

다변량 통계 분석을 이용한 결측 데이터의 예측과 센서이상 확인

이창규 · 이인범[†]

포항공과대학교 화학공학과
794-784 경북 포항시 남구 효자동 산 31
(2006년 10월 31일 접수, 2007년 1월 9일 채택)

Missing Value Estimation and Sensor Fault Identification using Multivariate Statistical Analysis

Changkyu Lee and In-Beum Lee[†]

Department of Chemical Engineering, POSTECH, San 31, Hyoja-dong, Nam-gu, Pohang 790-784, Korea
(Received 31 October 2006; accepted 9 January 2007)

요 약

최근 공정의 이상을 감지하고 진단하기 위한 공정 모니터링 시스템의 개발이 공정 시스템 분야에서 많은 주목을 받고 있다. 공정으로부터 얻어지는 데이터는 공정의 특성에 대한 유용한 정보를 제공하고 이는 공정의 모델링과 모니터링 그리고 제어에 사용된다. 현대의 화학 및 환경 공정은 고차원적인 특성과 변수간의 강한 상관관계와 동특성 그리고 비선형적 특성을 가지고 있어 모델 기반 접근을 통해 공정을 분석하는 것을 쉽지 않다. 이러한 모델 기반 접근의 한계를 극복하기 위해 많은 시스템 엔지니어와 연구자들이 주성분 분석법(principal component analysis, PCA) 또는 부분 최소 자승법(partial least squares, PLS)과 같은 다변량 분석을 접목한 통계 기반 접근법에 초점을 맞추고 있다. 또한 동 특성, 비선형성 등과 같은 특성을 가진 공정에 적용하기 위해 많은 다변량 분석법들이 보완되었다. 여기에서는 동적 주 성분 분석법(dynamic PCA)과 케노니컬 변수 분석법(canonical variate analysis)을 이용한 결측 데이터의 예측법과 공정 변수의 복원을 통한 센서 오작동의 판별법에 대해 언급해 보고자 한다.

Abstract – Recently, developments of process monitoring system in order to detect and diagnose process abnormalities has got the spotlight in process systems engineering. Normal data obtained from processes provide available information of process characteristics to be used for modeling, monitoring, and control. Since modern chemical and environmental processes have high dimensionality, strong correlation, severe dynamics and nonlinearity, it is not easy to analyze a process through model-based approach. To overcome limitations of model-based approach, lots of system engineers and academic researchers have focused on statistical approach combined with multivariable analysis such as principal component analysis (PCA), partial least squares (PLS), and so on. Several multivariate analysis methods have been modified to apply it to a chemical process with specific characteristics such as dynamics, nonlinearity, and so on. This paper discusses about missing value estimation and sensor fault identification based on process variable reconstruction using dynamic PCA and canonical variate analysis.

Key words: Process Monitoring, Multivariate Analysis, Missing Value Estimation, Sensor Fault Identification

1. 서 론

화학 공정은 그 특성상 수많은 단위 공정의 조합으로 이루어진 거대 장치 산업이며, 공정의 운전상에 있어서 유량, 온도, 압력, 농도 및 조성 등의 수많은 공정 변수들은 변수간의 강한 선형적 또는 비선형적인 상관관계를 가지고 있다. 이러한 고차원적인 상관관계를 분석하기 위하여 기존의 화학적 반응식을 기반으로 한 물질 및

에너지 수지를 이용하여 화학 공정 또는 단지를 모델링하고 모니터링 한다는 것은 사실상 불가능하다. 이러한 문제를 해결하기 위해 각 공정 변수의 실시간 데이터를 이용하여 통계적인 유의 구간을 설정하여 공정의 이상 유무를 판단하고 진단하는 통계 기반 공정 모니터링 시스템이 제안되었다. 초기의 통계 모니터링 시스템은 각각의 공정 변수마다의 유의 구간을 설정하고 실시간으로 측정되는 변수가 유의 구간을 벗어나게 되는 경우를 공정이상이라고 판단하는 단변수 모니터링 시스템을 사용하였다. 그러나 이러한 모니터링 시스템은 화학 공정으로부터 측정되는 공정 변수 모두를 일일이 감

