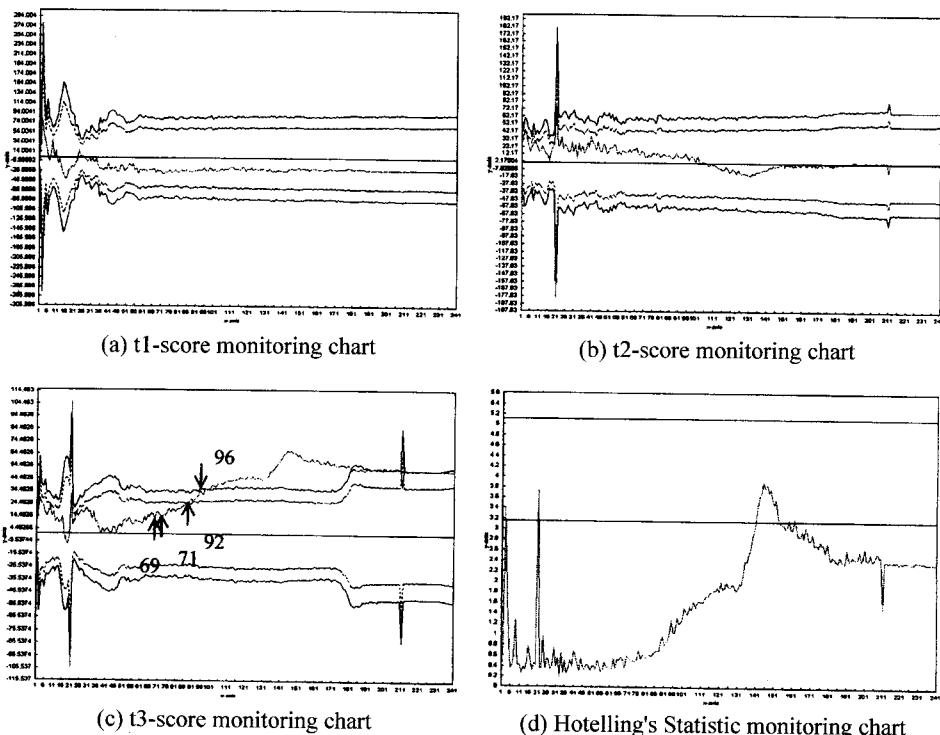


Fig. 11. The monitoring charts for abnormal batch(II).

운전원의 공정지식베이스와 결합하여 전문가 씨스템화할 수 있고 적절한 지침방향까지 정리되면 완전 자동화할 수 있게 된다.

이러한 진단의 검증을 위해 각 변수의 추이를 Fig. 10에 나타내었다. 반응기 온도와 압력, 급냉수의 열량, 교반기의 sealing water heat 그리고 환류 응축기의 냉각수 열량이 큰 variation을 나타냈지만 이들은 위에서 살펴보았듯이 환류 응축기의 이상으로 인해 전파되었을 가능성이 큼을 알 수 있다.

Fig. 11의 비정상회분(II)는 앞의 예에 비해 이상의 발생시간이 늦

다. 69번째 시간에서 응축기 냉각열의 variation에 문제가 있음이 처음 감지되었고 72번째와 91번째 시간에서는 반응기 온도와 압력에도 영향을 미치게 되었고 급기야 96번째 시간에서의 기여도를 보면 교반기까지 이상 현상의 영향을 받게 되었다. (II)의 경우도 결국 PVC 회분식 반응기의 문제점은 응축기에 보내주는 냉각수의 열량 제어나 응축기 내부에 fouling 때문일 것이라는 결론에 도달하게 된다.

