

Results and Discussions

제안된 방법은 반회분식 페니실린 생산공정의 시뮬레이션 벤치마크의 공정 모니터링에 응용되었다. 표1은 회분식 발효공정으로 이루어진 페니실린 공정에서 공정모니터링에 사용된 11개의 공정변수들을 나타내며 그림 2는 공정모니터링에 사용된 11개 측정 변수의 전형적인 trajectory profile을 나타낸다. 전체 55 배치가 MPCA/MICA 정상상태 모델 구축을 위해 사용되었다. 각 배치의 시간은 400시간이며 그 중 45시간은 pre-culture stage이고 나머지 355시간은 fed-batch stage이다. Reference data set은 three-way $\underline{\mathbf{X}}(I \times J \times K)$ 로서 I 는 55 batches, J 는 11 process variables 그리고 K 는 400 time points이다. NOC models 위해서, reference batch data $\underline{\mathbf{X}}$ 는 Nomikos and MacGregor's unfolding method을 사용하여 a matrix $\mathbf{X}(I \times JK)$ 로 펼쳐진다. 그 후에 \mathbf{X} 는 발효공정의 average trajectory를 제거하기 위해 mean-centered and scaled 된다.

Off-line Analysis of MPCA and MICA

MPCA모델은 crossvalidation방법을 사용하여 4개의 주성분을 선택했으며 전체 변화의 63%를 설명한다. MPCA의 온라인 모니터링을 위해 Macgregor가 제시한 PCA를 이용하여 missing data를 다루는 두번째 filling method를 사용하였다. MPCA와 비교를 위한 MICA모델은 4개의 독립성분을 선택했으며 MPCA/MICA모델 둘 다 99% control limit를 사용하여 테스트 하였다.

그림 3과 4는 55 reference batches에 대해 MPCA/MICA를 사용한 an end-of-batch analysis결과를 설명한다. 그림 3에서 MPCA의 T^2 chart 에서 모든 배치는 the 95% confidence limit아래 있어 모든 배치가 NOC모델에 기반함을 알 수

있다. 그러나, 그림 4의 MICA의 I^2 chart는 off-line analysis batches에서 4 와 10은 outliers로 판정되어서 이 배치들은 outside of the range of the model임을 알 수 있다. 그림 5와 6은 첫번째와 두번째 latent variable space에서 MPCA and MICA의 score plot을 나타낸다 (PC_1 vs. PC_2 and IC_1 vs. IC_2). The scatter plot of MPCA은 모든 55 batches가 normal batch operation하에 있음을 나타낸다. 한편 MICA scatter plot은 MPCA결과와 다른 형태를 나타내고 I^2 chart와 마찬가지로, batches 4 and 10은 outliers로서 분류된다. 이 차이는 MPCA는 feature extraction방법이 Gaussian distribution assumption에 기초하고 Schwartz chart에 기초하므로 작은 크기의 외란를 탐지하기는 어려운 특징을 지닌다. 반면에 MICA방법은 non-Gaussian distribution이라는 feature extraction방법에 기초하기 때문에 공정의 중요한 영향 인자를 더 잘 뽑아낸 모델에 기인한다. 따라서 만일 data set이 non-Gaussian component라도 포함하고 있으면 ICA방법은 PCA보다 더 좋은 feature extraction performance를 보인다. 따라서 공정에 더 영향을 끼치는 key component을 더 잘 뽑아내는 ICA가 공정데이터가 non-Gaussian를 포함할 경우 더 좋은 모니터링 결과를 보인다.

그림 7은 MPCA의 first score (t_1)가 a Gaussian distribution에 따르지 않고 대신 ‘*supergaussian distribution*’을 따름을 나타낸다. 즉 dominant component가 Gaussian보다는 0에 매우 가깝거나 매우커다란 값을 가짐을 알 수 있다. 따라서 Gaussian distribution가정에 기초한 T^2 and SPE charts of MPCA는 잘못된 모니터링 결과를 제공할 수 있다. 실제 많은 산업공정에서 데이터들은 ‘*supergaussian distribution*’형태로 알려져 왔다. 이러한 결과는 MPCA모니터링

방법이 false alarm을 발생할 수 있는 근본적인 단점을 가지고 있음을 나타낸다.

