

이상치 탐지 방법론을 활용한 반도체 가상 계측 결과의 신뢰도 추정

장필성¹ · 김동일² · 이승경² · 도승용³ · 조성준^{2*}

¹서울과학기술대학교 국제융합학부 IT Management 과정
/²서울대학교 산업공학과 /³삼성 SDS

Estimating the Reliability of Virtual Metrology Predictions in Semiconductor Manufacturing : A Novelty Detection-based Approach

Pilsung Kang¹ · Dongil Kim² · Seung-Kyung Lee² · Seungyong Doh³ · Sungzoon Cho²

¹IT Management Programme, International Fusion School, Seoul National University of
Science and Technology(Seoultech)

²Industrial Engineering, Seoul National University

³Samsung SDS

The purpose of virtual metrology (VM) in semiconductor manufacturing is to predict every wafer's metrological values based on its process equipment data without an actual metrology. In this paper, we propose novelty detection-based reliability estimation models for VM in order to support flexible utilization of VM results. Because the proposed model can not only estimate the reliability of VM, but also identify suspicious process variables lowering the reliability, quality control actions can be taken selectively based on the reliance level and its causes. Based on the preliminary experimental results with actual semiconductor manufacturing process data, our models can successfully give a high reliance level to the wafers with small prediction errors and a low reliance level to the wafers with large prediction errors. In addition, our proposed model can give more detailed information by identifying the critical process variables and their relative impacts on the low reliability.

Keywords: Virtual Metrology, Reliability, Novelty Detection, Semiconductor Manufacturing, Data Mining

1. 서론

최근 스마트폰을 포함한 다양한 디지털 기기의 급속한 성장으로 인하여 IT 기기의 핵심 부품인 반도체에 대한 수요가 급격하게 증가하고 있으며 이에 따라 정보를 처리하고 저장하는 시스템 및 메모리 반도체 관련 산업 또한 2009년부터 연평균 13.6%의 성장을 지속하여 2012년 기준 3,315억 달러의 시장 규모가 될 것으로 추산되고 있다. 반도체 제조 공정에서 투입되는

웨이퍼(wafer)는 일반적으로 약 500개의 개별 공정(operational process)를 거쳐 최종 제품으로 생산되는데, 제조 공정의 높은 복잡도와 향상되는 정밀도로 인하여 미세하고 정확한 품질관리가 요구되며, 이를 위해 제조사들은 계측(metrology)과 통계적 공정 관리(statistical process control; 이하 SPC) 기법에 기반한 품질 관리에 많은 노력을 기울이고 있다(Spanos *et al.*, 1992; Sachs *et al.*, 1995; Bergeret and Gall, 2003). 계측 기반 SPC의 일반적인 절차는 다음과 같다. 첫째, 개별 공정에서는 계획된 공

본 연구는 서울과학기술대학교 교내 학술연구 지원비로 수행되었습니다.

* 연락저자 : 조성준 교수, 151-742 서울시 관악구 신림동 산56-1번지 서울대학교 공과대학 산업공학과, Tel : 02-880-6275, Fax : 02-889-8654,
E-mail : zoon@snu.ac.kr

2011년 9월 21일 접수; 2012년 2월 14일 게재 확정.

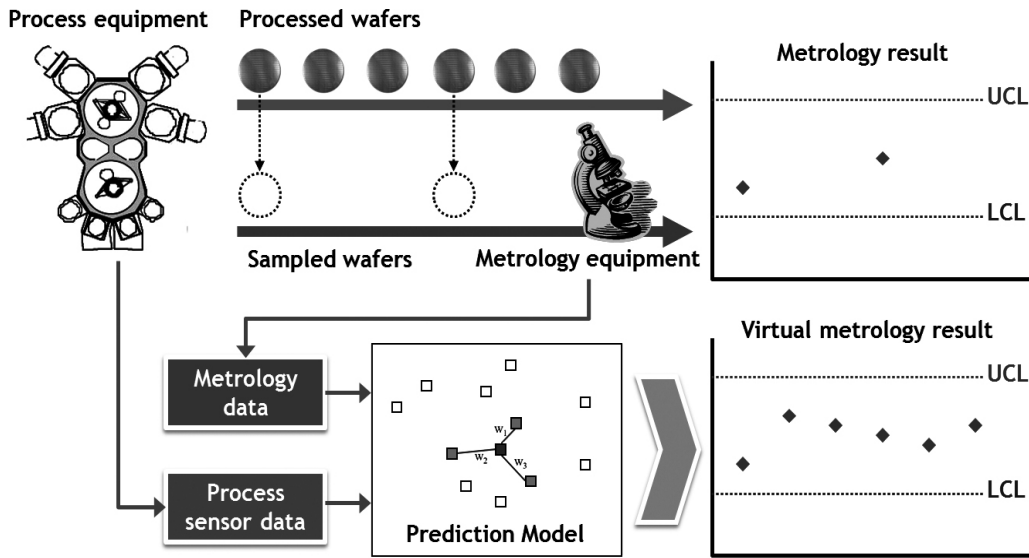


Figure 1. The difference between an actual metrology(top) and a virtual metrology(bottom)

정 레시피(recipe)를 사용하여 한 랫(lot)의 웨이퍼들에 대한 작업을 수행한다. 둘째, 해당 랫의 웨이퍼들이 정상적으로 작업이 수행되었는지를 판단하기 위하여 소수의 웨이퍼를 샘플링하여 계측 공정에 투입한다. 계측 공정에서는 샘플링된 웨이퍼들을 검사하여 그 결과가 공정에서 요구하는 기준을 충족할 경우 해당 웨이퍼가 속한 랫의 모든 웨이퍼를 정상으로 판별하여 후속 공정에 투입하고, 그렇지 않을 경우 재작업 또는 폐기 등의 후처리를 수행하게 된다. 계측 기반 SPC는 현재 대다수의 제조 공정에서 사용되는 품질 관리 기법이지만 다음과 같은 문제점을 내포하고 있다. 첫째, 품질 관리를 위한 별도의 계측 공정을 추가함으로써 전체 생산 시간이 증가한다(Chen *et al.*, 2005; Su *et al.*, 2007). 계측 공정은 작업 공정과는 달리 웨이퍼에 특정한 가공을 하여 가치를 증가시키는(value-added) 공정이 아니기 때문에 전체 공정 수를 증가시킬 뿐만 아니라, 샘플링 웨이퍼에 대한 계측 결과 분석이 완료될 때까지 해당 웨이퍼가 추출된 랫은 이후 공정에 투입되지 못하기 때문에 전반적인 생산 시간의 증가를 초래하게 된다. 둘째, 계측 공정은 샘플링 기법을 사용하기 때문에 모든 웨이퍼에 대한 품질 관리가 불가능하게 된다(Chen *et al.*, 2006; Lin *et al.*, 2006). 샘플링 기반 품질 관리는 계측 공정에 투입되는 웨이퍼가 해당 웨이퍼가 속한 랫의 모든 웨이퍼의 품질을 대표하는 것을 가정한다. 그러나 공정의 변동성을 포함한 여러 가지 이유로 인해 동일한 랫의 웨이퍼들 사이에도 품질의 편차가 존재하는 것이 일반적이다.

