

# Markov Chain Monte Carlo를 이용한 반도체 결함 클러스터링 파라미터의 추정

하정훈<sup>†</sup> · 장준현 · 김준현

홍익대학교 정보컴퓨터공학부

## Estimation of Defect Clustering Parameter Using Markov Chain Monte Carlo

Chunghun Ha<sup>†</sup> · Jun Hyun Chang · Joon Hyun Kim

School of Information and Computer Engineering, Hongik University, Seoul, Korea

Negative binomial yield model for semiconductor manufacturing consists of two parameters which are the average number of defects per die and the clustering parameter. Estimating the clustering parameter is quite complex because the parameter has not clear closed form. In this paper, a Bayesian approach using Markov Chain Monte Carlo is proposed to estimate the clustering parameter. To find an appropriate estimation method for the clustering parameter, two typical estimators, the method of moments estimator and the maximum likelihood estimator, and the proposed Bayesian estimator are compared with respect to the mean absolute deviation between the real yield and the estimated yield. Experimental results show that both the proposed Bayesian estimator and the maximum likelihood estimator have excellent performance and the choice of method depends on the purpose of use.

**Keywords** : Yield Model, Clustering Effect, Parameter Estimation, Markov Chain Monte Carlo, Bayesian Estimation

### 1. 서 론

다이수율(die yield)은 웨이퍼에서 제조가 가능한 총 다이(die)의 개수와 양품인 다이 개수의 비율이다. 수율은 반도체의 제조공정에 따라 라인수율, 다이수율, 조립수율, 검사수율, 품질수율 등으로 분류할 수 있으며, 총 누적 수율은 이들의 곱으로 표현된다. 총 누적수율에 가장 많은 영향을 미치는 것은 다이수율 이므로 다이수율을 관리하는 것은 총 누적수율관리의 핵심이다[4]. 다이 수준에서의 불량은 총 발생 불량량의 약 50%를 차지하고 있으며, 이의 주원인은 오

염으로 알려져 있다[10].

다이에서 발생하는 불량은 대부분이 결함(defects) 또는 결점(faults)에 기인하므로 다이수율은 결점에 기초한 Poisson 분포를 이용한 Poisson 수율모델로 예측이 가능하다[8, 15]. 그러나 Poisson 수율모델은 다이의 면적이 작은 경우(또는 웨이퍼 상에 결함이 고르게 분포되어 있는 경우)에는 예측력이 우수하지만, 다이의 면적이 큰 경우(또는 웨이퍼 상에 결점이 군집되어 있는 경우)에는 예측수율이 실제수율보다 작은 현상이 발생한다. 이러한 현상을 클러스터링 효과(clustering effect)라 부른다. Poisson 수율모델에서는 다이 당(또는 단위면적 당) 평균 결점수를 상수

논문접수일 : 2009년 06월 12일    논문수정일 : 2009년 08월 18일    게재확정일 : 2009년 08월 21일

<sup>†</sup> 교신저자 chunghun.ha@hongik.ac.kr

※ 이 논문은 2007년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음(KRF-2007-331-D00545).

로 가정하므로 하나의 웨이퍼 상에 있는 모든 다이는 동일한 결점 발생 가능성을 가진다. 다이의 면적이 작으면 하나의 다이에 다수의 결점이 존재할 확률이 낮으므로 전체적인 웨이퍼의 관점에서는 결점이 균등하게 분포되어 있는 것처럼 보인다. 그러나 다이의 면적이 증가하면 일부 다이는 다수의 결점을 포함하고 일부 다이는 결점을 포함하지 않게 된다. 이 때, 다이 당 결점수의 평균과 분산이 모두 변화하며 평균은 웨이퍼 당 다이의 개수, 분산은 결점의 군집정도에 따라 변화한다. 하나의 다이에 한 개 이상의 결점은 해당 다이의 불량이라는 동일한 결과를 초래하므로 실제수율이 예측수율보다 커진다.

Poisson 수율모델에서 클러스터링 효과를 반영하기 위해, 다이 당 평균 결점수를 상수가 아닌 Gamma 분포를 따르는 확률변수로 가정하는 모델이 음이항(negative binomial, 이하 NB) 수율모델이다. Poisson 수율모델에서는 파라미터가 다이 당 평균 결점수 하나이므로 평균과 분산의 변화를 모두 반영할 수 없으나, NB 수율모델은 파라미터가 다이 당 평균 결점 수와 클러스터링 파라미터(clustering parameter 또는 over-dispersed parameter)로 두 개를 가지고 있으므로 클러스터링 효과를 반영할 수 있다.

수율모델로서 NB 모델을 이용할 경우는 두 개의 파라미터를 동시에 추정해야 한다. 다이 당 평균 결점수는 FAB공정에서 웨이퍼 당 결함의 개수를 inline 또는 offline으로 모니터링 하고 defect size distribution과 critical area를 이용하여 검출이 가능하다[8, 11]. 클러스터링 파라미터의 추정은 용이하지 않은 데, 그 이유는 클러스터링 파라미터는 용어의 명확성과 달리 그 물리적 의미가 명확하지 않아 직접적으로 측정하기 어렵기 때문이다. 클러스터링 파라미터는 군집의 정도를 나타내는 파라미터이나 Gamma 분포의 파라미터와 Poisson 분포의 파라미터가 혼합된 파라미터이므로 결점의 군집 정도와 비례하여 나타나지 않을 뿐만 아니라 군집의 정도를 측정하는 기준 또한 존재하지 않는다. 따라서 클러스터링 파라미터는 별도의 추정과정을 수행해야 한다.

통계적인 과정을 통하여 파라미터를 추정하는 전통적인 방법으로는 적률추정법(method of moment estimation: 이하 MME)과 최우추정법(maximum likelihood estimation: 이하 MLE)이 있다. Cunningham[4]은 클러스터링 파라미터의 추정방법으로서 MME에 기반한 방법을 제시하였다. MME 클러스터링 파라미터 추정치를 이용한 NB 모델 예측수율은 Poisson 모델의 예측수율보다는 실제수율에 근사한 결과를 보였다. 그러나 실제수율과는 다소 차이가 있었으며, 임의의 클러스터링 파라미터 값을 적용한 경우 오히려 실제수율에 근사한 결과를 보이는 경우도 존재하였다. 또한 음수의 추정치를 발생하여 적용이 용이

하지 않은 경우가 발생하였다. MLE는 유연하며 불변성이 있어 일반적으로 가장 선호하는 파라미터 추정법이나, 목적하는 파라미터의 분포함수가 복잡한 경우 추정치를 계산하기 어려운 단점이 있다. NB 수율모델의 클러스터링 파라미터를 MLE를 이용하여 추정할 경우에도 비선형 최적화를 통하여 별도의 해를 구해야 한다.