[†]To whom correspondence should be addressed.
E-mail: iblee@postech.ac.kr

시해야 한다는 단점과 함께 공정 변수간의 상관관계를 고려하지 못하는 한계성을 드러내면서 다변량 통계 분석을 통한 공정 모니터링 시스템의 개발이 활기를 띠게 되었다[1]. 고차원 공정 데이터의 상관관계를 분석하기 위해 주성분 분석법(principal component analysis, PCA)이나 부분 최소 자승법(partial least squares, PLS)과 같은 다변량 분석 방법을 도입하기 시작 하였고 이를 기반으로 동적 특성 또는 비선형적인 특성을 가지는 여러 가지 공정에 적용하기 위해 많은 이론들이 개발되어 왔다[2-4]. 또한 이러한 이론들은 공정의 이상을 감지하는데 그치지 않고 공정의 이상을 확인하거나 공정 데이터의 수집 과정에서 발생하는 데이터의 결측에 관한 문제를 해결하기 위한 방법으로 발전하였다. 본 연구에서는 PCA와 CVA를 기반으로 하여 결측 데이터가 존재하는 공정 데이터를 어떻게 처리하여 공정 시스템에 유용한 모델을 찾아내는지와 실시간적으로 공정 데이터를 복원하는 방법을 통해 여러 공정의 측정 장치의 이상 유무를 확인하는 방법에 대해 언급하고자 한다.

2. 결측 데이터 처리를 통한 공정 모델링

언급한 바와 같이 데이터 마이닝에 기반한 공정 모니터링 시스템의 구축에서 가장 중요한 문제는 정상 공정 데이터의 확보라 할 수 있다. 하지만 실제 공정 데이터를 수집하는 과정에서 원치 않은 요인들로 인해 데이터의 결측이 일어나는 경우가 많다. 결측된 공정 데이터의 처리를 위하여 다변량 분석을 통한 통계 기반 접근법 이외에도 물질 및 에너지 수지 등과 같은 수학적 모델링을 이용한 접근 방법들이 제안되었다. 하지만 전연한 바와 같이 모델링에 기반한 접근법은 대상 공정에 대한 모델의 부재 및 모델에 대한 불확실성 등의 이유로 다변량 분석을 통해 이러한 결측 데이터를 처리할 수 있는 연구가 꾸준히 진행되어 왔다. 이전 연구[5]에 따르면, 결측 데이터의 정도가 전체 데이터의 20% 미만일 때 소개되는 방법이 유용하게 적용된다. 일반적으로 결측 데이터의 처리 과정을 보면, 소위 EM(expectation-maximization) 알고리즘이라고 불리는 두

가지의 단계를 거쳐 반복 계산함으로써 결측된 데이터값을 찾아간다. E단계(예측 단계)에서는 결측 데이터값을 모델로부터 구하는 단계이며, M단계(최대화 단계)에서는 공정 데이터를 이용하여 공정 시스템 구축에 필요한 모델링을 하는 단계이다. Fig. 1은 PCA와 CVA에 기반한 EM 알고리즘을 이용한 불완전한 데이터의 예측 방법을 간단하게 설명하고 있다.

공정의 특성에 맞는 어떤 모델링 방법을 적용시키느냐에 따라서 그리고 어떤 예측 방법을 채택하느냐에 따라서 불완전한 데이터를 채택하는 여러 가지 방법으로 나누어진다. 다변량 기반 공정 모니터링 시스템의 일반적인 모델링 방법으로 PCA, PLS, 그리고 캐노니컬 변수 분석법(canonical variate analysis, CVA)[6, 7] 등을 들 수가 있으며, 이러한 방법의 선택은 EM 알고리즘의 반복 계산 과정 중 M단계와 연관된다. 결측 데이터의 예측 방법 또한 여러 방법의 선택에 따라 달라지지만, 일반적으로는 확률적인 기대치를 최대화시키는 방법이 가장 좋은 방법으로 알려져있다. 그러나 이 방법은 데이터의 상관 관계 분석에 있어 랭크 부족(rank deficiency)으로 인한 역행렬 문제(inversion problem)가 생기는 데이터에 대해서는 적용할 수 없다는 단점을 가지고 있다[8]. 또 다른 방법으로는 모델공간에서 제외된 잔차공간(residual space)이나 공정의 잡음으로 해석되는 공간의 값을 최소화 시킴으로써 보다 모델에 근접한 데이터값으로 예측하는 방법이 있다. 이외에도 여러 접근 방법이 있지만, 본 연구에서는 이 두 가지의 방법을 기준으로 결측된 불완전한 데이터의 예측 방법을 설명하고자 한다.