이러한 계측 기반 SPC의 단점을 해결하기 위하여 최근 가상 계측(virtual metrology, VM)에 대한 관심이 집중되고 있다. 가상 계측의 목적은 실제로 계측 공정을 수행하지 않고서도 모든 웨이퍼에 대한 계측 결과를 예측하는 것이며, 실 계측(actual metrology)와의 차이는 <Figure 1>에 나타난 바와 같다. 실 계측의 경우, 공정 설비(process equipment)에 투입되어 해당 공정의 레시피에 따라서 작업이 수행된 한 랫의 웨이퍼들

중에서 일부만을 추출하여 품질 관련 지표(확대, 축소, 변형, 회전 등)를 측정한다. 측정된 지표들이 허용 범위 이내에 존재할 경우 해당 웨이퍼가 추출된 랫의 모든 웨이퍼들을 정상으로 판별하여 후속 공정에 투입하고, 측정값이 허용 범위를 벗어날 경우 그 정도에 따라서 추가 작업 등을 진행한다. 반면에, 가상 계측은 추출된 웨이퍼들의 공정 설비 데이터를 설명 변수(explanatory variables)로 하여 해당 웨이퍼들을 실제로 계측함으로써 얻은 품질 지표들을 목표 변수(target variables)로 하는 예측 모델(prediction model)을 구축하게 된다. 이 예측 모델이 공정 설비 데이터와 계측 품질 지표들과의 관계를 명확히 규명하는 경우에는 실제 계측 공정을 수행하지 않더라도 모든 개별 웨이퍼에 대한 계측 정보들을 얻을 수 있게 된다.

반도체 제조 공정에서 좋은 가상 계측 모델을 구축할 경우 다음과 같은 긍정적인 효과를 기대할 수 있다. 첫째, 모든 웨이퍼에 대한 계측 정보를 가지고 공정 관리를 수행함으로써 최종 수율을 향상시킬 수 있다(Chang *et al.*, 2006). 둘째, 실제 계측 공정에 투입되는 웨이퍼의 수를 크게 감소시킴으로써 전체 생산 시간의 단축과 계측 공정에 소요되는 비용을 감소시킬 수 있다(Chen *et al.*, 2005). 마지막으로 개별 웨이퍼에 대해 연속적인 모니터링을 수행함으로써 프로세스 변화를 실시간으로 탐지할 수 있고(Lin *et al.*, 2006), 배치(batch) 단위가 아닌 개별 웨이퍼 단위(wafer by wafer, run-to-run; R2R)의 공정 관리를 가능하게 한다(Khan *et al.*, 2007).

앞서 언급된 다양한 장점으로 인하여 가상 계측 모델 구축에 관련된 많은 연구가 진행되어 왔으며, 이들은 크게 (1) 새로운 예측 알고리즘 개발 및 효과적인 공정 변수 선택을 통한 예측 성능 향상과 (2) 가상 계측과 Run-to-Run 제어(이하 R2R 제어)의 결합을 통한 실시간 공정 관리 시스템 구축으로 구분될 수 있다. 예측 성능 향상에 관련된 대표적인 연구들은 다음과 같다. Besnard and Toprac(2006)은 정규분포를 따르지 않

는 공정 설비 변수들을 제거한 뒤 의사 결정 회귀(decision tree regression) 기법을 예측 알고리즘으로 사용하여 실제 계측 결과와 약 85%의 상관관계를 나타내는 가상 계측 모델을 구축하였으며, Lin *et al.*(2006)은 주성분 분석(principal component analysis; PCA)을 사용하여 변수의 수를 감소시킨 뒤 radial basis function(RBF) 네트워크를 이용한 회귀 모델을 구축하여 1% 이내의 MAPE(mean absolute percentage error)를 달성하였다. Pan *et al.*(2011)은 다수의 공정 장비(tool)가 동일한 작업을 수행할 때 발생하는 변동성을 클러스터링 기법과 다변량 분산분석(multi-variate analysis of variance)을 사용하여 가상 계측 모델에 반영함으로써 상대적으로 단순한 선형 회귀 분석을 예측 모델로 사용함에도 불구하고 매우 낮은 MAPE를 나타낼 수 있음을 보였고, Lynn *et al.*(2010)은 각 공정 설비 변수의 중요도를 반영한 부분 최소 제곱 회귀(weighted partial least squares regression) 기법을 사용하여 가상 계측의 예측 정확도 향상을 추구하였다.

가상 계측과 R2R 제어의 결합을 통한 실시간 공정 관리 시스템 구축에 관련하여 Qin *et al.*(2006)은 가상 계측 및 불량 탐지(fault detection control; FDC)를 결합한 전체 공정 수준(fab-wide)의 R2R 제어 프레임워크를 제시하고 이의 성공을 위해 필수적으로 해결해야 하는 예측 모델 업데이트, FDC 계측의 결합 등의 주요 이슈들을 제기하였다. 또한 Khan *et al.*(2007)은 개별 공정에 가상 계측 모델을 내재화(embed)하고 연속적인 두 공정 사이의 가상 계측 모델들 간의 정보 교류가 가능하게 함으로써 가상 계측 결과의 정확도 향상과 R2R 제어의 유연성 향상을 동시에 도모하였다.

이와 같이 가상 계측에 관련된 기존 연구들은 가상 계측 자체의 정확도를 향상시키는 방법론을 개발하는 것과 가상 계측의 결과가 정확하다는 가정 하에 이를 활용하여 R2R 제어를 구현함으로써 전체 반도체 공정의 생산성을 향상시키는 것에 초점을 맞추고 있다. 그러나 가상 계측 모델은 많은 장점이 있는 반면에 다음과 같이 두 가지의 위험(risk)을 내포하고 있다. 첫 번째 위험은 모델 위험(model risk)이다. 즉, 가상 계측 모델이 정확한 예측을 하지 못함으로써 발생하는 위험으로, 정확하지 않은 계측 결과를 바탕으로 공정 제어를 수행할 경우 바람직하지 않은 방향으로 최종 결과가 도출될 수 있다. 두 번째 위험은 데이터 위험(data risk)이다. 정확한 가상 계측 모델은 모델 구축 시 사용된 공정 설비 데이터와 유사한 새로운 웨이퍼에 대한 예측은 매우 높은 정확도로 수행될 수 있음을 의미한다. 반면에 모델 구축 시 사용되지 않은 공정 설비 값을 갖는 웨이퍼에 대해서는 어느 정도 일반화를 통한 예측은 가능할 지라도 정확한 예측이 어려운 점이 있으며, 이러한 결과를 바탕으로 공정 제어를 하는 것은 매우 큰 위험을 초래할 수 있다. 모델 위험을 감소시키기 위해서 이미 앞서 언급된 바와 같이 많은 연구자들에 의해 정확한 예측 알고리즘을 개발하기 위한 다양한 연구가 시도되고 있는 반면에, 데이터 위험을 감소시키기 위한 연구는 거의 시도되지 않고 있는 실정이다. 따라서

본 연구에서는 데이터 위험을 감소시키기 위하여 이상치 탐지 기법을 활용하여 가상 계측 결과에 대한 신뢰도를 추정하는 방법론을 제안한다. 또한 신뢰도가 낮게 추정될 경우, 이의 원인이 되는 공정 설비 변수를 규명하고 그 영향력을 측정하는 기법을 제안한다. 이를 통하여 가상 계측 결과를 추정된 신뢰도에 따라 취사선택하여 공정 제어에 활용함으로써 반도체 공정의 생산성 향상을 기대할 수 있다.

본 연구의 나머지 부분은 다음과 같이 구성된다. 제 2장에서는 이상치 탐지 기반 가상 계측 신뢰도 추정 방법론을 제안한다. 제 3장에서는 제안된 방법론의 효과를 분석하기 위하여 실제 반도체 공정 설비 데이터를 사용한 실험 설계를 설명한다. 제 4장에서는 제안된 방법론의 효과를 분석하고 시사점을 도출하며, 마지막으로 제 5장에서는 결론과 함께 본 연구의 한계점을 언급하고 이를 개선하기 위한 향후 연구 방향에 대한 논의를 할 것이다.