⌘ 1. Variables used in the monitoring of the benchmark model

No.	Variables
1	Aeration rate(L/h)
2	Agitator power(W)
3	Substrate feed rate(L/h)
4	Substrate Feed Temperature (K)
5	Dissolved Oxygen Concentration (g/L)
6	Culture Volume (L)
7	Carbon Dioxide Concentration (g/L)
8	pH
9	Fermentor Temperature (K)
10	Generated Heat (kcal)
11	Cooling Water Flow Rate (L/h)

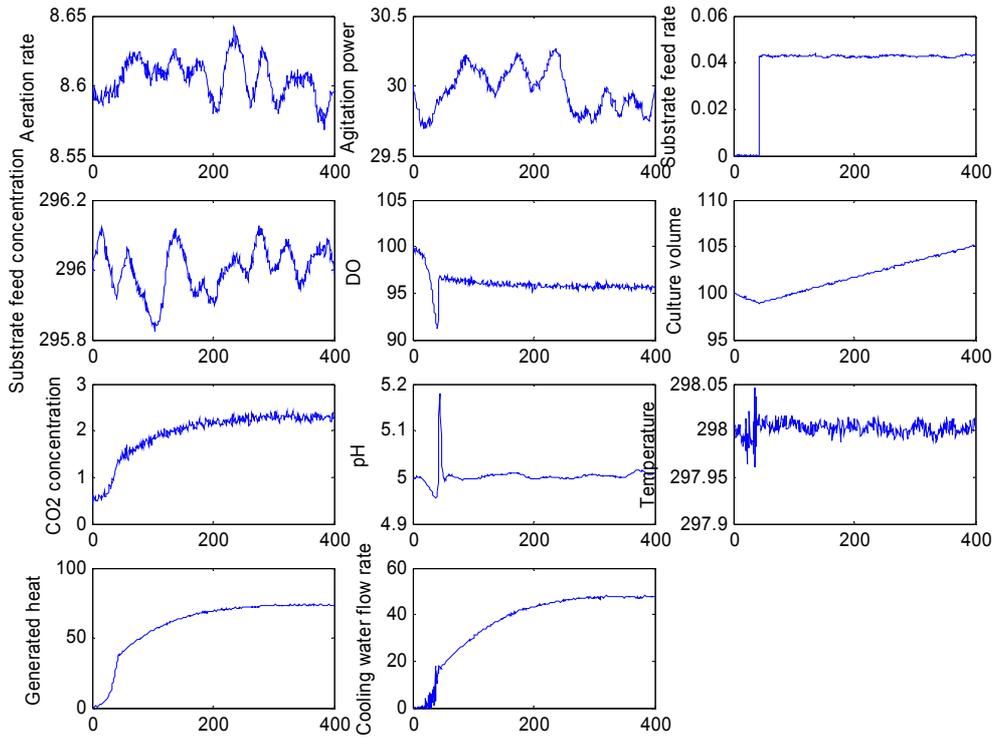


그림 2. Typical batch profiles of 11 variables during penicillin fermentation.

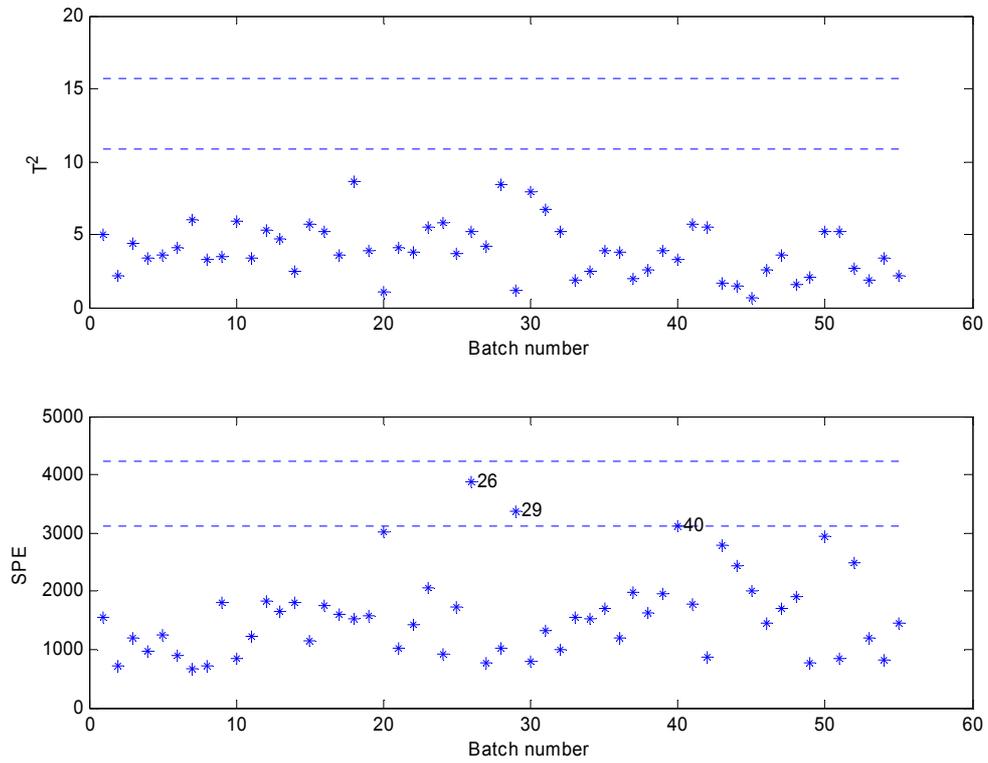


그림 3. T^2 and SPE charts of an end-of-batch analysis based on MPCA for 55 reference batches. The dotted lines correspond to 95 and 99% confidence limits.