파라미터 추정의 다른 대안으로는 최근 많은 주목을 받고 있는 베이지안 추론방법이 있다. 베이지안 추론은 과거의 경험 또는 직관에 의하여 사전분포를 정의하고 추정하고자 하는 파라미터에 대한 사후분포를 구하여 사후평균(posterior mean), 사후최빈값(posterior mode), 또는 사후중앙값(posterior median)을 추정량으로 하는 방법이다. 사후최빈값은 사후분포를 최대화하는 값이므로 만약 사전분포가 평평하다면 MLE와 동일하게 되며, 사후최빈값을 추정량으로 사용할 경우 일반화 최대우도 추정량(generalized maximum likelihood estimator)이라고도 한다. 베이지안 추론방법은 적은 데이터로도 좋은 추정치를 도출할 수 있으며 지속적으로 업데이트가 가능한 장점을 가지고 있으나, 사후분포가 표준분포가 아니면 이중 적분이 있는 주변밀도함수를 구해야 하므로 계산이 매우 어려운 단점이 있다.

Markov Chain Monte Carlo(MCMC)는 임의의 상태에서 시작한 마코프 체인은 충분한 기간이 지난 후에 유일한 정상분포를 가지며 이를 유지한다는 마코프 체인의 특성을 이용하여 정상분포 상태에서 샘플을 추출한 후 이를 몬테카를로 적분으로 계산하는 방법이다[1]. MCMC를 이용하면 베이지안 추정을 적용하는 데 필요한 이중적분이 있는 주변밀도함수를 용이하게 계산할 수 있다.

기존의 연구는 대부분 결점을 이용한 수율 모델에 집중하고 있으나[12, 13, 15, 16], 수율모델을 실제 공정에 적용하기 위해서는 파라미터 추정이 필수적이다. NB 수율모델의 경우 사후분포가 표준분포가 아니므로 주변밀도함수를 별도로 계산해야 한다. 이에 본 논문에서는 MCMC 방법을 이용하여 사후분포를 도출하는 베이지안 추론방법에 기반한 클러스터링 파라미터 추정방법을 제시하고자 한다. 또한 예측수율 비교를 통하여 대표적인 파라미터 추정방법인 MME, MLE, 그리고 베이지안 추정법 중 클러스터링 파라미터를 추정하는 데 가장 적합한 추정법을 판별하고자 한다.

본 논문의 구성은 다음과 같다. 우선 제 2장에서는 다이 수율의 NB 수율모델링 과정을 설명하고 클러스터링 파라미터가 도출되는 과정을 설명한다. 제 3장에서는 클러스터링 파라미터를 추정하기 위한 세 가지 방법론에 대하여 설명한다. 제 4장에서는 본 논문에서 제시하는 MCMC를 이용한 베이지안 추론 방법에 대하여 설명하고

5장에서는 모의실험을 통하여 각 방법에 대하여 파라미터 추정을 실시하고 반도체에 적합한 클러스터링 파라미터 추정법에 관하여 논한다. 마지막으로 제 6장에서는 본 논문의 결과를 요약하고 향후 연구방향을 제시한다.

## 2. 다이 수율모델링

다이 수율은 웨이퍼 상 결점의 개수를 이용하므로 이산분포를 사용하여 모델링한다. 설계와 상이한 것을 결함(defects)이라고 하고 결함 중에서 불량을 야기하는 치명적인 결함을 결점(faults)라고 하는 데, 결함과 결점간의 관계는 수리적으로 결함크기분포(defect size distribution)와 critical area를 통하여 모델링 할 수 있다. 이에 대한 상세한 설명은 Ha[8]의 논문을 참고하기 바란다. 본 논문에서는 결점으로부터 수율을 예측하는 과정만 다루므로 모든 결함은 결점으로 간주하여 논의를 전개한다.

다이 당 발생한 결점의 개수를 확률변수  $F$ 로 정의하자. 결점의 발생이 Poisson 분포를 따르고, 다이 당 평균 결점 개수가 상수인  $\lambda$ 라고 가정할 때, 하나의 다이에서  $k$ 개의 결점이 발생할 확률은 다음과 같다.

$$f(k|\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

다이 수율은 다이에 한 개의 결점도 없을 확률이므로

$$Y_{FOR} = \Pr(F=0|\lambda) = e^{-\lambda}$$

로 계산할 수 있다. 이것을 Poisson 수율모델이라 한다. 클러스터링 효과가 있는 경우, 웨이퍼 상 동일한 결점의 개수에 대하여 결점을 포함하지 않는 다이의 개수와 다수의 결점을 포함하는 다이의 개수가 동시에 증가한다. 따라서 다이 당 결점의 개수에 대한 평균보다 분산이 커지게 되고 예측수율이 실제수율보다 작게 나타난다. Poisson 수율모델은 하나의 파라미터  $\lambda$ 만을 사용하므로, 평균과 분산이 다른 경우 이를 반영할 방법이 존재하지 않는다. 최근에는 Poisson 분포에서 무결점의 확률이 Poisson 분포보다 큰 경우를 반영하기 위해 Zero-Inflated Poisson 분포를 응용하는 연구가 진행되었다[2, 6].

반면에 compounded Poisson 수율모델은 두 개의 파라미터를 사용하므로 분산이 평균보다 큰 경우에 이를 반영하기 용이한 장점이 있다. compounded Poisson 수율모델을 고려하기 위해 평균 defect의 개수  $\lambda$ 가 상수가 아니고 확률변수라고 가정하자. 일반적으로 가장 많이 사용하는 Poisson-Gamma 모델(NB 수율모델)은  $\lambda$ 가 Gamma 분포 ( $G(\alpha, \beta)$ )를 따른

다고 가정한다. 이 때,  $\lambda$ 의 확률분포함수는 다음의 식과 같다.

$$f(\lambda|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\beta}, \quad \lambda \geq 0$$

여기서  $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ 이다. 이를 Poisson 모델에 적용하면, 확률밀도함수는

$$\begin{aligned} f(k|\alpha, \beta) &= \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!} f(\lambda|\alpha, \beta) d\lambda \\ &= \frac{\Gamma(\alpha+k)\beta^k}{\Gamma(k+1)\Gamma(\alpha)(1+\beta)^{\alpha+k}} \end{aligned}$$

가 되고, 이는  $NB(\alpha, 1/(1+\beta))$ 의 확률밀도함수와 동일하다. 다이 수율은 결점의 개수가 0일 확률이므로

$$Y_{NB} = f(0|\alpha, \beta) = (1+\beta)^{-\alpha}$$

가 된다. 이를 NB 수율모델이라 한다.