2. 이상치 탐지 기반 가상 계측 결과 신뢰도 및 설비 변수 영향력 추정

본 연구에서 제안하는 이상치 탐지 기반 가상 계측 결과 신뢰도 및 설비 변수 영향력 추정 시스템의 개념 및 기존 가상 계측 시스템과의 차이는 <Figure 2>와 같다. <Figure 1>의 기존 가상 계측 시스템에서는 계측이 수행된 웨이퍼들의 계측 결과와 설비 데이터를 사용하여 예측 모델(prediction model)을 구축하고, 새로운 웨이퍼의 설비 데이터가 수집되면 이를 예측 모델에 투입하여 계측 결과를 예측하게 된다. 그러나 본 연구에서 제안하는 가상 계측 결과 신뢰도 및 설비 변수 영향력 추정 시스템에서는 예측 모델 구축에 사용된 설비 데이터만을 이용하여 이상치 탐지(novelty detection) 모델을 구축하는 단계가 추가된다. 새로운 웨이퍼의 설비 데이터가 수집되면 이 데이터는 예측 모델과 이상치 탐지 모델에 동시에 투입되어 예측 모델에서는 계측 결과를 예측하게 되고, 이상치 탐지 모델에서는 해당 웨이퍼의 데이터와 예측 모델 구축 당시의 설비 데이터와의 유사도를 추정한다. 이상치 탐지 모델에 의해 유사도가 높게 나타날 경우 가상 계측 모델 구축 데이터와 동질의(homogeneous) 데이터로 판단하여 가상 계측 결과에 높은 신뢰도를 부여하고, 유사도가 낮게 나타날 경우 이질적(heterogeneous) 데이터로 판단하여 가상 계측 결과에 낮은 신뢰도를 부여한다. 또한 이상치 탐지 모델에서는 데이터의 유사성에 대한 각 설비 변수의 영향력 또한 추정하여 가상 계측 결과의 신뢰도가 낮게 부여될 경우 그 원인이 되는 설비 변수와 영향력을 동시에 제공한다.

2.1 가상 계측 모델

본 연구에서는 다중 선형 회귀 분석(multiple linear regression;

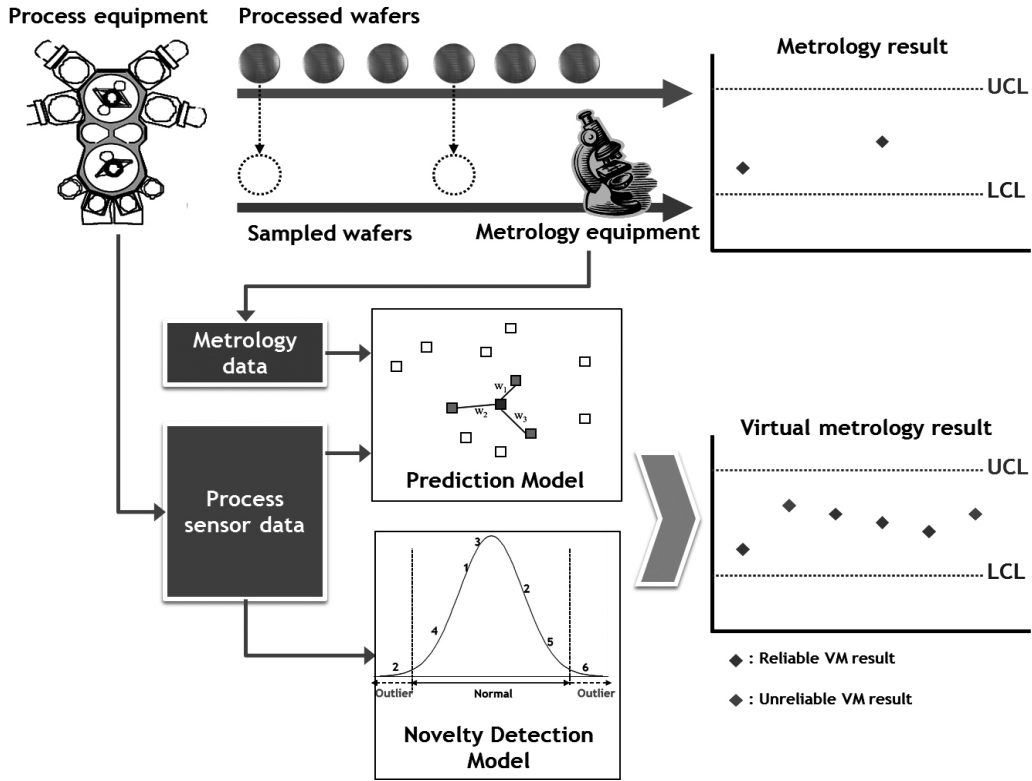


Figure 2. The proposed novelty detection-based virtual metrology prediction reliability and process sensor parameter sensitivity estimation system

이하 MLR)을 가상 계측 시스템의 예측 알고리즘으로 사용하였다. MLR은 수집된 데이터를 바탕으로 두 개 이상의 독립변수 (independent variables)와 종속변수(dependent variable)의 관계를 선형식(linear equation)으로 추정하는 방법론으로써, 인공 신경망 (artificial neural network; ANN), 지지 벡터 회귀(support vector regression; SVR) 등의 복잡한 기계 학습 방법론에 비해 단순한 구조와 효율적인 계산 복잡도 및 적은 사용자 지정 모수(user-specific parameter)를 갖는 장점으로 인해 여러 분야의 예측 시스템에 널리 사용되고 있다. 본 연구에서는 가상 계측 모델의 정확도 향상보다는 예측 결과의 신뢰도 및 예측 변수의 영향력 추정에 초점을 맞추고 있기 때문에 기본적인 예측 알고리즘인 MLR을 사용하였다.

m 개의 독립 변수를 갖는 n 개의 웨이퍼에 대해 MLR의 회귀 식은 다음과 같이 표현된다.

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m. \quad (1)$$

여기서 y 는 예측하고자 하는 계측 변수이고, x_i 는 예측에 사용되는 공정 설비 변수이며, β_i 는 i 번째 설비 변수에 대한 회귀 계수이다. 이를 행렬식의 형태로 표현하면 식 (2)와 같다.

$$\hat{y} = \mathbf{X}\beta, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1m} \\ 1 & x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}. \quad (2)$$

이와 같은 MLR의 회귀 계수 $\beta = (\beta_0, \beta_1, \dots, \beta_m)^T$ 는 식 (3)의 잔차 제곱합을 최소로 하는 최소 제곱 회귀법(ordinary least squares; OLS)을 사용하여 식 (4), 식 (5)와 같이 추정할 수 있다.

$$E = \frac{1}{2} \sum_{i=1}^n e_i^2 = \frac{1}{2} \sum_{i=1}^n (\mathbf{y} - \mathbf{X}\beta)^T (\mathbf{y} - \mathbf{X}\beta). \quad (3)$$

$$\frac{\partial E}{\partial \beta} = \mathbf{X}^T \mathbf{y} - \mathbf{X}^T \mathbf{X} \beta = 0. \quad (4)$$

$$\beta = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}. \quad (5)$$

2.2 이상치 탐지 기반 입력 데이터 유사도 및 설비 변수 영향력 추정 모델

본 연구에서는 가상 계측 테스트 웨이퍼의 설비 데이터와 가상 계측 모델 구축에 사용된 웨이퍼들의 설비 데이터와의 동질성(homogeneity)을 측정하여 이를 바탕으로 가상 계측 결과에 대한 신뢰도를 부여하고, 신뢰도에 대한 각 설비 변수들의 영향력을 측정하기 위해 이상치 탐지(novelty detection) 기법을 사용하였다. 이상치 탐지 기법은 주어진 데이터를 모두 하나의 범주로 가정하고 학습한 뒤 새로운 데이터에 대해 기존 데이터와의 유사도를 측정하여 충분히 유사할 경우 동일 범주로 판단하고 그렇지 않을 경우 다른 범주로 판단하는 기계 학습(machine learning) 기법으로써, 특정 범주에 속하는 데이터를 수집하기 어려운 상황이나 범주 간 불균형이 심한 경우 이범주 분류(binary classification) 기법에 비해 효과적인 분류