2-1. PCA를 이용한 불완전한 데이터의 처리

공정의 모델링을 위한 방법으로 PCA를 선택하였다면, 전연한 바와 같이 결측 데이터의 예측 방법을 몇 가지 생각해 볼 수 있다. 첫째로 측정된 데이터의 조건부 기대치(conditional expectation)의 값을 결측 데이터로 예측하는 방법이다[5]. 조건부 기대치를 결측 데이터의 예측값으로 사용하기 때문에 조건부 평균 치환법(conditional mean replacement, CMR)이라고 불리는 이 방법은 초기치를 이용하여 PCA를 이용하여 모델링하고 이 모델과 결측 데이터를 제외한 나머지 데이터를 조건으로 보고 결측 데이터를 예측하게 된다. 예측된 데이터로 다시 PCA를 통해 모델링을 하여 Fig. 1에서 보듯 반복적으로 계산하여 수렴하는 값을 구하여 결측 데이터를 처리한다. 이 방법은 알고 있는 데이터를 이용하여 결측치를 선형회귀하는 방법과 동일하므로 PCA에 기반한 KDR(known data regression)이라고도 불린다[8]. 알고 있는 데이터의 정보를 모두 이용하므로 예측 능력이 상당히 뛰어나다는 장점이 있지만, 알고 있는 데이터의 선형회귀를 통해 결측 데이터를 예측함으로써 공정 변수간의 상관관계가 아주 강할 경우에는 언급한 바와 같이 역행렬 문제가 발생할 수 있다는 단점을 가지고 있다. 두번째 방법으로는 모델 공간을 제외한 잔차공간의 거리를 최소화하여 모델에 근접한 데이터를 찾아내는 방법이다. 이는 잔차공간의 거리를 목적함수로 두고 이를 결측 데이터가 존재하는 공정 변수로 편미분하여 최소값을 구할 수 있다. 이 방법은 예측 성능에서는 CMR보다는 떨어지지만, CMR에서 발생하는 역행렬 문제는 피해 갈 수 있다는 장점을 가지고 있다. 잔차공간의 거리를 최소화하여 얻어지는 결과는 실제 결측치를 포함한 데이터를 모델 공간상에 투영(projection)하여 얻어지는 값과 동일하여 이 방법을 모델 공간 투영법(projection to the model plane, PMP)이라고도 한다[8].

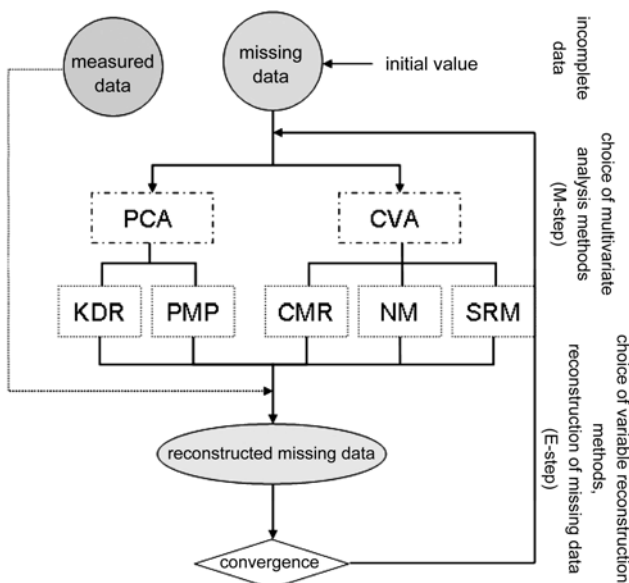


Fig. 1. PCA and CVA based EM algorithm to deal with incomplete data.

Table 1. Comparison of missing data treatment methods according to analysis strategy

PCA		CVA	
model		projected vector and variance to principal components	state space model
reconstruction method	KDR	linear regression of score vector and measured data	CMR
	PMP	minimization in residual space	NM
			SRM
			minimization in state residual space

NM : noise minimization, SRM: state residual minimization

2-2. CVA를 이용한 불완전한 데이터의 처리

동특성이 강한 공정이라면 DPCA나 소개한 CVA의 분석 방법을 채택하는 것이 타당하다. CVA는 시스템 확인(system identification) 법의 한 방법으로 연구되었던 방법으로 이를 공정 모니터링에 이용한 연구들이 최근 많이 제안되었으며 공정이상의 감지 및 확인 성능에서는 DPCA보다 탁월하다고 연구되어 있다[6, 7, 9]. DPCA와 마찬가지로 CVA를 이용하여 결측 데이터를 예측 방법을 생각해 볼 수 있다. CVA를 이용한 결측 데이터 처리 방법 역시 PCA와 같이 크게 두 가지로 볼 수 있다. 우선 첫째로 알고 있는 데이터를 기반으로 하여 상태 공간 모델(state space model)의 상(state)을 예측하는 방법이다[10]. 이 방법은 PCA를 기반으로 한 CMR과 같은 원리의 접근 방법이며, 알고 있는 데이터의 상관관계가 강할 경우 랭크 결여로 인한 역행렬 문제가 발생할 수 있다는 단점을 가지고 있다. 두 번째의 접근 방법 또한 PCA의 접근 방법과 유사한 방법이지만 최소화시키는 목적함수가 PCA와는 달리 두 가지를 생각해 볼 수 있다. CVA에서는 PCA에서의 잔차공간처럼 상(state)으로 선택되지 못하는 공간 즉 잔여 상태 공간(residual state space)뿐 아니라 상의 예측 및 출력 신호의 예측에서 발생하는 잡음 공간(noise space)이 함께 존재한다. 즉, PCA는 잔차공간의 최소화하는 하나의 방법만이 존재하지만, CVA는 잔여 상태 공간 뿐 아니라 잡음 공간에서 발생하는 크기를 최소화하는 이 두 가지 방법이 존재한다. 물론 이 두 가지 방법을 조합도 생각할 수 있다. Table 1은 PCA와 CVA를 통한 모델과 각각의 결측 데이터 방법의 특성을 비교하여 설명하고 있다.