$NB(\alpha, 1/(1+\beta))$  분포의 평균은  $\alpha\beta$ 이고 분산은  $\alpha\beta(\beta+1)$ 이다. NB 수율모델의 평균은 평균 결점의 개수이므로  $\lambda = \alpha\beta$ 로 하고,  $\alpha$ 와  $\lambda$ 를 파라미터로 하는 확률밀도함수로 변형하면,

$$f(k|\lambda, \alpha) = \frac{\Gamma(\alpha+k)}{\Gamma(k+1)\Gamma(\alpha)} \left(\frac{\lambda}{\lambda+\alpha}\right)^k \left(\frac{\alpha}{\lambda+\alpha}\right)^\alpha$$

이고 이는  $NB(\alpha, \alpha/(\lambda+\alpha))$ 와 같다. 이때의 평균은  $\lambda$ 이고 분산은  $\lambda + \lambda^2/\alpha$ 이므로 클러스터링 효과와 같이 분산이 평균보다 큰 경우에도 적합하게 사용이 가능하다. 클러스터링 효과가 존재할 경우, 결점에 기초한 수율모델은 NB 수율모델이 Poisson 모델에 비해 적합한 모델임은 이미 다른 연구에 의해서 입증이 되었고 가장 일반적으로 사용되고 있다. NB 모델을 사용하기 위해서는 두 개의 파라미터, 즉 다이 당(또는 단위면적 당) 평균 결점수  $\lambda$ 와 클러스터링 파라미터  $\alpha$ 를 추정해야 한다.  $\lambda$ 는 공정에서 automatic optical inspection(AOI)설비를 이용하여 웨이퍼 당 결점의 개수를 측정함으로써 용이하게 도출이 가능하다. 그러나 클러스터링 파라미터  $\alpha$ 의 경우는 용이하지 않다.

## 3. 클러스터링 파라미터 $\alpha$ 의 추정

클러스터링 파라미터  $\alpha$ 는 다양한 방법으로 추정이 가

능하나 대표적인 것으로는 MME, MLE, 그리고 베이지안 추론방법이 있다.

### 3.1 MME

데이터  $x_1, x_2, \dots, x_n$ 을 다이 당 결점의 개수라 하고, 각각은 iid이며,  $NB(\alpha, \alpha/(\lambda + \alpha))$ 로부터의 관측 값이라고 하자.  $NB(\alpha, \alpha/(\lambda + \alpha))$ 의 평균과 분산은 각각  $\lambda$ 와  $\lambda + \lambda^2/\alpha$ 이므로, 다음과 같이 표현할 수 있다.

$$\begin{aligned} \lambda &= \bar{x} \\ \lambda + \lambda^2/\alpha &= s^2 \end{aligned}$$

여기서  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$  이고  $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ 이다. 위의 두 등식을  $\alpha$ 에 대하여 정리하면, 클러스터링 파라미터  $\alpha$ 의 MME를 다음과 같이 구할 수 있다.

$$\hat{\alpha} = \frac{(\bar{x})^2}{s^2 - \bar{x}} \tag{1}$$

파라미터  $\alpha$ 는  $G(\alpha, \beta)$  분포의 파라미터 특성에 의하여 양수이어야 하므로, MME  $\hat{\alpha}$  또한 양수이어야 한다. 이는  $s^2 - \bar{x}$ 가 양수임을 의미하므로 분산이 평균보다 클 경우에만 MME는 유효하다. 그러나 실제 데이터를 적용하면 분산이 평균보다 작은 경우가 존재하므로 Cunningham[4]은 이를 수정하여 다음의 추정량을 제시하였다.

$$\hat{\alpha} = a \operatorname{ve} \left[ \left( \frac{\bar{x}}{s} \right)^2 \right] \frac{1}{1 - \operatorname{ave}(\bar{x}/s^2)}$$

MME는 표본평균과 표본분산을 이용하여 계산에 의해서 도출이 가능하므로 값을 구하기 용이한 반면, 파라미터  $\alpha$ 의 추정치가 분산의 변화에 매우 민감하게 반응하는 단점이 있다.

### 3.2 MLE

데이터  $x_1, x_2, \dots, x_n$ 을 다이 당 결점의 개수라 하고, 각각은 iid이며,  $NB(\alpha, \alpha/(\lambda + \alpha))$ 를 따른다고 하자. 그러면 우도함수는 다음과 같다.

$$l(\mathbf{x}|\lambda, \alpha) = \left[ \prod_{i=1}^n \frac{\Gamma(\alpha + x_i)}{\Gamma(x_i + 1)\Gamma(\alpha)} \right] \left( \frac{\lambda}{\lambda + \alpha} \right)^{n\bar{x}} \left( \frac{\alpha}{\lambda + \alpha} \right)^{n\alpha}$$

여기서,  $\mathbf{x}$ 는 벡터  $(x_1, x_2, \dots, x_n)$ 를 의미하고  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ 이다. 양변에 자연로그를 취하여, 로그우도함수를 구하면

$$\begin{aligned} \ln l(\mathbf{x}|\lambda, \alpha) &= \sum_{i=1}^n \ln \frac{1}{\Gamma(x_i + 1)} + \sum_{i=1}^n \sum_{j=0}^{x_i-1} \ln \left( 1 + \frac{j}{\alpha} \right) \\ &+ n\bar{x} \ln \lambda + n(\bar{x} + \alpha) [\ln(\alpha) - \ln(\lambda + \alpha)] \end{aligned} \tag{2}$$

와 같이 표현할 수 있다.