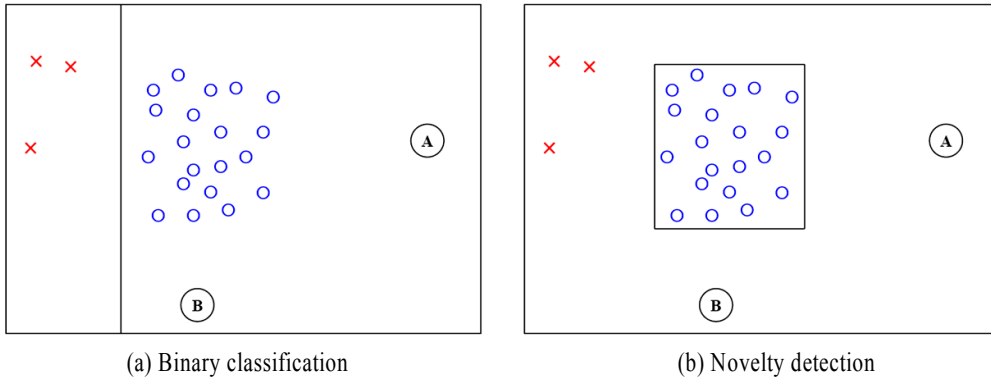


Figure 3. Class boundary of binary classification and novelty detection(o : normal data, x: novelty data, reprinted from Lee, 2007)

기법이다(Kang and Cho, 2008). 이범주 분류와 이상치 탐지의 차이는 <Figure 3>과 같다. 일반적인 이범주 분류(binary classification) 기법은 정상 데이터와 비정상 데이터가 충분하다는 가정 하에, 두 범주를 가장 잘 구분할 수 있는 분류 기준을 생성한다. 따라서 <Figure 3>(a)의 A와 B는 이범주 분류 알고리즘에 의해 대부분 정상으로 분류된다. 그러나 <Figure 3>의 경우 비정상 데이터가 정상 데이터에 비하여 매우 적기 때문에 분류 알고리즘이 비정상 데이터의 속성을 충분하게 학습하지 못하는 문제점이 발생한다. 반면에 이상치 탐지 기법은 정상 데이터만을 학습하여 분류 경계면을 생성하기 때문에 <Figure 3>(b)의 A와 B는 비정상 데이터로 분류된다.

이상치 탐지를 이용한 가상 계측 결과의 신뢰도 부여 및 설비 변수 영향력 측정은 다음과 같이 수행될 수 있다. 우선 가상 계측 모델 구축에 사용된 설비 데이터를 정상 데이터로 가정하고 이상치 탐지 모델을 구축한다. 새로운 웨이퍼가 가상 계측 모델에 투입되면, 해당 웨이퍼의 설비 데이터를 구축된 이상치 탐지 모델에 투입하여 가상 계측 모델 구축 당시의 설비 데이터들과의 유사도를 측정한다. 만일 새로운 웨이퍼의 설비 데이터가 가상 계측 모델 구축 웨이퍼들의 설비 데이터들과 유사할 경우에는 예측 모델의 정확도가 높을 것으로 판별하여 해당 웨이퍼에 대한 가상 계측 예측 결과에 대한 신뢰도를 높게 부여한다. 반면에, 새로운 웨이퍼의 설비 데이터가 가상 계측 모델 구축 당시의 설비 데이터들과 유사하지 않을 경우에는 예측 모델의 정확도를 보장할 수 없으므로 해당 웨이퍼에 대한 가상 계측 예측 결과에 대한 신뢰도를 낮게 부여한다. 또한 신뢰도가 낮게 부여된 경우에는 사용된 이상치 탐지 알고리즘을 통해 유사도가 낮게 추정된 원인이 되는 설비 변수를 탐지하고 그 영향력을 추정하여 그 결과를 공정 제어에 활용할 수 있게 된다.

본 연구에서는 가우시안 분포 추정(Gaussian density estimator; 이하 Gauss)과 K-평균 군집화(K-means clustering; 이하 KMC)의 두 가지 이상치 탐지 방법론을 사용하여 가상 계측 결과에 대한 신뢰도 및 설비 변수 영향력을 추정하였다. Gauss는 대표적인 분포 추정 기반 이상치 탐지 방법론으로써, 모든 정상 데이터가 식 (6)과 같이 동일한 가우시안 분포로부터 생성되었다고 가정하고 해당 분포의 평균(μ)과 공분산(Σ)을 추정하여

새로운 데이터에 대한 확률 밀도를 계산하여 정상 여부를 판별하는 기법이다(Barnett and Lewis, 1994).

$$p(\mathbf{x}) = (2\pi)^{-d/2} |\Sigma|^{-1/2} \times \exp\left[-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1}(\mathbf{x} - \mu)\right] \quad (6)$$

Gauss를 이상치 탐지 기법으로 사용하여 가상 계측 결과의 신뢰도를 추정할 경우, 각 설비 변수들의 영향력은 다음과 같이 추정할 수 있다.

$$j\text{번째 설비 변수}(x_j)\text{의 영향력} : \frac{w_j}{\sum_{i=1}^m w_i}, \quad w_i = \frac{(x_i - \bar{x}_i)^2}{\sigma_i^2} \quad (7)$$

여기서 \bar{x}_i 는 가상 계측 구축 웨이퍼들의 i 번째 설비 변수의 평균이고, σ_i^2 는 가상 계측 구축 웨이퍼들의 i 번째 설비 변수의 분산이다. 따라서 j 번째 설비 변수의 영향력은 해당 변수가 가상 계측 모델 구축 당시 변동성이 적고, 테스트 웨이퍼의 값이 가상 계측 구축 웨이퍼들의 평균값과 상이할수록 증가하게 된다.

KMC는 대표적인 군집화 기반 이상치 탐지 기법으로써, 정상 데이터(X)들은 식 (8)에 나타난 바와 같이 몇 개의 동질적인 그룹(cluster)으로 구분할 수 있다고 가정한다.

$$X = C_1 \cup C_2 \cup \dots \cup C_K, \quad C_i \cap C_j = \phi, \quad i \neq j \quad (8)$$

KMC는 정상 데이터를 가장 잘 표현할 수 있는 최적의 그룹을 식 (9)와 같이 그룹 내 제곱 오차(within-cluster sum of squared error)을 최소화하여 찾는다.

$$\arg \min_{\mathbf{c}} \sum_{i=1}^K \sum_{\mathbf{x}_j \in C_i} \|\mathbf{x}_j - \mathbf{c}_i\|^2 \quad (9)$$

여기서 \mathbf{c}_i 는 i 번째 군집(C_i)의 중심이다. 식 (9)를 통하여 최적 그룹의 수와 각 그룹의 중심(centroid)이 결정되면 새로운 데이터에 대하여 식 (10)과 같이 가장 가까운 그룹과의 유사도를 측정하여 정상 여부를 판별한다(Tax, 2001).

$$\|\mathbf{x}_j - \mathbf{c}_i\|^2, \quad \text{where } \|\mathbf{x}_j - \mathbf{c}_i\|^2 \leq \|\mathbf{x}_j - \mathbf{c}_k\|^2 \quad (10)$$

KMC는 Gauss와는 달리 모든 데이터가 단일 분포에서 생성되는 것을 가정하지 않기 때문에, <Figure 4>와 같이 데이터가 정규 분포의 형태를 나타내지 않는 경우에 매우 효과적으로 사용될 수 있다. KMC를 이상치 탐지 기법으로 사용하여 가상 계측 결과의 신뢰도를 추정하는 경우에는 각 설비 변수들의 영향력은 식 (11)과 같이 가장 가까운 군집과의 거리에 대한 각 설비 변수의 기여도로 추정할 수 있다.