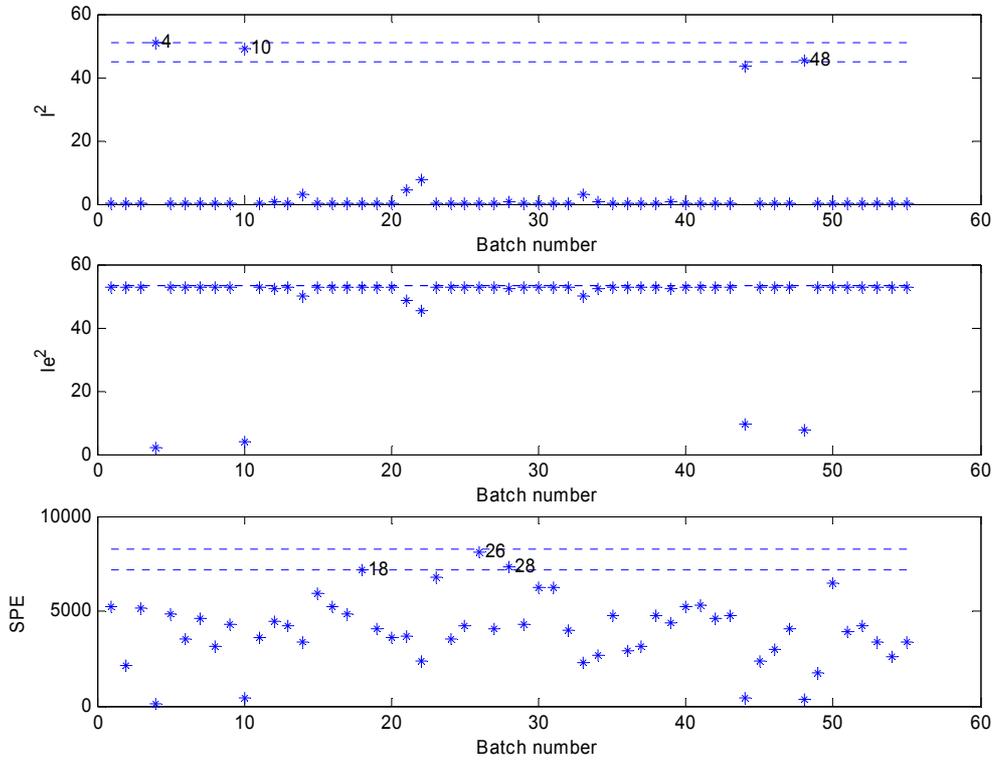


그림 4. I^2 , I_e^2 , and SPE charts of an end-of-batch analysis based on MPCA for 55 reference batches. The dotted lines correspond to 95 and 99% confidence limits.

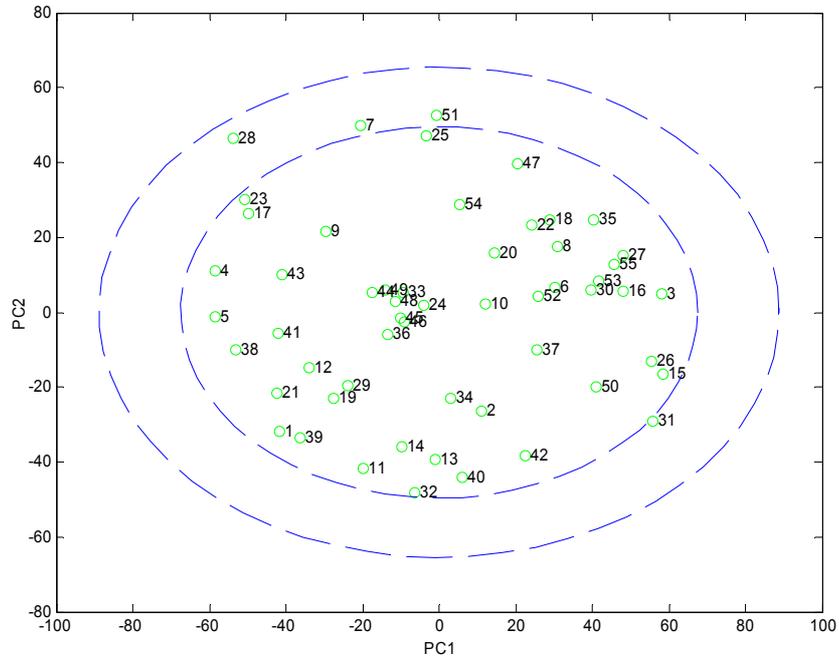


그림 5. Score plot of 55 reference batches obtained using MPCA in the space of the first and second latent vectors. The dotted lines correspond to 95 and 99% confidence limits which define the normal operation region.