파라미터  $\lambda$ 에 대한 최대우도추정량을 구하기 위해 로그우도함수를  $\lambda$ 에 대하여 미분하고 1계 조건을 구하면,

$$\frac{\partial}{\partial \lambda} \ln l(\mathbf{x}|\lambda, \alpha) = n\bar{x} \frac{1}{\lambda} - n(\bar{x} + \alpha) \frac{1}{\lambda + \alpha} = 0$$

가 된다. 따라서  $\lambda$ 의 MLE  $\hat{\lambda}$ 는  $\bar{x}$ , 즉 다이 당 평균 결점의 개수가 된다.

파라미터  $\alpha$ 에 대한 최대우도추정량을 구하기 위해 로그우도함수를  $\alpha$ 에 대하여 미분하면,

$$\begin{aligned} \frac{\partial}{\partial \alpha} \ln l(\mathbf{x}|\lambda, \alpha) &= \sum_{i=1}^n \sum_{j=0}^{x_i-1} \left( \frac{1}{\alpha + j} - \frac{1}{\alpha} \right) \\ &+ \frac{n\lambda(\alpha + \bar{x})}{\alpha(\alpha + \lambda)} + n[\ln(\alpha) - \ln(\lambda + \alpha)] = 0 \end{aligned} \tag{3}$$

가 되므로 MLE  $\hat{\alpha}$ 를 구하기 위해서는  $\hat{\lambda} = \bar{x}$ 를 대입하고, 식 (3)의 근을 구하면 된다. MLE는 일반적으로 정확한 파라미터를 추정한다. 그러나 NB 수율모델의 경우는 식 (3)이 비선형이므로 비선형최적화 방법을 별도로 적용해야 하는 단점이 있다.

### 3.3 베이지안 추론

Poisson 수율모델을 베이지안 추정방법으로 구하기 위해서 데이터  $x_1, x_2, \dots, x_n$ 을 다이 당 결점의 개수라 하고, 각각은 iid이며,  $Poisson(\lambda)$ 를 따른다고 하자. 이 때, 우도함수는 다음과 같다.

$$l(\mathbf{x}|\lambda) \propto \prod_{i=1}^n e^{-\lambda} \lambda^{x_i} = e^{-n\lambda} \lambda^{n\bar{x}}$$

$\lambda$ 에 대한 사전 분포로서  $G(\alpha, \beta)$  분포를 고려하면,

$$\pi(\lambda) \propto \lambda^{\alpha-1} e^{-\lambda/\beta}$$

이고, 사후분포는

$$f(\lambda|\mathbf{x}) \propto \lambda^{(n\bar{x}+\alpha-1)} e^{-\lambda(n+1/\beta)}$$

가 된다. 즉, 사후분포는  $G(n\bar{x}+\alpha, \beta/(n+1))$  분포가 되며,  $\lambda$ 의 베이지안 추정량은 사후평균을 사용하면  $\hat{\lambda} = \frac{n\bar{x}+\alpha}{n+1/\beta}$  이다. 사후최빈값을 사용하면,  $\hat{\lambda} = \frac{n\bar{x}+\alpha-1}{n+1/\beta}$  이 된다. 이 방법은 Poisson 모델을 근간으로 하므로 파라미터  $\lambda$ 에 대한 추정은 가능하나, 파라미터  $\alpha$ 에 대한 추정은 불가능하다.

NB 수율모델의 베이지안 추정을 위하여, 데이터  $x_1, x_2, \dots, x_n$ 을 다이 당 결점의 개수라 하고, 각각은 iid이며,  $NB(\alpha, \alpha/(\lambda+\alpha))$ 를 따른다고 하자. 이때, 우도함수는 식 (2)와 같다.  $\lambda$ 에 대한 사전분포로서  $G(a, b)$  분포를 고려하면,

$$\pi_\lambda(\lambda) \propto \lambda^{a-1} e^{-\lambda/b}$$

이고,  $\alpha$ 에 대한 사전분포로서  $\text{Exp}(\theta)$ 를 고려하면,

$$\pi_\alpha(\alpha) \propto \theta e^{-\theta\alpha}$$

가 된다.  $\alpha$ 와  $\lambda$ 는 독립이라고 가정하면, 결합사전분포  $\pi(\lambda, \alpha)$ 는  $\pi_\lambda(\lambda)\pi_\alpha(\alpha)$ 가 되고 사후분포는

$$f(\lambda, \alpha|\mathbf{x}) = \frac{l(\mathbf{x}|\lambda, \alpha)\pi(\lambda, \alpha)}{\int_0^\infty \int_0^\infty l(\mathbf{x}|\lambda, \alpha)\pi(\lambda, \alpha)d\lambda d\alpha}$$

가 되어 이중 적분이 있는 주변밀도함수를 구해야 한다.

#### 4. MCMC를 이용한 베이지안 추정

제 3장에서 적용한 베이지안 추정방법은 사후분포의 계산이 매우 어려워 실제로 적용하기 어렵다. 따라서 본 절에서는 MCMC, 특히 Gibbs Sampler를 이용한 파라미터  $\alpha$ 의 추정방법에 관하여 논한다.

##### 4.1 계층적 베이지스를 적용한 사후분포의 추출

베이지안 추론법을 NB 수율모델에 적용할 경우 사전 분포 간에 독립으로 가정해야 하며, 이중적분이 있는 주변밀도함수를 계산해야 하는 어려움이 따른다. 이를 해결할 수 있는 방법이 초모수(hyperparameter)를 이용한 계층적 베이지스(hierarchical Bayes) 방법이다[3].