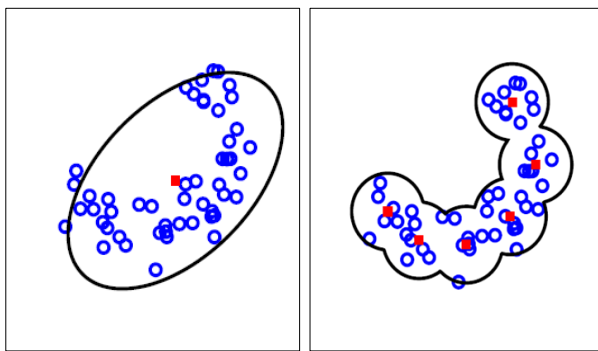
$$j\text{번째 설비 변수}(x_j)\text{의 영향력} : \frac{w_j}{\sum_{i=1}^m w_i}, \quad w_i = \frac{(x_i - \mathbf{c}_{ki})^2}{\|\mathbf{x} - \mathbf{c}_k\|^2} \quad (11)$$

KMC를 이상치 탐지 기법으로 사용할 경우, 새로운 웨이퍼와 가장 가까운 군집 중심과의 거리가 증가할수록 가상 계측 결과의 신뢰도는 낮게 추정된다. 따라서 군집 중심의 값과 차이가 크게 나타나는 설비 변수일수록 낮은 가상 계측 신뢰도에 대한 기여도가 증가하게 된다.

3. 실험 설계

3.1 데이터

본 연구에서 제안하는 이상치 탐지 기반 가상 계측 결과 신



(a) Gauss (b) KMC
Figure 4. Class boundary generated by Gauss and KMC (reprinted from Lee, 2007).

뢰도 및 설비 변수의 영향력 추정 모델의 효과를 검증하기 위하여 실제 반도체 제조 공정에서 수집된 데이터를 사용하여 실험을 수행하였다. 두 개의 photo-lithography 공정 설비(process equipment)에서 약 4개월 간에 걸쳐 117개의 공정 설비 변수와 4개의 주요 계측 변수(Y1, Y2, Y3, Y4)를 수집하였으며, 예방 보수(preventive maintenance) 시점을 기준으로 <Table 1>과 같이 총 8개의 구간으로 나누고, 각 구간에서 실제 계측 결과가 정상 영역(허용 오차 이내)에 속하는 초기 100개의 웨이퍼를 사용하여 가상 계측 및 이상치 탐지 모델을 구축하고, 이후 웨이퍼들을 대상으로 테스트를 수행하여 가상 계측 신뢰도의 효과를 측정하였다.

3.2 공정 설비 변수 선택

최초에 공정 설비로부터 수집되는 설비 변수는 총 117개로서, 모델 구축에 사용되는 웨이퍼의 수에 비해 매우 많은 편이다. 또한 상수 값을 나타내거나 서로 높은 상관관계를 나타내는 등 예측 모델 구축에 적합하지 않은 설비 변수들도 다수 포함되어 있기 때문에, 본 연구에서는 유전자 알고리즘(genetic algorithm; GA, Yang and Honavar, 1998)을 사용하여 가상 계측에 필수적인 설비 변수 집합을 선택하였다. 유전자 알고리즘을 이용한 변수 선택의 절차는 <Figure 5>와 같다. 먼저 일정 수의 초기 집단(population)을 생성하고 예측 알고리즘을 사용하여

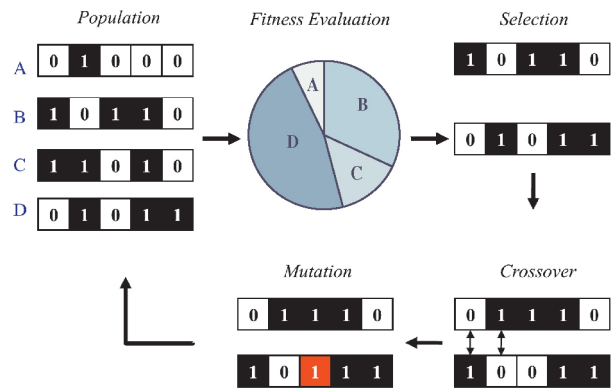


Figure 5. Variable selection by genetic algorithm(GA)(reprinted from Kang *et al.*, 2011)

Table 1. The number of wafers for two equipments in eight periods

장비	구간 1	구간 2	구간 3	구간 4	구간 5	구간 6	구간 7	구간 8
EQ1	230	172	137	167	452	818	138	195
EQ2	226	180	136	170	450	816	138	195

Table 2. The number of selected input variables for each equipment-period combination

장비	구간 1	구간 2	구간 3	구간 4	구간 5	구간 6	구간 7	구간 8
EQ1	12	9	8	12	22	26	24	19
EQ2	11	11	8	13	17	24	13	15

각 염색체들의 적합도(fitness)를 평가한다. 평가된 적합도를 바탕으로 부적합하거나 경쟁력이 없는 해를 제거하는 선택(selection), 좋은 염색체들의 결합을 가능하게 하는 교배(crossover), 낮은 확률로 임의의 값을 변화시키는 돌연변이(mutation) 등의 연산을 수행하여 새로운 세대의 집단을 생성한다. 유전자 알고리즘은 이러한 과정을 일정 세대 수 경과 또는 해의 변화 미비 등의 종료 조건을 만족할 때까지 수행하여 최적의 설비 변수 집합을 찾는 과정이며, 이를 통해 각 구간별로 선택된 공정 설비 변수의 수는 <Table 2>에 나타나 있다. 장비 및 구간별로 다소 차이는 있으나 유전자 알고리즘을 통하여 전체 설비 변수 대비 최소 78%(EQ1, 구간 6 : 26개), 최대 93%(EQ1, EQ2, 구간 3 : 8개)의 변수를 감소시킨 것을 확인할 수 있다.

3.3 알고리즘 매개변수, 신뢰도 추정 및 가상 계측 성능 평가 지표

본 연구에서는 MLR을 사용하여 가상 계측 모델을 구축하고 Gauss와 KMC의 이상치 탐지 알고리즘을 사용하여 가상 계측 결과에 대한 신뢰도를 부여하고 설비 변수의 영향력을 판별하였다. MLR을 이용하여 가상 계측 모델 구축 시, 계측 결과가 허용치를 벗어나는 불량 웨이퍼들은 모두 제거하고 계측 결과가 허용치 이내인 정상 웨이퍼들만을 이용하여 학습 데이터를 생성하였다. 또한, MLR과 Gauss는 매개 변수(parameter)가 없는 알고리즘이나 KMC의 경우에는 군집의 수(K)를 매개 변수로 하는 알고리즘이다. 따라서 KMC의 군집의 수는 가상 계측 모델 구축에 사용된 설비 데이터를 바탕으로 {2, 3, 5, 7, 10}개의 후보 군집 수를 사용하여 KMC를 실행한 후, 최적 군집 수를 결정하는 기준인 BIC(Bayesian information criterion, Schwarz, 1978)와 AIC(Akaike information criterion, Akaike, 1974) 사용하여 군집화의 적합도를 평가하였으며, 최종적으로 3개의 군집 수를 사용하였다.

새로운 웨이퍼의 설비 데이터가 이상치 탐지 모델에 투입되면 해당 웨이퍼에 대한 신뢰도는 다음과 같이 이범주(high/low)로 추정하였다. 먼저 가상 계측 모델 구축에 사용된 웨이퍼들을 사용하여 이상치 탐지 모델을 구축하고 학습 웨이퍼들에 대한 이상치 스코어를 산출하여 상위 5%(정상 데이터와 유사하지 않은 정도가 가장 큰 5%)에 해당하는 웨이퍼의 스코어를 기준(threshold)으로 설정한다. 새로운 웨이퍼의 이상치 스코어가

기준 이상의 값을 갖는 경우에는 가상 계측 결과의 신뢰도가 낮다고 추정하고, 기준 이하의 값을 갖는 경우에는 가상 계측 결과의 신뢰도가 높다고 추정한다. 즉 가상 계측 모델 구축에 사용된 웨이퍼들의 95% 이상과 유사하다고 판별된 새로운 웨이퍼에 대해서만 가상 계측 결과를 신뢰할 수 있다고 판단하고, 그렇지 않을 경우에는 가상 계측 모델 구축 시, 충분한 데이터의 속성을 학습하지 못한 것으로 간주하여 가상 계측 결과를 신뢰하기 어렵다고 판단하는 것이다.