3. 공정변수 복원을 통한 센서이상 확인

화학 공정의 운전상에서 발생하는 비이상성의 원인은 작동기(actuator)나 센서, 입출력 신호의 외란 등 여러 가지가 있다. 공정 모니터링 시스템에서는 공정의 이상을 빠르게 감지하는 것도 중요한 문제이지만 이러한 이상이 어디서 발생하는지 판단하는 과정 또한 중요한 문제이다. 기존의 공정이상에 대한 확인 방법으로 제안된 방법이 바로 기여도 도표(contribution chart)이다. 이 방법은 공정이상의 기여하는 공정 변수의 기여도를 표시함으로 해서 공정의 이상을 확인하는 방법인데, 공정이상의 전파로 인해 발생하는 모든 변수의 값이 비정상적으로 인식되어 명확한 공정이상 변수를 제시하지 못하는 한계점을 가진다. 이를 극복하기 위해 우선적으로 고려해야 할 부분이 공정이상의 특징들이다. 공정의 이상은 특성에 따라 크게 두 가지로 나뉘어 진다[11]. 첫 번째가 단순이상(simple fault)이다. 단순이상은 공정 이상이 발생한 위치에서만 계속 발생하는 형태로서 한 위치에서 발생한 공정의 이상이 다른 공정변수로 전파가 이루어지지 않는 형태이며 대표적으로 센서이상(sensor fault)을 들 수 있다. 두 번째는 한 곳에서 발생한 이상이 다른 곳으로 전파되는

형태인데 이를 복합이상(complex fault)이라고 한다. 예를 들어 작동기와 같은 공정의 전체에 영향을 미치는 요소에 이상이 발생하였을 경우, 공정의 이상이 작동기에 한해 발생하기는 하지만 다른 공정의 여러 요소들에 영향을 미치게 된다. 또 다른 예로 센서이상이라고 하더라도 센서가 제어기와 같은 공정 운전 전반에 영향을 미치는 공정변수와 연관이 되어 있다면 이는 복합이상으로 해석하는 것이 타당하다. 복합이상에는 그 위치가 동일하더라도 전파되는 이상의 패턴이 상당히 복잡하여 해석하기가 힘들다. 복합이상을 확인하기 위한 방법으로 일반적으로 제안된 방법은 패턴 분류(pattern classification)나 SDG(signed directed graph)와 같은 방법들이 제안되었다. 단순이상은 복합이상에 비해 전파되는 경로를 해석할 필요가 없기 때문에 이상의 확인이 비교적 용이하다. 단순이상은 복합이상인 여러 위치에서 동시 이상이 발생하는 경우를 다중이상(multiple fault)라고 한다.

다변량 통계 분석에 기반한 공정 모니터링 시스템에서 가장 기본적으로 갖춰져야 할 부분은 공정의 정상 운전 데이터의 확보이며, 센서 오작동의 확인은 이러한 기본 조건을 만족시키기 위한 것에 연구 목적의 의의를 두고 있다. 본 연구에서는 한 곳에서 발생하는 단순이상의 대표적인 경우인 센서 오작동에 대해 실시간으로 확인 가능한 공정변수 복원에 기반한 센서 오작동 확인 방법에 대해 언급하고자 한다. 공정변수 복원을 이용한 센서 오작동의 확인은 앞서 언급한 결측 데이터의 예측 방법과 동일한 방법을 사용한다. 결측 데이터 처리 과정에서는 E단계와 M단계를 반복 계산을 하지만, 선택한 다변량 분석 방법에 맞게 공정이 모델링되었다고 한다면 M단계는 생략된다. 즉, 실시간 측정되는 데이터를 이용하여 연속적으로 모든 공정 변수들에 대한 예측치를 계산할 수 있게 되고 현재 측정되는 값과의 비교를 통하여 센서의 이상을 확인 할 수 있다. 일반적인 비교의 수단은 실시간으로 복원된 공정 측정치의 공정 모니터링 인자(process monitoring indices)를 이용한다. 복원된 데이터의 신뢰 구간은 선형 통계적인 관계식을 이용하여 유도할 수도 있지만, 정상으로 수집된 데이터의 경험적인 신뢰 구간(empirical reference distribution, ERD)[12]을 이용하는 것도 하나의 방법이라고 할 수 있다. Fig. 2는 공정 데이터의 복원을 통한 센서이상 감지 시스템의 구조를 설명하고 있다.