NB 수율모델을 계층적 베이지스 방법으로 구하기 위해

서 데이터  $x_i$ 를  $i$ 번째 다이의 결점의 개수라 하고, Poisson ( $\lambda_i$ )를 따른다고 하자. 각각의 다이에서 결점의 발생이 독립적으로 이루어진다고 가정하면, 우도함수는 다음과 같다.

$$l(\mathbf{x}|\lambda) = \prod_{i=1}^n \frac{\lambda_i^{x_i} e^{-\lambda_i}}{x_i!}$$

여기서,  $\lambda = (\lambda_1, \dots, \lambda_n)$ 이다.  $\lambda_i$ 에 대한 사전 분포로서  $G(\alpha, \beta)$  분포를 고려하면,

$$f(\lambda_i|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda_i^{\alpha-1} e^{-\lambda_i/\beta}$$

이다. 여기서  $\alpha$ 와  $\zeta (=1/\beta)$ 에 대한 초사전분포(hyperprior distribution)로서 각각  $\text{Exp}(c)$ 와  $\text{Uniform}(a, b)$ 를 고려하면 다음과 같다.

$$f(\alpha|c) = c e^{-c\alpha}$$

$$f(\zeta|a, b) = \frac{1}{b-a}$$

따라서 파라미터  $\theta = (\lambda_1, \lambda_2, \dots, \lambda_n, \alpha, \zeta)$ 에 대한 사후분포는

$$f(\theta|\mathbf{x}) \propto l(\mathbf{x}|\lambda) f(\lambda|\alpha, \zeta) f(\alpha|c) f(\zeta|a, b) \quad (4)$$

$$\propto \left( \frac{1}{\Gamma(\alpha)} \right)^n (\zeta^\alpha)^n \left( \prod_{i=1}^n \lambda_i^{x_i+\alpha-1} \right) (e^{-n\lambda(1+\zeta-\alpha)})$$

가 된다. 여기서,  $\bar{\lambda} = \frac{1}{n} \sum_{i=1}^n \lambda_i$ 이다.

위에서 구한 계층적 베이지스를 이용한 사후분포는 다차원의 결합확률분포가 복잡하여 사후분포를 직접 계산하기 어려우므로 각 변수의 조건부 확률분포로부터 랜덤표본을 반복적으로 생성하여 정상상태에서 결합확률 밀도함수로 수렴하는 Gibbs Sampler를 이용하면 효과적이다.

##### 4.1 Gibbs Sampler를 이용한 사후분포의 계산

Gibbs Sampler[5]는 Metropolis-Hastings(MH) 알고리즘 [9]과 함께 정상분포를 갖는 마코프 체인을 형성하는 대표적인 방법이다. MH 알고리즘은 1차원 변수에 대한 마

코프 체인을 형성하는 데 적합하고 Gibbs Sampler는 다차원 변수에 대한 마코프 체인을 형성하는 데 적합하다.

Gibbs Sampler를 적용하기 위해서는 식 (4)의 사후분포로부터 각 파라미터에 대한 조건부 확률분포를 구해야 한다.  $\lambda_i$ 의 조건부 확률밀도함수를 구하면 다음과 같다.

$$f(\lambda_i | \tilde{\lambda}_i, \alpha, \zeta, \mathbf{x}) \propto \lambda_i^{x_i + \alpha - 1} e^{-\lambda_i(1 + \zeta)}$$

여기서,  $\tilde{\lambda}_i = (\lambda_1, \lambda_2, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_n)$ 이다. 따라서  $\lambda_i$ 는 모든  $i$ 에 대하여  $G(x_i + \alpha, 1/(1 + \zeta))$ 를 따른다.  $\zeta$ 의 조건부 확률밀도함수는

$$f(\zeta | \lambda, \alpha, \mathbf{x}) \propto \zeta^{\alpha n} e^{-\zeta n \bar{\lambda}}$$

이므로  $\zeta$ 는  $G(\alpha n + 1, 1/(n \bar{\lambda}))$ 를 따른다.  $\alpha$ 의 조건부 확률밀도함수는

$$f(\alpha | \lambda, \zeta, \mathbf{x}) \propto \left(\frac{1}{\Gamma(\alpha)}\right)^n (\zeta^{\alpha n}) \left(\prod_{i=1}^n \lambda_i^{\alpha - 1}\right) (e^{-\alpha})$$

가 되나 이는 표준 분포가 아니므로 샘플링하기 어렵다. 그러나 log-concave 함수이므로 Gilks and Wild가 제시한 adaptive rejection sampling(ARS)[7]을 이용하면 샘플링이 가능하다.

각 파라미터의 조건부 확률밀도함수를 이용한 Gibbs Sampler 알고리즘은 <그림 1>과 같이 구성할 수 있다.

### 5. 시뮬레이션

클러스터링 파라미터의 추정방법의 비교를 위해서는

```

Set  $\theta_0 = \theta^{(0)}$ ;
For  $i = 0 : (N-1)$  do
     $\lambda_1^{(i+1)} \sim f(\lambda_1 | \tilde{\lambda}_1, \alpha, \zeta, \mathbf{x})$ ;
     $\lambda_2^{(i+1)} \sim f(\lambda_2 | \tilde{\lambda}_2, \alpha, \zeta, \mathbf{x})$ ;
    :
     $\lambda_n^{(i+1)} \sim f(\lambda_n | \tilde{\lambda}_n, \alpha, \zeta, \mathbf{x})$ ;
     $\zeta^{(i+1)} \sim f(\zeta | \lambda, \alpha, \mathbf{x})$ ;
     $\alpha^{(i+1)} \sim f(\alpha | \lambda, \zeta, \mathbf{x})$ ;
end
    
```

<그림 1> Gibbs Sampler

기존의 방법에 의한 파라미터의 추정치와 MCMC를 이용한 베이저안 추정치의 정확도를 비교해야 한다. 정확도는 실제수율과 추정된 파라미터를 이용한 예측수율의 비교를 통하여 이루어 질 수 있다. 본 논문에서는 실제와 근사한 실험 모델을 구성하여 이를 검증하였다.

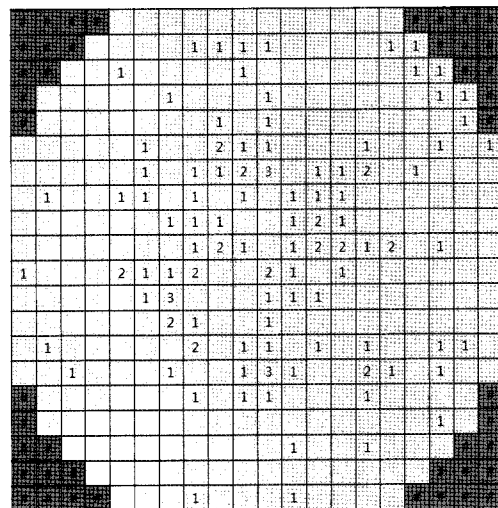
### 5.1 실험대상

실험 대상은 실제 데이터를 사용한 Cunningham[4]의 논문에서 제시된 웨이퍼를 확장하여 적용하였다. 실험 대상이 되는 웨이퍼는 4인치 웨이퍼로서 현재 일반적으로 적용하고 있는 8인치 또는 12인치 웨이퍼 보다 작으나 본 논문에서 적용하고자 하는 방법은 웨이퍼의 크기에는 관계없으므로 4인치 웨이퍼로 가정하고 진행하였다. 각 웨이퍼는 다이 면적에 따라 <표 1>과 같이 격자의 크기에 따라 다양한 면적의 다이를 생산할 수 있다. 여기서 다이의 개수는 격자 중에서 <그림 2>와 같이 외곽에 활용 불가능한 다이를 제거한 개수이다.