추정된 신뢰도의 수준에 따라 다음과 같이 두 가지의 지표를 사용하여 가상 계측 모델의 예측 성능을 측정하였다. 첫 번째 지표는 평균 절대 오차(mean absolute error, 이하 MAE)로써, 식 (12)와 같이 실제 계측 값과 가상 계측에 의해 예측된 값 사이의 절대 오차를 측정한다.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (12)$$

여기서 n 은 전체 웨이퍼의 수이고 y_i 는 i 번째 웨이퍼의 실제 계측 값이며, \hat{y}_i 는 i 번째 웨이퍼의 가상 계측 결과이다. 두 번째 지표는 허용 오차 범위 이내 비율(percentage of absolute error, 이하 PARE)로써, 식 (13)과 같이 실제 계측 값과 가상 계측에 의해 예측된 계측 값 사이의 절대 오차가 일정 수준 이내인 웨이퍼의 비율이다.

$$PARE = \frac{1}{n} \sum_{i=1}^n I(|y_i - \hat{y}_i| < \alpha) \quad (12)$$

여기서 $I(\cdot)$ 는 지시 함수로써 괄호 안의 조건을 만족하면 1을 반환하고 그렇지 않을 경우 0을 반환한다. α 는 공정 관리의 수준에 따라 결정되는 가상 계측에 대한 허용 오차로써, 본 연구에서는 실제 공정에서 사용하는 값인 0.003을 이용하여 PARE를 측정하였다.

4. 실험 결과

<Table 3>은 두 이상치 탐지 알고리즘(Gauss, KMC)에 의해 가상 계측 결과의 신뢰도가 높게 부여된 웨이퍼들의 비율을 각

Table 3. The proportion of the test wafers with high reliability level assigned by two novelty detection models for each equipment-period combination

장비	이상치 탐지	구간								Summary	
		1	2	3	4	5	6	7	8	평균	편차
EQ1	Gauss	0.708	0.819	0.973	1.000	0.963	0.981	0.895	1.000	0.917	0.098
	KMC	0.785	0.778	0.973	1.000	0.946	0.985	0.921	0.990	0.922	0.084
EQ2	Gauss	0.667	0.950	0.944	1.000	0.991	0.986	0.895	0.958	0.924	0.102
	KMC	0.770	0.913	0.917	0.971	0.974	0.992	0.974	0.990	0.938	0.069

구간별로 나타낸 것이다. 구간에 따라 편차는 있으나 평균적으로 91~94%에 해당하는 테스트 웨이퍼의 설비 데이터가 Gauss와 KMC에 의해 가상 계측 구축에 사용된 웨이퍼의 설비 웨이퍼와 유사한 것으로 판별되어 높은 신뢰도를 부여받은 것을 알 수 있다. 또한 이상치 탐지 알고리즘을 비교하면 KMC가 Gauss에 비하여 다소 높은 비율로 새로운 웨이퍼에 대해 높은 신뢰도를 부여하는 것을 확인할 수 있으며(EQ1 : Gauss-0.917, KMC-0.922, EQ2 : Gauss-0.924, KMC-0.938), 구간별 편차 또한 적게 나타나는 것을 알 수 있다(EQ1 : Gauss-0.098, KMC-0.084, EQ2 : Gauss-0.102, KMC-0.069).

<Figure 6>은 두 이상치 탐지 알고리즘에 의해 부여된 신뢰도에 따른 가상 계측 결과의 구간별 MAE의 평균을 나타낸 것이다. <Figure 6>(a)~<Figure 6>(c)의 4번 구간과 <Figure 6>(a)의 8번 구간에는 신뢰도가 낮게 부여된 웨이퍼의 비율이 0이기 때문에 MAE가 측정되지 않았다. <Figure 6>으로부터 우리는 이상치 탐지 기법을 이용한 가상 계측 결과에 대한 신뢰도 부여를 통해 가상 계측 오차가 작은 웨이퍼들과 오차가 상대적으로 큰 웨이퍼들을 성공적으로 구분할 수 있음을 확인할 수 있다. 두 설비(EQ1, EQ2) 모두 대부분의 구간에서 높은 신뢰도를 부여받은 웨이퍼들의 평균 MAE가 낮은 신뢰도를 부여받은 웨이퍼들의 평균 MAE에 비하여 낮게 나타난다. 구간에 따라 다소 차이는 있으나, Gauss를 사용하여 가상 계측 결과의 신뢰도를 추정할 경우 신뢰도가 높게 부여된 웨이퍼들은 두 구간(EQ2의 구간 7, 구간 8)을 제외하고 최소 7%(EQ1의 구간 1), 최대 59%(EQ1의 구간 5) 낮은 MAE를 나타내고 있다. 또한 KMC를 사용하여 가상 계측 결과의 신뢰도를 추정할 경

우에는 신뢰도가 높게 부여된 웨이퍼들은 한 구간(EQ2의 구간 7)을 제외하고 최소 4%(EQ2의 구간 2), 최대 65%(EQ2의 구간 6) 낮은 MAE를 나타낸다.

이상치 탐지 기법을 활용하여 기반 가상 계측 결과에 신뢰도를 부여하는 방법의 효과는 <Figure 7>의 가상 계측 모델의 구간별 평균 PARE에서도 확인할 수 있다. 두 설비 모두 대부분의 구간에서 높은 신뢰도를 부여받은 웨이퍼들의 PARE가 낮은 신뢰도를 부여받은 웨이퍼들의 PARE에 비하여 높게 나타난다. Gauss를 이상치 탐지 기법으로 사용한 경우, EQ2의 구간 2와 구간 7을 제외하고 높은 신뢰도를 부여받은 웨이퍼들은 낮은 신뢰도를 부여받은 웨이퍼들에 비해 최소 3%(EQ2의 구간 8), 최대 21%(EQ2의 구간 5) 높은 PARE를 나타낸다. 한편, KMC를 이상치 탐지 기법으로 사용한 경우, EQ2의 구간 2, 구간 7, 구간 8을 제외하고 최소 7%(EQ1의 구간 7), 최대 42%(EQ2의 구간 6) 높은 PARE를 나타내고 있다. 이 외 주목할 사항은 MAE와 PARE의 측면에서 볼 때, 테스트 웨이퍼의 숫자가 증가할수록 가상 계측 신뢰도의 효과가 높게 나타나는 점이다. 제 3.1절에서 언급한 바와 같이 본 연구에서는 모든 구간에 대하여 최소 100개의 웨이퍼를 이용하여 이상치 탐지 및 가상 계측 모델을 구축하고 이후 웨이퍼들을 사용하여 성능 평가를 수행하였다. 이 때, 웨이퍼의 수가 가장 많은 구간은 두 장비 모두 구간 5와 구간 6(<Table 1> 참조)이다. <Figure 6>과 <Figure 7>의 결과를 보면 해당 구간에서는 다른 구간보다 높은 신뢰도를 부여받은 웨이퍼들의 가상 계측 결과가 낮은 신뢰도를 부여받은 웨이퍼들에 비하여 매우 우수하게 나타나는 것을 확인할 수 있다.