4. 공정 데이터 복원을 통한 센서이상 확인 예제

본 연구에서 언급된 공정 데이터 복원을 통한 센서이상의 확인 방법을 간단한 공정 시뮬레이션을 통해 적용해 보았다. 적용된 대상 공정은 1차 반응이 이루어지는 CSTR(continuous stirred tank reactor)로 공정 모니터링을 위한 측정 변수로 9개의 센서를 사용하였다. 공정의 구조도와 측정 변수 및 간단한 물성은 Fig. 3과 Table 2

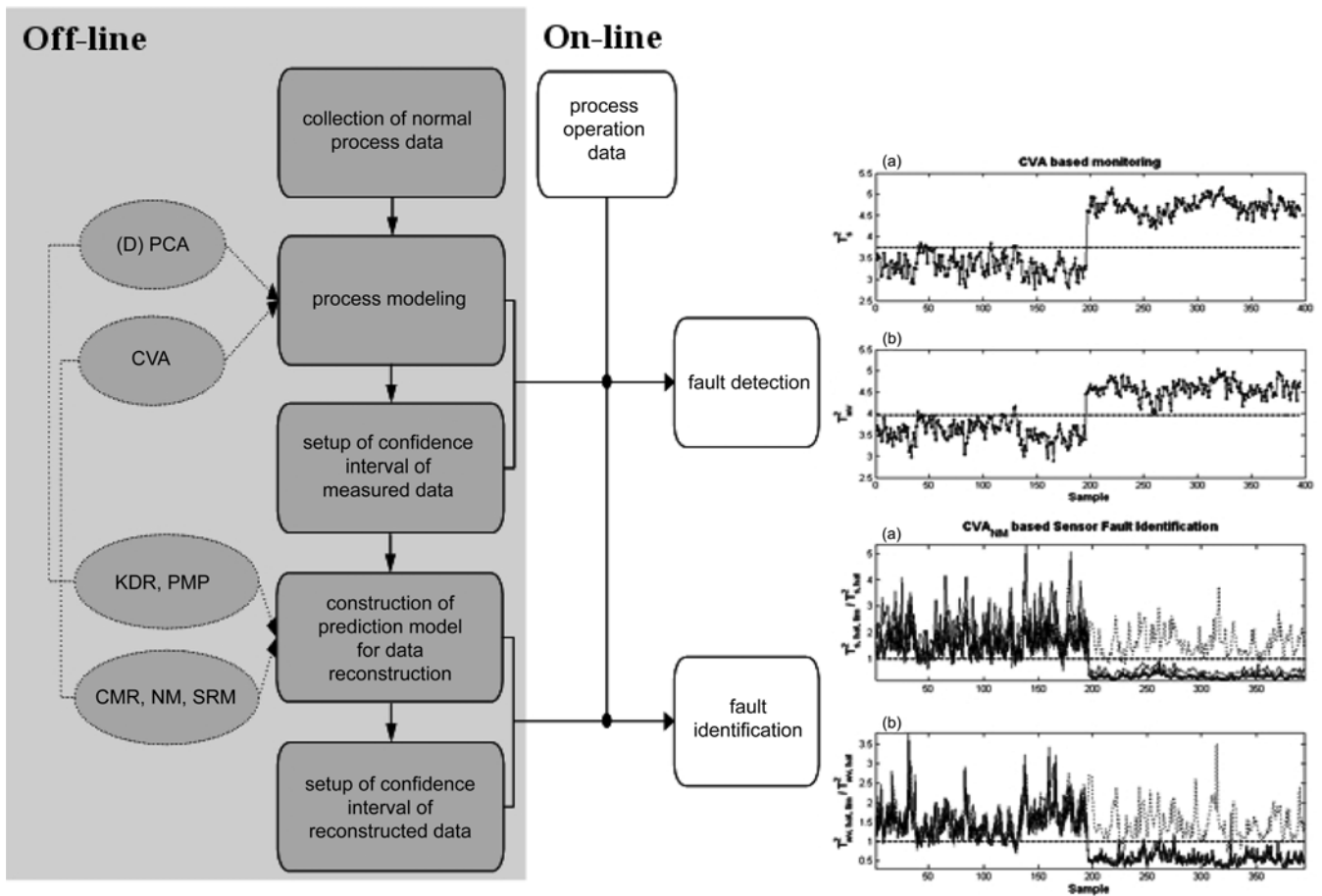


Fig. 2. Structure of sensor fault identification system based on process variable reconstruction.