마지막으로 지금까지의 분석에 대한 신뢰성 문제를 생각해 보지 않을 수 없다. 이는 SPE 차트를 통해 확인된다. 이상의 진단을 위해

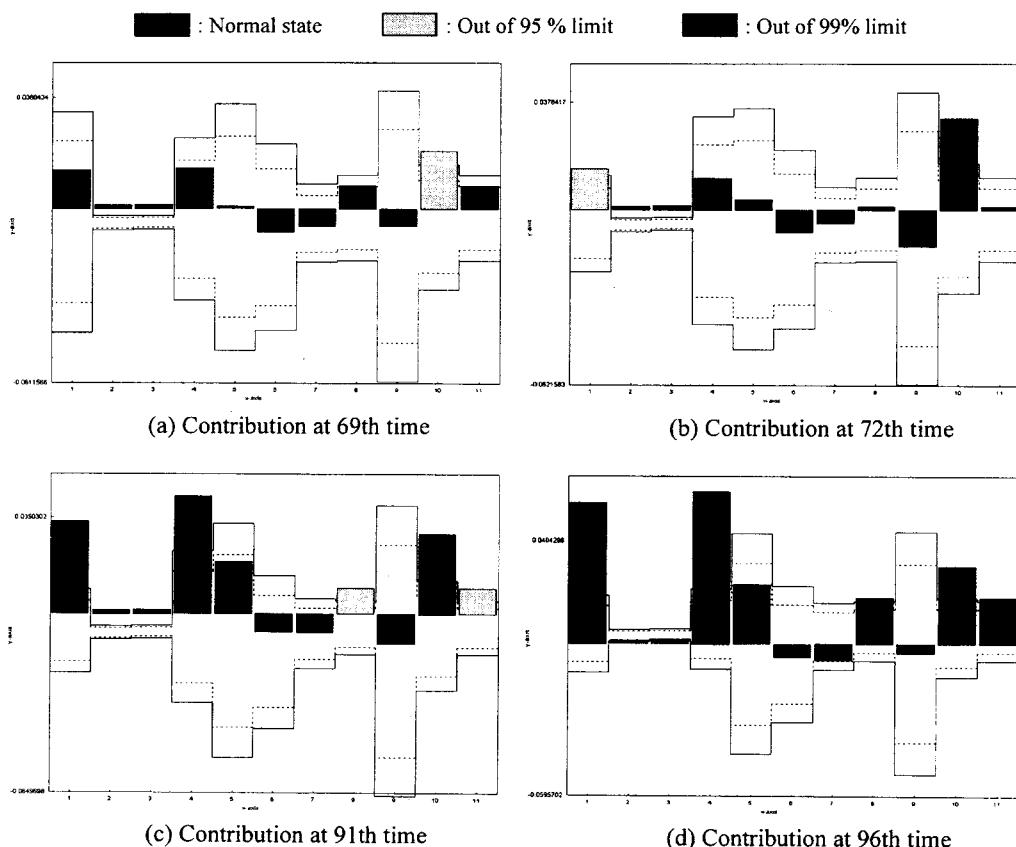


Fig. 12. Contribution charts for abnormal batch(II).

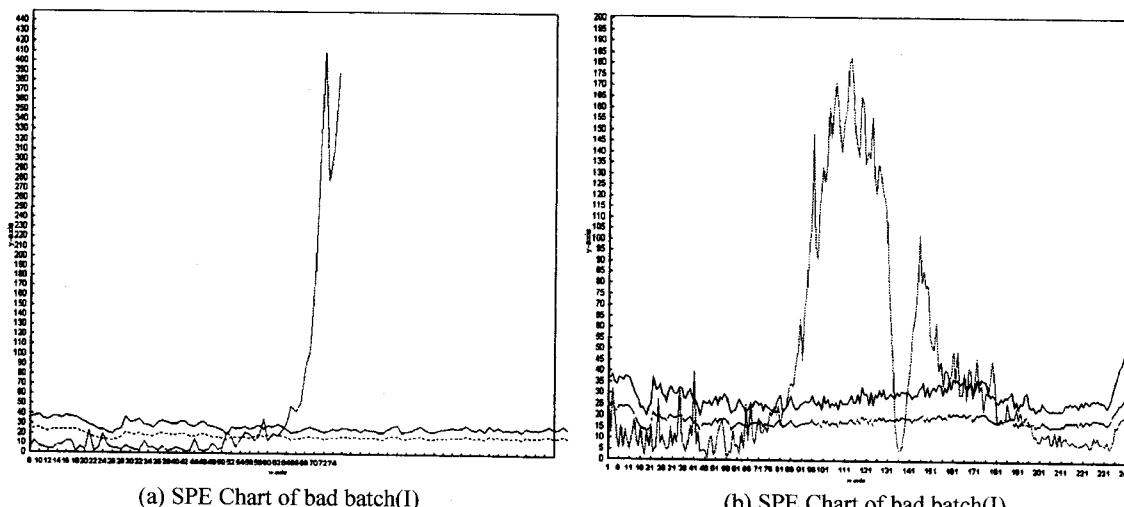


Fig. 13. SPE monitoring charts for abnormal batches.