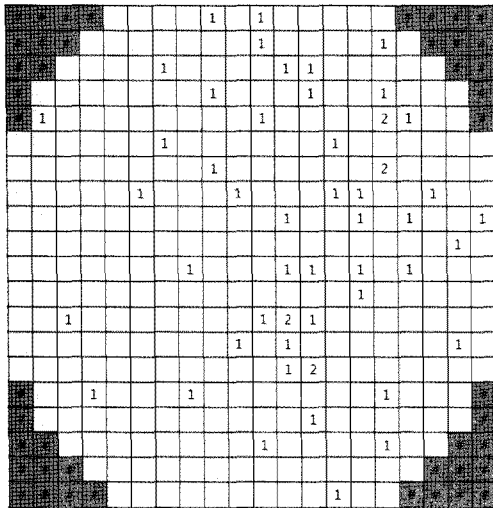
실험은 총 6개의 웨이퍼에 대하여 실시하였으며, 그 중 2개의 웨이퍼(Wafer1, Wafer2)는 Cunningham[4]에 제시된 데이터를 동일하게 사용하였고, 나머지 4개의 웨이퍼는 임의의 wafer map을 형성하여(wafer map은 다르나

<표 1> Number of dies per wafer and die area

격자	다이의 개수	다이의 면적( $cm^2$ )
40×40	1,536	0.0452
20×20	384	0.0181
10×10	96	0.72
8×8	58	1.19



<그림 2> Wafer map for Wafer1 (20×20)



<그림 3> Wafer map for *Wafer2* (20×20)

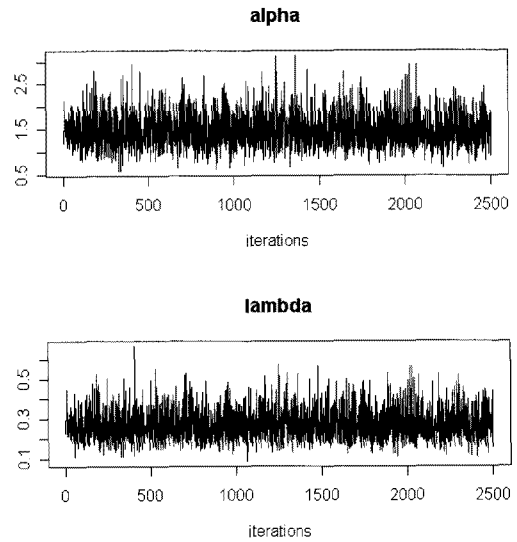
<표 2> Faults information for 2 wafers

faults per die	Wafer1 Grid size				Wafer2 Grid size			
	40	20	10	8	40	20	10	8
0	1419	287	41	18	1484	334	54	28
1	117	80	29	14	50	46	33	19
2	0	14	10	8	2	4	7	5
3	0	3	4	7	0	0	1	4
4	0	0	7	3	0	0	1	1
5	0	0	3	3	0	0	0	1
6	0	0	1	2	0	0	0	0
7	0	0	1	1	0	0	0	0
8	0	0	0	2	0	0	0	0
9	0	0	0	0	0	0	0	0
>10	0	0	0	0	0	0	0	0
합계	117	117	117	117	54	54	54	54

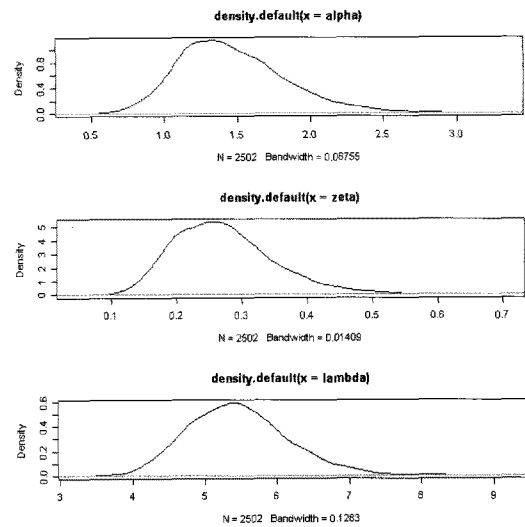
결점의 개수는 Cunningham[4]과 동일하다. 실험에 사용하였다. 대표적인 예로서 Wafer 1과 Wafer 2의 결점의 위치 및 결점에 대한 정보는 <그림 2, 3>와 <표 2>에 제시되어 있다. 각 웨이퍼는 각각 총 결점의 개수가 다르며 (Wafer 1은 117개, Wafer 2는 54개), 동일한 웨이퍼 일지라도 다이의 개수가 달라지면 격자의 위치도 변경되므로 다이 당 결점의 개수가 달라진다.

### 5.2 클러스터링 파라미터의 추정

클러스터링 파라미터  $\alpha$ 는 MME, MLE, 그리고 MCMC를 이용한 베이저안 방법을 이용하여 각각 추정하였다.



<그림 4> MCMC Samples of  $\alpha$  and  $\lambda$



<그림 5> Posterior distributions of  $\alpha$ ,  $\zeta$ , and  $\lambda$

MME와 MLE는 제 3장에서 제시한 방법으로 추정하였으며, MCMC를 이용한 베이저안 추정은 제 4장에서 제시한 방법을 적용하여 ARS가 기본적으로 내장된 WinBUGS(Bayesian interface using Gibbs sampling)[14]와 R2WinBUGS[17]를 이용하여 추정하였다. 베이저안 방법은 과거의 경험을 적용할 수 있는 장점이 있으므로 초사전분포의 파라미터  $c$ 는 Cunningham[4]에서 추정된  $\alpha$ 의 평균값을 이용하여 초기치를 설정하였다( $c=1/ave(\alpha)=0.771$ ). 초사전분포의 파라미터  $a$ 와  $b$ 는 별다른 사전정보가 없으므로 임의로  $a=0$ ,  $b=10$ 을 적용하였다. MCMC는 3개의 마코프 체인에 대하여 5,000개를 샘플을 생성한 후 2,500개를 번인으로 제거하고 <그림 4>와 같이 2,500개의 샘플을 사용하여 <그림 5>와 같이  $\alpha$ ,  $\zeta$ , 그리고  $\lambda$ 에 대한 사후