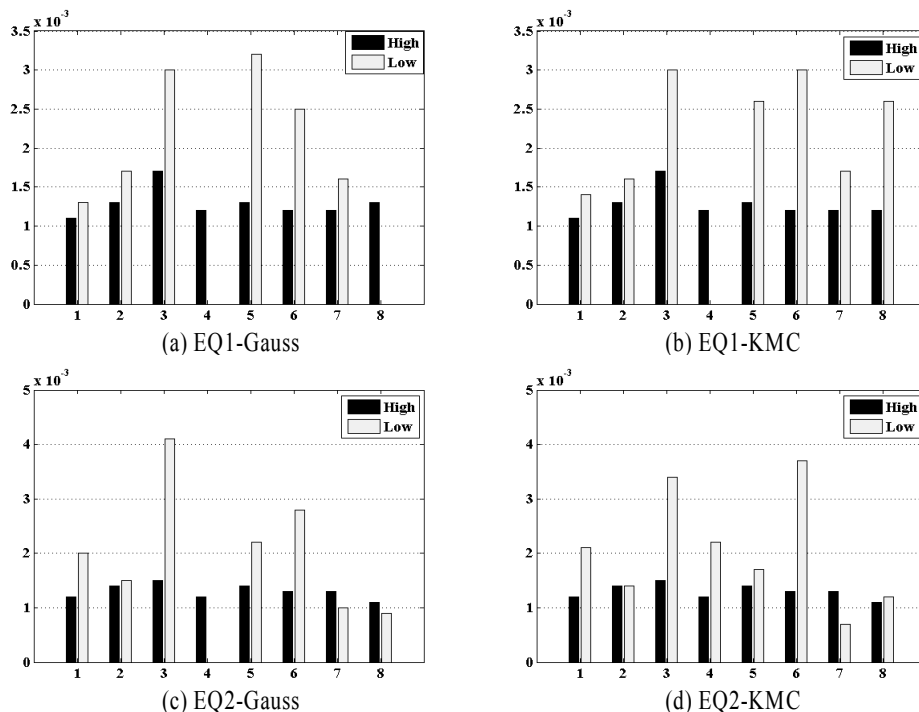


Figure 6. The average MAEs of the wafers with respect to the reliability level(x-axis : period, y-axis : MAE)

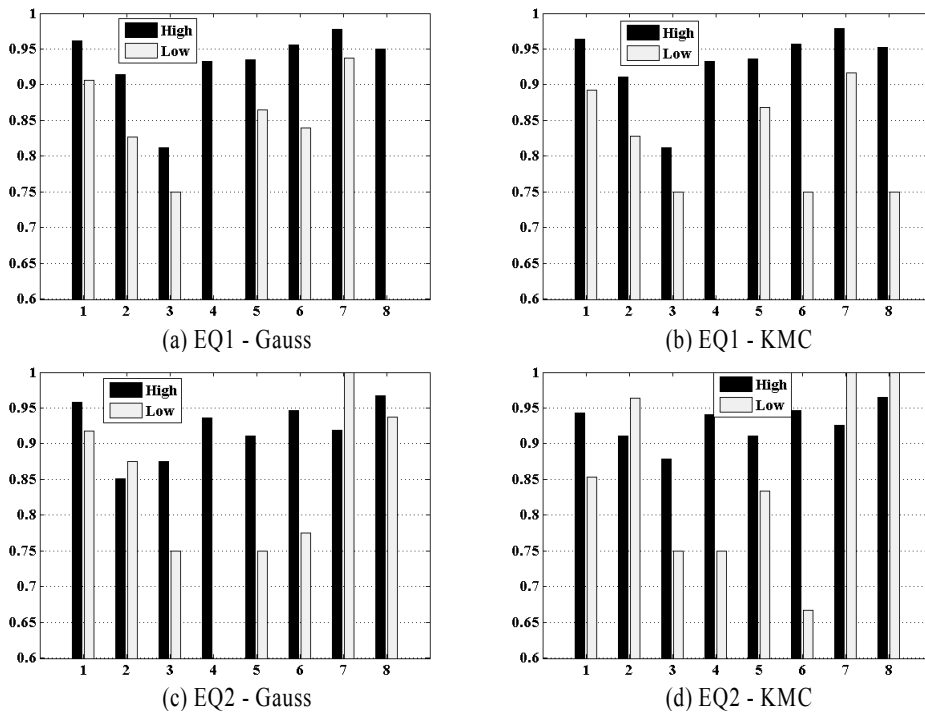
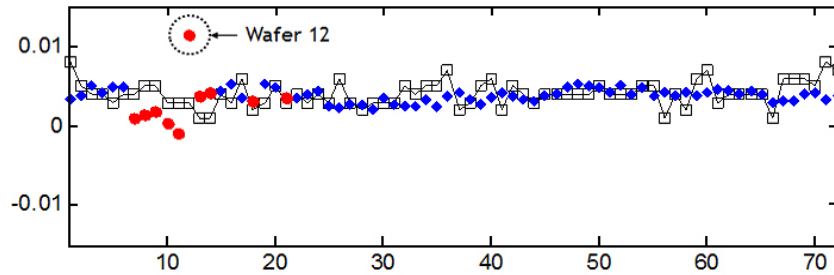
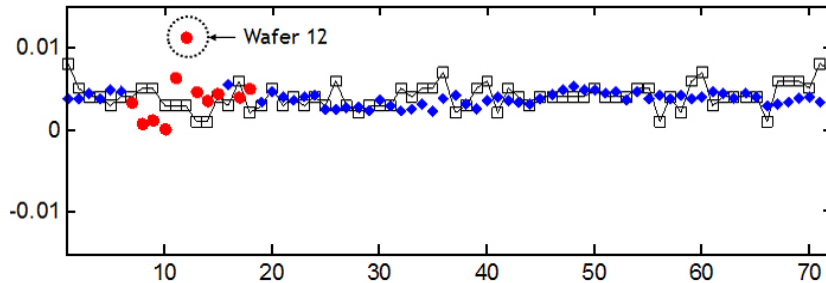


Figure 7. The average PAREs of the wafers with respect to the reliability level(x-axis : period, y-axis : PARE)



(a) Gauss



(b) KMC

Figure 8. The actual and predicted metrological values of Y2 in period 2 for EQ1(□ : actual metrological value, ◆ : virtual metrology with high reliability, ● : virtual metrology with low reliability, x-axis : wafer number, y-axis : normalized metrological value)

Table 4. The contribution of each process sensor parameter for a low reliability of 12-th wafer in period 2 for EQ1(%)

기법	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉
Gauss	0.53	0.36	0.01	0.09	0.18	<u>98.72</u>	0.01	0.04	0.07
KMC	1.11	0.74	0.08	0.06	0.44	<u>97.24</u>	0.02	0.30	0.02

<Figure 8>은 EQ1의 구간 2, 계측 변수 Y2에 대한 실제 계측 결과와 신뢰도에 따른 가상 계측 결과를 도시한 것이다. 두 이상치 탐지 기법에 의해 높은 신뢰도를 부여받은 웨이퍼들의 가상 계측 결과는 실제 계측 결과와 큰 차이를 보이지 않는 것을 볼 수 있으며, 낮은 신뢰도를 부여 받은 웨이퍼들의 가상 계측 결과는 일반적으로 높은 신뢰도를 부여받은 웨이퍼에 비하여 실제계측 정보와의 차이가 크게 나타나는 것을 알 수 있다. 특히, <Figure 8>(a), <Figure 8>(b)의 12번 웨이퍼의 경우에는 다른 웨이퍼들에 비해 특히 가상 계측 결과와 실제 계측 결과와의 차이가 매우 크게 나타나는 것을 확인할 수 있다.