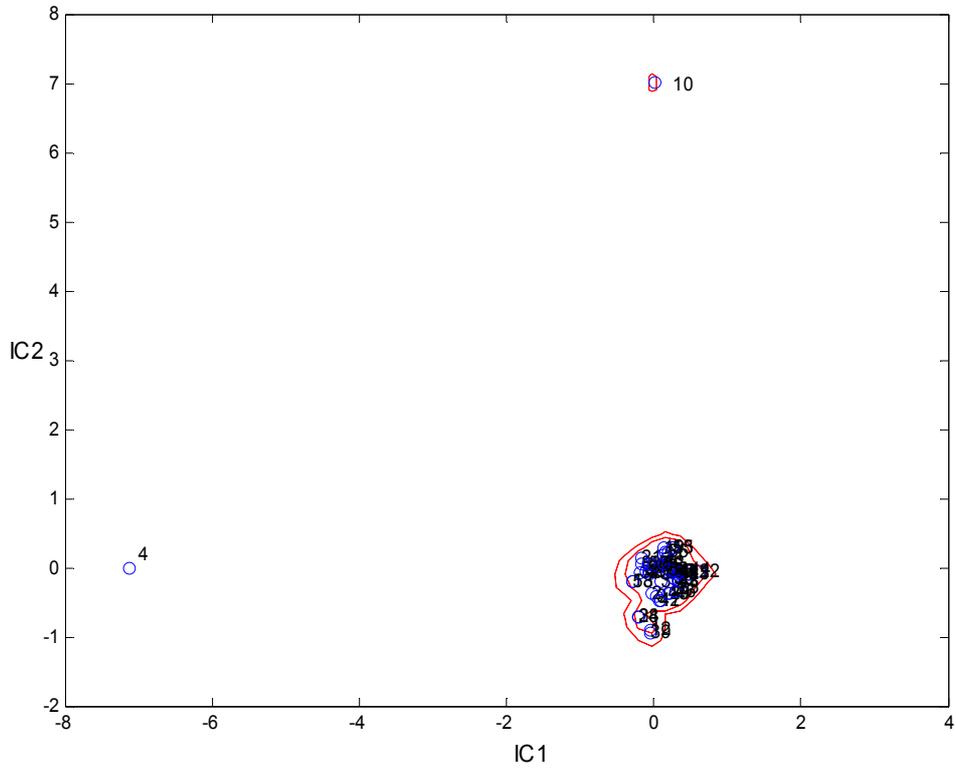


그림 6. Score plot of 55 reference batches obtained using MICA in the space of the first and second latent vectors. The dotted lines correspond to 95 and 99% confidence limits which define the normal operation region.

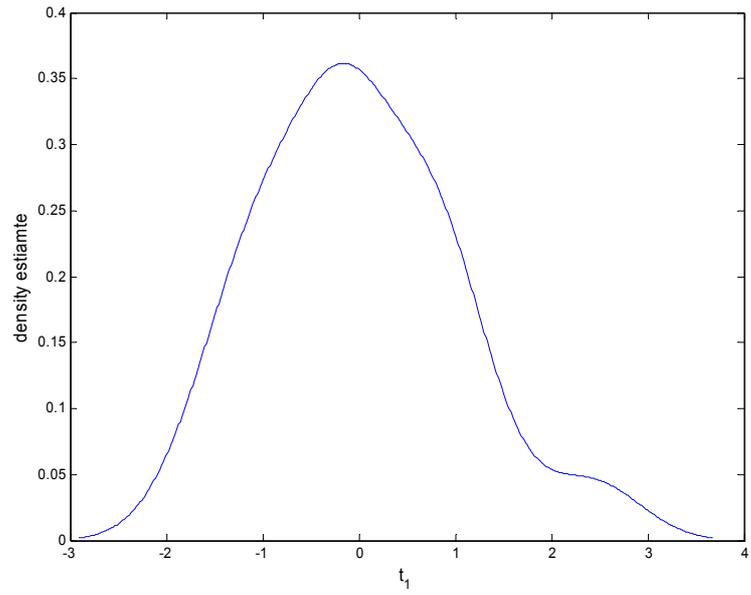


그림 7. Density estimate of the first score (t_1) obtained from MPCA.