<표 3> Comparison results of various estimation methods with fabricated wafers

항 목		Wafer 1				Wafer 3				Average(6 wafers)			
		40 X 40	20 X 20	10 X 10	8 X 8	40 X 40	20 X 20	10 X 10	8 X 8	40 X 40	20 X 20	10 X 10	8 X 8
데이터	총 다이의 개수/wafer	1536	384	96	58	1536	384	96	58	1536.00	384.00	96.00	58.00
	양품 다이의 개수/wafer	1421	283	40	18	1487	336	66	34	1377.00	266.17	39.33	18.33
	D0(defects/wafer)	117	117	117	117	49	49	49	49	166.67	166.67	166.67	166.67
	A(cm <sup>2</sup> )	0.045	0.181	0.720	1.190	0.045	0.181	0.720	1.190	0.045	0.181	0.720	1.190
파라미터 추정치	fixed $\alpha$	0.700	0.700	0.700	0.700	0.700	0.700	0.700	0.700	0.700	0.700	0.700	0.700
	Cunningham $\alpha$	negative	41.658	1.324	1.332	negative	negative	0.784	0.822	negative	13.152	1.885	1.409
	MLE $\alpha$	100.00	54.180	1.275	1.145	100.00	100.00	0.671	0.650	100.00	44.170	1.880	1.137
	MLE $\lambda$	0.076	0.305	1.219	2.017	0.032	0.128	0.510	0.845	0.109	0.434	1.736	2.874
	MCMC $\alpha$ (mean)	0.695	2.281	1.611	1.397	0.272	1.052	1.077	0.961	1.000	2.041	1.626	1.408
	MCMC $\lambda$ (mean)	0.080	0.310	1.206	2.004	0.034	0.133	0.503	0.824	0.112	0.438	1.727	2.852
	MCMC $\alpha$ (mode)	0.703	2.253	1.366	1.215	0.261	1.013	0.662	0.685	1.013	2.000	1.378	1.145
	MCMC $\lambda$ (mode)	0.078	0.309	1.177	1.956	0.033	0.133	0.468	0.748	0.112	0.433	1.691	2.771
실제 수율		0.925	0.737	0.417	0.310	0.968	0.875	0.688	0.586	0.896	0.693	0.410	0.316
예측 수율	Poisson 모델	0.927	0.737	0.296	0.133	0.969	0.880	0.600	0.430	0.900	0.675	0.295	0.175
	Murphy 모델	0.927	0.743	0.334	0.185	0.969	0.881	0.613	0.456	0.901	0.686	0.329	0.212
	NB 모델 fixed $\alpha$	0.930	0.777	0.494	0.387	0.969	0.889	0.682	0.575	0.908	0.741	0.480	0.383
	NB 모델Cunningham $\alpha$	-	0.738	0.421	0.293	-	-	0.675	0.559	-	0.607	0.401	0.292
	NB 모델 MLE $\alpha, \lambda$	0.927	0.738	0.425	0.312	0.969	0.880	0.684	0.582	0.900	0.696	0.412	0.315
	NB 모델 MCMC $\alpha, \lambda$	0.930	0.751	0.404	0.287	0.970	0.887	0.658	0.545	0.904	0.701	0.399	0.294
	NB 모델 MCMC $\alpha, \lambda$ (mean)	0.927	0.748	0.406	0.289	0.968	0.882	0.662	0.552	0.901	0.698	0.401	0.297
	NB 모델 MCMC $\alpha, \lambda$ (mode)	0.929	0.749	0.428	0.312	0.969	0.883	0.702	0.603	0.902	0.701	0.422	0.326
MAD	Poisson 모델	0.002	0.000	0.121	0.177	0.001	0.005	0.087	0.157	0.003	0.022	0.115	0.141
	Murphy 모델	0.002	0.006	0.083	0.126	0.001	0.006	0.074	0.130	0.004	0.014	0.081	0.104
	NB 모델 fixed $\alpha$	0.005	0.040	0.077	0.077	0.001	0.014	0.006	0.012	0.012	0.048	0.072	0.071
	NB 모델Cunningham $\alpha$	-	0.001	0.005	0.018	-	-	0.013	0.027	-	0.005	0.014	0.027
	NB 모델 MLE $\alpha, \lambda$	0.002	0.001	0.008	0.002	0.001	0.005	0.003	0.004	0.003	0.004	0.007	0.015
	NB 모델 MCMC $\alpha, \lambda$	0.005	0.014	0.013	0.023	0.002	0.012	0.029	0.041	0.008	0.010	0.016	0.030
	NB 모델 MCMC $\alpha, \lambda$ (mean)	0.002	0.011	0.010	0.022	0.000	0.007	0.026	0.035	0.005	0.007	0.015	0.027
	NB 모델 MCMC $\alpha, \lambda$ (mode)	0.004	0.012	0.011	0.001	0.001	0.008	0.014	0.017	0.005	0.008	0.012	0.017

분포를 도출하였다.  $\alpha$  및  $\lambda$ 에 대한 추정치로는 사후평균(mean)과 사후최빈값(mode)을 모두 계산하여 수율모델에 적용하였다.

### 5.3 실험 결과

다양한 수율모델에 대한 클러스터링 파라미터의 성능 비교를 위하여 Poisson 모델, Murphy 모델, 그리고 NB 모델에 대한 예측수율을 계산하였다. 또한 클러스터링 파라미터 추정방법 간 비교를 위하여 임의의 값( $\alpha = 0.7$ ),

MME, MLE, 사후평균을 이용한 베이지안, 사후최빈값을 이용한 베이지안 방법을 비교하였다. 비교척도는 4개의 격자에 대해 각각 도출한 파라미터를 적용한 예측수율에 대하여 mean absolute deviation(MAD)를 사용하였다.