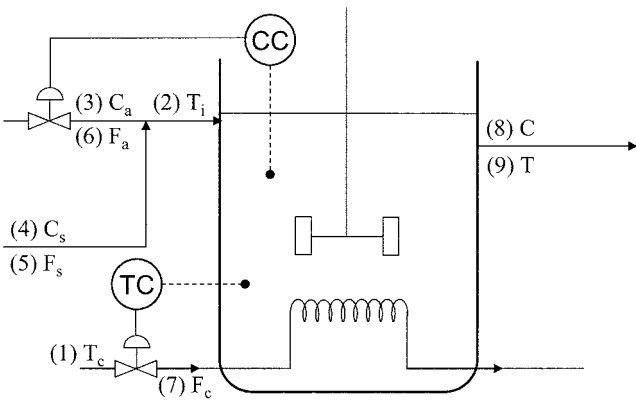


Fig. 3. The simulated CSTR process.

에 설명되어 있다. 공정 시뮬레이터에 대한 자세한 설명은 여기서는 생략하였다[11, 13].

센서 이상의 종류는 더 많은 경우를 따져 볼 수 있지만 센서가 완전히 측정을 하지 못하는 경우(completely broken), fouling 등에 의해 갑작스레 측정의 정확도가 떨어지거나(precision degradation), 시간이 지남에 따라 점점 실제 값에서 동떨어지는 값을 나타내는 경우(drifting), 또는 실제 값에서 편중된 값을 측정 하는(bias) 4가지

Table 2. Information of CSTR process

process variable
state variable : C, T
control variable : C, T
manipulating variable : F_s, F_c
measured variable : $T_c, T_i, C_a, C_s, F_s, F_a, F_c, C, T$
CSTR model information
$V=1 \text{ m}^3; \rho=106 \text{ g/m}^3; \rho_c=106 \text{ g/m}^3; C_p=1 \text{ cal/g/K}; C_{pc}=1 \text{ cal/g/K};$ $k_0=1,010 \text{ min}^{-1}; a=1.678 \times 10^6 \text{ cal/min}; b=0.5; \Delta H_r=-1.3 \times 10^7 \text{ cal/kmol}$

정도의 현상을 고려해 볼 수 있다. 본 예제에서는 모델링 방법으로 CVA를 선택하였으며, 모델링을 위한 지연 시간(lag time) 수는 4로, 정상 공정 운전하에 수집된 1,000개의 샘플 데이터를 데이터 분석을 통하여 전체 상의 90% 이상을 표현하는 상의 수를 상공간으로 선택하였다[10]. 센서의 이상 시나리오로 센서가 정상적으로 작동되다가 정확도가 저하되는 현상을 모사하였다. 공정 변수의 복원 방법으로는 CMR를 선택하였으며, 복원 데이터의 신뢰 구간은 수집된 샘플 데이터를 기준으로 95%의 범위로 설정하였다. Fig. 4[9]는 공정 모사의 결과와 모니터링 방법을 이용하여 공정의 이상을 감지하고 센서 이상을 확인한 결과를 보여 주고 있다.