가상 계측 결과의 신뢰도가 낮게 추정된 경우, 제 2.2절에서 설명한 바와 같이 낮은 신뢰도에 대한 각 설비 변수의 영향력을 추정할 수 있다. 예를 들어 <Figure 8>(a), <Figure 8>(b)의 12번 웨이퍼의 경우, 해당 구간에서 선택된 설비 변수의 수는 <Table 2>에서 나타난 바와 같이 총 9개이다. 이를 편의상 X_1, X_2, \dots, X_9 이라고 지칭하면, 각 설비 변수 X_i 의 Gauss와 KMC 결과에 대한 영향력은 각각 식 (7)과 식 (11)에 의해 구할 수 있으며, 그 결과는 <Table 4>에 나타난 바와 같다. <Table 4>를 통해서 우리는 9개의 공정 설비 변수 중에서 여섯 번째 설비 변수(X_6)에 의해 12번 웨이퍼의 가상 계측 신뢰도가 낮게 추정되었음을 알 수 있다. Gauss와 KMC 모두 X_6 이 낮은 신뢰도에 기여하는 비율이 97% 이상이며, 대부분의 나머지 설비 변수들은 영향력이 1% 이하로 나타난다. 이는 X_6 을 제외한 나머지 설비 변수들의 측정 수치는 가상 계측 모델 구축 시 사용된 웨이퍼들의 측정 수치와 크게 차이가 나지 않으나, X_6 의 수치가 일반적인 웨이퍼들과는 매우 큰 차이가 있음을 의미한다.

5. 결론 및 향후 연구 방향

본 연구에서는 가상 계측 결과의 활용성을 높이기 위하여 이상치 탐지 방법론을 사용한 가상 계측 결과의 신뢰도 및 설비 변수의 영향력 추정 방법론을 제안하고 실제 반도체 공정 데이터를 이용하여 그 효과를 검증하였다. 두 가지의 이상치 탐지 알고리즘을 사용하여 가상 계측 결과의 신뢰도를 추정한 결과, 높은 신뢰도가 부여된 웨이퍼들의 가상 계측 결과가 낮은 신뢰도가 부여된 웨이퍼들의 가상 계측 결과에 비하여 MAE 측면에서 최대 65%, PARE 측면에서 최대 42% 향상된 것으로 나타났으며, 개별 웨이퍼의 혐의 설비 변수 및 영향력 추정을 통해 낮은 신뢰도에 대한 원인 규명이 가능하다는 것도 입증하였다. 이러한 결과를 바탕으로 이상치 탐지 알고리즘을 활용함으로써 가상 계측 결과를 선택적으로 공정 관리에 적용하는 것이 가능하게 됨을 알 수 있다.

본 연구는 이상치 탐지 방법론을 이용하여 가상 계측 결과의 신뢰도를 추정하고자 하는 초기 연구로서 의의를 갖으며, 본 연구 결과를 통해 추후 진행되어야 할 연구 방향은 다음과 같다. 첫째, 본 연구에서는 가장 기본적인 예측 및 이상치 탐지 알

고리즘인 MLR과 Gauss, KMC를 이용하여 가상 계측 모델을 구축하고 가상 계측 결과에 대한 신뢰도를 부여하였다. 최근 다양한 기계 학습 및 데이터마이닝 예측 알고리즘이 가상 계측 모델 구축 이용되고 있으며, 다수의 이상치 탐지 알고리즘 또한 여러 응용 분야에서 성공적으로 활용되고 있다. 따라서 본 연구에서 사용된 방법론 이외에 보다 많은 예측 및 이상치 탐지 알고리즘을 사용하여 가상 계측 결과에 대한 신뢰도 부여 방법론의 효과를 비교, 분석하는 연구가 필요하다. 둘째, 가상 계측의 최종적인 목표는 개별 웨이퍼에 대한 미세 공정 관리를 가능하게 하는 R2R 시스템의 구축이다. 따라서 부여된 신뢰도에 따라 가상 계측 결과를 R2R 관리에 적용하는 방법론을 연구하고 이의 효과를 분석하는 연구 또한 진행되어야 할 것이다.

참고문헌

Akaike, H. (1974), A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, **19**(6), 716-723.

Barnett, V. and Lewis, T. (1994), *Outliers in Statistical Data*, Wiley and Sons, USA.

Bergeret, F. and Gall, C. L. (2003), Yield Improvement Using Statistical Analysis of Process Data, *IEEE Transactions on Semiconductor Manufacturing*, **16**(3), 535-542.

Besnard, J. and Toprac, A. (2006), Wafer-to-Wafer Virtual Metrology Applied to Run-to-Run Control, *Proc. of the Third ISMI Symposium on Manufacturing Effectiveness*.

Chang, Y.-J., Kang, Y., Hsu, C.-L., Chang, C.-T., and Chan, T. Y. (2006), Virtual Metrology Technique for Semiconductor Manufacturing, *Proc. Int. Joint Conf. on Neural Network(IJCNN 2006)*, 5289-5293.

Chen, P., Wu, S., Lin, J., Ko, F., Lo, H., and Wang, J. (2005), Virtual Metrology : A Solution for Wafer to Wafer Advanced Process Control, *Proc. IEEE Int. Symp. on Semiconductor Manufacturing (ISSM 2005)*, 155-157.

Chen, Y.-T., Yang, H.-C., and Cheng, F.-T. (2006), Multivariate Simulation Assessment for Virtual Metrology, *Proc. IEEE Int. Conf. on Robotics and Automation(ICRA 2006)*, 1048-1053.

Kang, P. and Cho, S. (2008), A Hybrid Novelty Score and Its Use in Keystroke Dynamics-based User Authentication, *Pattern Recognition*, **42**(11), 3115-3127.

Kang, P., Kim, D., Lee, H.-J., Doh, S., and Cho, S. (2011), Virtual Metrology for Run-to-Run Control in Semiconductor Manufacturing, *Expert Systems with Applications*, **38**(3), 2508-2522.

Khan, A. A., Moyné, J. R., and Tilbury, D. M. (2007), An Approach for Factory-wide Control Utilizing Virtual Metrology, *IEEE Transactions on Semiconductor Manufacturing*, **20**(4), 364-375.

Lee, H.-J. (2007), *Novelty Detection for Class Imbalance : LVQ-based Algorithm and Its Application to Security and CRM*, Ph.D Thesis, Seoul National University, Republic of Korea.

Lin, T.-H., Hung, M.-T., Lin, R.-C., and Cheng, F.-T. (2006), A Virtual Metrology Scheme for Predicting CVD Thickness in Semiconductor Manufacturing, *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA 2006)*, 1054-1059.

Lynn, S., Ringwood, J., and MacGearailt, N. (2010), Weighted Windowed

- PLS Models for Virtual Metrology of an Industrial Plasma Etch Process, *Proc. IEEE Int. Conf. on Industrial Technology(ICIT 2010)*, 309-314.
- Pan, T.-H., sheng, B.-Q., Wong, D. S.-H., and Jang, S.-S. (2011), A Virtual Metrology System for Predicting End-of-Line Electrical Properties using a MANCOVA Model with Tools Clustering, *IEEE Transactions on Industrial Informatics*, **7**(2), 187-195.
- Qin, S. J., Cherry, G., Good, R., Wang, J., and Harrison, C. A. (2006), Semiconductor Manufacturing Process Control and Monitoring : A Fab-wide Framework, *Journal of Process Control*, **16**(3), 179-191.
- Sachs, E., Hu, A., and Ingolfsson, A. (1995), Run by Run Process Control : Combining SPC and Feedback Control, *IEEE Transactions on Semiconductor Manufacturing*, **8**(1), 26-43.
- Schwarz, G. E. (1978), Estimating the Dimension of a Model, *Annals of Statistics*, **6**(2), 461-464.
- Spanos, C. J., Guo, H.-F., Miller, A., and Levine-Parril, J. (1992), Real-Time Statistical Process Control Using Tool Data, *IEEE Transactions on Semiconductor Manufacturing*, **5**(4), 308-318.
- Su, A.-J., Jeng, J.-C., Huang, H.-P., Yu, C.-C., Hung, S.-Y., and Chao, C.-K. (2007), Control Relevant Issues in Semiconductor Manufacturing : Overview with Some New Results, *Control Engineering Practice*, **15**(10), 1268-1279.
- Tax, D. M. J. (2001), *One-class Classification*, Ph.D. Thesis, Delft University of Technology, Netherlands.
- Yang, J. and Honavar, V. (1998), Feature Subset Selection Using a Genetic Algorithm, *IEEE Intelligent Systems*, **13**(2), 44-49.