구한 contribution 차트의 시간내에서 SPE 값들은 우리가 정해 놓은 신뢰도 구간내에 들어가야 한다. 만약 그렇지 않다면 모델 구성을 다시 하던지 아니면 그만큼의 위험성을 안고 진단을 하게 되는 것이다. 위의 분석에서는 Fig. 13에서 비정상 회분(I)과 (II)의 SPE 값은 이상이 처음 감지되는 시점이 한계선 안에 존재하므로 신뢰할 만한 분석임을 알 수 있다.

5. 결 론

본 연구에서는 다변량 통계 기법 중 데이터를 압축하여 축소공간에서 모니터링을 수행하기 위해 PCA를 도입하였다. 이를 이용하여 이상이 몇 번째 PC에서 일어날 가능성이 제일 큰지를 확인하고 이를 기반으로 이상을 감시하고 감지된 이상의 원인이 되는 변수들을 contribution 값을 이용해 찾아내는 방법을 제안하였으며, 사례 연구를 통하여 공정의 이상이 정확하게 감지됨을 조업 상황과 비교해 볼으로써 확인하였고, 이러한 방법을 통해 얻은 t-score 값에 대한 각 변수의 기여도를 회분식 공정의 진단에 효과적으로 활용하였다. 각 이상이 감지되는 회분의 기여도나 t-score 값의 signal은 "fault library"에 저장함으로써 공정에서 자주 발생하는 이상에 대한 corrective action을 위한 정보를 실시간으로 제공함으로써 이상감시 및 진단의 완전자동화를 이룰 수 있다.

감 사

본 연구는 과기부 1997년도 엔지니어링 핵심 공통 기반기술 개발 과제(97-EG-22-01-A)의 지원에 의해 수행되었으며 연구비를 지원하여 준 현대정보기술(주)에 감사드립니다. 또한 포항공대 공정산업의 지능자동화 연구센터를 통해 재정적 지원을 해주신 한국과학재단에 사의를 표합니다.

사용기호

D	: Hotelling's statistic
E	: residual matrix of a new batch
<u>E</u>	: residual matrix for historical database
F _{R,I,R}	: F-distribution with R and I_R degrees of freedom
g	: constant associated with the distribution of the squared prediction error
h	: degree of freedom for chi-squared distribution
i	: index for batches
I	: total number of batches
j	: index for measurement variables
J	: total number of variable measurements

k	: index for time intervals
K	: total number of time intervals
p	: loading vector
P	: loading matrix
r	: index for principal components
R	: total number of principal components
S	: covariance matrix of t scores
SPE	: squared prediction error
t	: score for a new batch
t	: score vector
X _{new}	: new batch
<u>X</u>	: historical database

그리아스 문자

α	: level of significance
χ^2_h	: chi-squared distribution with h degrees of freedom

참고문헌

1. Burgess, R. H.: "Manufacture and Processing of PVC", Applied Science Publishers(1980).
2. Johnson, R. A. and Wichern, D. W.: "Applied Multivariate Statistical Analysis", 3rd ed., Prentice Hall(1992).
3. Kassidas, A., MacGregor, J. F. and Taylor, P. A.: *AICHE J.*, **44**, 864 (1998).
4. Kourtzi, T., Nomikos, P. and MacGregor, J. F.: *J. Proc. Cont.*, **5**, 277(1995).
5. Kosanovich, K. A., Dahl, K. S. and Pivoso, M. J.: *Ind. Eng. Chem. Res.*, **35**, 138(1996).
6. Kresta, A. V., MacGregor, J. F. and Marlin, T. E.: *Can. J. Chem. Eng.*, **69**, 35(1991).
7. Lopes, J. A. and Menezes, J. C.: Proc. FOCAPO, Snowbird, Utah (1998).
8. MacGregor, J. F., Jackle, C., Kiparissidas, C. and Koutoudi, M.: *AICHE J.*, **40**, 826(1994).
9. Marlin, E. B., Morris, A. J. and Zhang, J.: *IEE Proc-Control Theory Anal.*, **143**, 132(1996).
10. Nomikos, P. and MacGregor, J. F.: *Technometrics*, **37**, 41(1995).
11. Nomikos, P. and MacGregor, J. F.: *AICHE J.*, **40**, 1361(1994).
12. Nomikos, P. and MacGregor, J. F.: *Chem. and. Int. Lab. Sys.*, **30**, 97(1995).
13. Vedam, H. and Venkatasubramanian, V.: IFAC Conf., Session 4A (1998).