Fig. 4(a)는 샘플링 시간 200 이후에 냉각수 유량의 측정 센서가

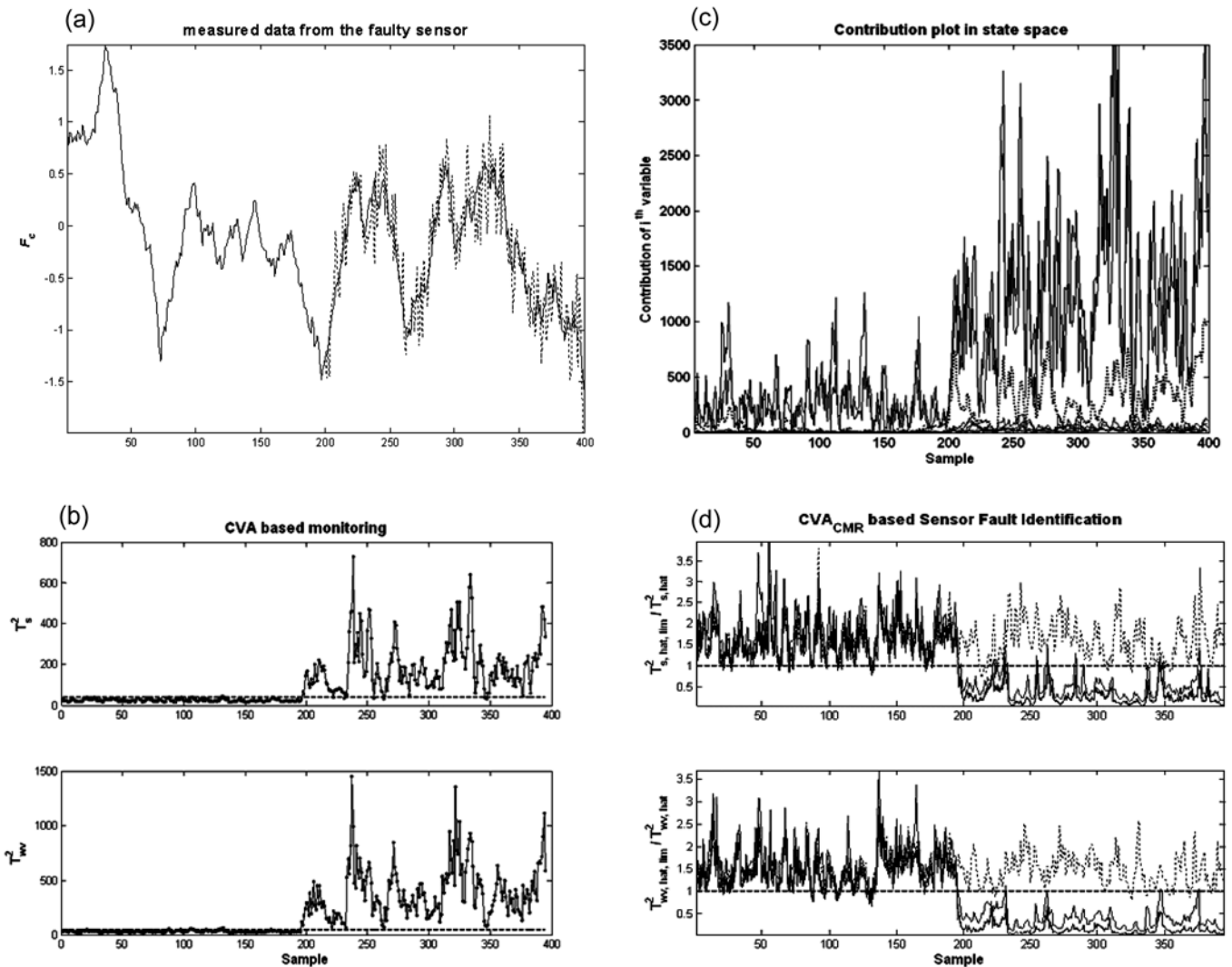


Fig. 4. Results of sensor fault detection and identification.

정확하게 측정하지 못하고 있음을 보여 주고 있다. Fig 4(b)는 CVA에 기반한 공정 이상 감지의 결과를 보여 주고 있다. Fig 4(b)의 첫 번째 그림은 상공간에서의 공정 이상 감지를, 두 번째 그림은 잡음 공간에서의 공정 이상을 감지한 결과이며 두 공간에서 모두 샘플링 시간 200 이후에 나타나는 공정의 이상을 감지하고 있다. 변수 복원을 통한 센서의 이상 감지 결과가 Fig. 4(d)에 도식화되어 있다. 상공간에서의 변수 복원을 통한 센서이상의 확인 결과를 Fig. 4(d)의 첫 번째 그림에 표현되어 있으며, 정상적으로 측정하고 있는 센서(실선)와 오작동되는 센서(점선)가 명확하게 구분되고 있다. Fig. 4(d)의 두 번째 그림에서 잡음공간에서의 센서 이상 감지 역시 성공적으로 이루어 지고 있음을 보여 주고 있다. 기존에 제안된 CVA에 기반한 기여도표를 Fig. 4(c)에 도식화하였다. 이 접근 방법으로는 센서의 이상 위치(점선)를 명확하게 구분할 수 없다. 변수 복원을 통한 공정의 이상 감지가 기존의 제안된 방법보다 센서이상 확인 성능이 보다 좋음을 확인 할 수 있다. CVA를 기반한 접근법이 여러 가지의 센서이상에 대해서 탁월한 센서이상 감지 성능을 보이고 있음을 이전 연구를 통해 알 수 있다[9].

5. 결 론

다변량 통계기반 공정 모니터링 시스템의 구축에 있어 가장 중요시 되는 부분 중의 하나인 공정에서 발생하는 결측 데이터의 복원 방법과 이러한 복원 방법을 응용하여 오작동 센서의 이상 확인 방법에 대해 본 연구에서는 언급하고 있다. 대표적인 다변량 분석 방법인 PCA와 CVA가 결측 데이터의 처리 방법인 EM 알고리즘에 어떻게 적용되는지에 대해 언급하였으며, 각 다변량 분석 방법을 이용한 여러 가지 결측 데이터의 복원 방법에 대한 원리에 대해 언급하였다. 또한 결측 데이터의 복원 방법을 센서 오작동 확인에 응용하는 원리에 대해서도 언급하였다. 공정이상 확인 측면에서 기존에 제안되었던 방법에 대한 한계성[15]이 밝혀짐에 따라 공정의 특성에 맞는 다변량 분석 방법의 선택과 채택된 분석 방법의 장점과 단점을 올바르게 이해하고 대상 공정에 올바르게 적용한다면 해석하기 힘든 화학 공정의 안정성 확보와 양질의 제품을 생산할 수 있는 향상된 공정 모니터링 시스템을 구축할 수 있을 것이다.

감 사

본 연구는 Brain Korea 21 지원으로 수행되었으며 지원에 감사를 드립니다.

참고문헌

1. MacGregor, J. F. and Kourti, T., "Statistical Process Control of Multivariate Processes," *Cont. Eng. Prac.*, **3**(3), 403-414(1995).
2. Choi, S. W., Lee, C., Lee, J.-M., Park, J. H. and Lee, I.-B., "Fault Detection and Identification of Nonlinear Processes Based on Kernel PCA," *Chem. Int. Lab. Sys.*, **75**, 55-67(2005).
3. Cho, J.-H., Choi, S. W., Lee, D. and Lee, I.-B., "Fault Identification for Process Monitoring Using Kernel Principal Component Analysis," *Chem. Eng. Sci.*, **60**, 279-288(2005).
4. Ku, W., Storer, R. H. and Georgakis, C., "Disturbance Detection and Isolation by Dynamic Principal Component Analysis," *Chem. Int. Lab. Sys.*, **30**, 179-196 (1995).
5. Nelson, P. R. C., Taylor, P. A. and MacGregor, J. F., "Missing Data Methods in PCA and PLS; Score Calculations with Incomplete Observations," *Chem. Int. Lab. Sys.*, **35**, 45-65(1996).
6. Russell, E. L., Chiang, L. H. and Braatz, R. D., "Fault Detection in Industrial Processes Using Canonical Variate Analysis and Dynamic Principal Component Analysis," *Chem. Int. Lab. Sys.*, **51**, 81-93 (2000).
7. Negiz, A. and Cinar, A., "Statistical Monitoring of Multivariable Dynamic Processes with State Space Model," *AIChE J.*, **43**, 2002-2020(1997).
8. Arteaga, F. and Ferrer, A., "Dealing with Missing Data in MSPC: Several Method, Different Interpretations, Some Examples," *J. Chemometrics*, **16**, 408-418(2002).
9. Lee, C., Choi, S. W. and Lee, I.-B., "Variable Reconstruction and Sensor Fault Identification Using Canonical Variate Analysis," *J. Process Control*, **16**, 747-761(2006).
10. Larimore, W. E., "Canonical Variate Analysis in Identification, Filtering, and Adaptive Control," Proceedings of IEEE Conference on Decision and Control, Honolulu, Hawaii, 596-604(1990).
11. Yoon, S. and MacGregor, J. F., "Fault Diagnosis with Multivariate Statistical Models Part I: Using Steady State Fault Signatures," *J. Process Control*, **11**, 387-400(2001).
12. Willemain, T. R. and Runger, G. C., "Designing Control Charts Using an Empirical Reference Distribution," *J. Quality Technology*, **28**, 31-38(1996).
13. Marlin, T. E., *Process Control*, McGraw-Hill, New York(1995).
14. Conference report, "Abnormal Situation Detection and Projection Methods-industrial Applications," *Chem. Int. Lab. Sys.*, **76**, 215-220(2005).
15. Lee, C., Choi, S. W., Lee, J.-M. and Lee, I.-B., "Sensor Fault Identification in MSPM Using Reconstructed Monitoring Statistics," *Ind. Eng. Chem. Res.*, **43**(15), 4293-4304(2004).