첫 번째 실험은 6개의 웨이퍼를 대상으로 4종류의 격자에 대하여 모두 실시하였으며, 결과 중 Wafer1, Wafer3, 그리고 6개 wafer의 평균에 대한 결과를 <표 3>에 정리하였다. 여기서 MCMC를 이용하여 도출한 파라미터는 3가지 경우, MCMC  $\alpha$ (mean) and MLE  $\lambda$ , MCMC  $\alpha$ (mean) and MCMC  $\lambda$ (mean), MCMC  $\alpha$ (mode) and MCMC  $\lambda$ (mode), 로



<표 4> Comparison results of various estimation methods with fabricating wafers

항 목		Average(6 wafers)				standard deviation(6 wafers)				Average	
		40 X 40	20 X 20	10 X 10	8 X 8	40 X 40	20 X 20	10 X 10	8 X 8	average	std
데 이 터	총 다이의 개수/wafer	1536.000	384.000	96.000	58.000						
	양품 다이의 개수/wafer	1377.000	266.167	39.333	18.333						
	D0(defects/wafer)	166.667	166.667	166.667	166.667						
	A(cm <sup>2</sup> )	0.045	0.181	0.720	1.190						
	$\lambda$	0.109	0.434	1.736	2.874						
파 라 미 터 추 정 치	fixed $\alpha$	0.700	0.700	0.700	0.700					0.700	
	Cunningham $\alpha$	-	13.152	1.885	1.409					5.482	
	MLE $\alpha$	100.000	44.170	1.880	1.137					36.797	
	MCMC $\alpha$ (mean)	1.000	2.041	1.626	1.408					1.519	
	MCMC $\alpha$ (mode)	1.013	2.000	1.378	1.145					1.384	
실제 수율		0.896	0.693	0.410	0.316					0.579	
예 측 수 율	Poisson 모델	0.900	0.675	0.295	0.175					0.511	
	Murphy 모델	0.901	0.686	0.329	0.212					0.532	
	NB 모델 fixed $\alpha$	0.908	0.741	0.480	0.383					0.628	
	NB 모델 Cunningham $\alpha$	-	0.680	0.378	0.291					0.450	
	NB 모델 MLE $\alpha, \lambda$	0.900	0.676	0.379	0.315					0.567	
	NB 모델 MCMC $\alpha$ (mean), $\lambda$	0.906	0.703	0.390	0.291					0.573	
	NB 모델 MCMC $\alpha$ (mode), $\lambda$	0.906	0.703	0.404	0.315					0.582	
M A D	Poisson 모델	0.003	0.022	0.115	0.141	0.003	0.028	0.057	0.035	0.070	0.031
	Murphy 모델	0.004	0.014	0.081	0.104	0.004	0.015	0.042	0.028	0.051	0.022
	NB 모델 fixed $\alpha$	0.012	0.048	0.072	0.071	0.012	0.032	0.033	0.042	0.051	0.030
	NB 모델 Cunningham $\alpha$	-	0.019	0.041	0.032	-	0.022	0.012	0.020	0.030	0.018
	NB 모델 MLE $\alpha, \lambda$	0.003	0.021	0.041	0.022	0.003	0.026	0.012	0.023	0.022	0.016
	NB 모델 MCMC $\alpha$ (mean), $\lambda$	0.009	0.011	0.031	0.031	0.010	0.004	0.011	0.020	0.021	0.011
	NB 모델 MCMC $\alpha$ (mode), $\lambda$	0.009	0.011	0.019	0.022	0.010	0.005	0.017	0.023	0.016	0.014

나누어 NB 수율모델에 적용하였다.

MAD 관점에서 MLE로 추정된 파라미터를 이용한 NB 모델이 가장 실제수율에 근접하고 있으며, 사후최빈값과 사후평균을 이용한 NB 모델도 MLE와 근소한 차이(평균적으로 1% 수율오차 이내)로 정확하게 수율을 예측하고 있다. 반면 MME는 분산이 평균보다 작아 음수가 도출되는 경우가 자주 발생하여 효과적이지 못한 방법으로 판단되었다. 베이저안 방법 내에서는 사후최빈값이 사후평균보다 다소 좋은 결과를 도출하였다. 그러나 MCMC를 이용한 베이저안 방법은 random sample을 이용하므로 실행 시 마다 결과가 조금씩 변하기 때문에 사후최빈값이 사후평균보다 반드시 성능이 좋다고 판단하기는 어렵다.  $\alpha$ 와  $\lambda$  모두 베이저안 추정치를 이용한 것이  $\alpha$ 만 베이저안 추정치를 적용한 것 보다는 성능이 우수하였는데, 이는 사후분포를 위한 마코프 체인의 샘플링과정에서  $\alpha$ 와

$\lambda$ 가 서로 상호보완적인 역할을 함을 알 수 있었다.

Poisson 모델과 Murphy 모델은 40 × 40과 20 × 20 격자에서 우수한 성능을 보이고 10 × 10과 8 × 8 격자에서는 매우 좋지 않은 성능을 보이고 있다. 이는 격자가 클수록 (다이의 면적이 증가할수록) 클러스터링 효과가 증가하여 예측수율이 작아지기 때문이다. 실제로 격자가 커질수록 MAD는 증가하고 있으며 이는 제 1장에서 언급한 Poisson 수율모델의 단점과 일치한다.

MLE는 전반적으로 가장 좋은 성능을 보이고 있으나 유의해야 할 점이 있는데, 그것은 MLE  $\alpha$ 의 추정치의 과도한 크기이다. 예를 들어, <표 3>의 Wafer3의 MLE  $\alpha$ 의 추정치는 40 × 40과 20 × 20 격자에서 100의 크기를 가지고 있다. MLE를 수치적으로 계산할 때 최대 추정치를 100으로 제한한 것을 감안하면, 실제 추정치는 100보다 큰 값을 가지고 있고 예상된다. 이렇게 편차가 큰

〈표 5〉 Comparison of various estimation methods

항 목	MLE	Bayesian (사후평균)	Bayesian (사후최빈값)
파라미터 추정	1	3	2
예측 정확성	3	2	1
예측 강건성	3	1	2

이유는 해당 격자의 관측치를 적용한 로그우도함수가 매우 완만하여 100보다 큰 값에서 최대치를 갖기 때문이다. 추정치의 큰 편차는 향후 변동성을 증가시키는 원인으로 작용할 가능성이 있음을 주의해야 한다.

지금까지는 이미 생산된 wafer의 결점 데이터로부터 얼마나 정확한 파라미터를 추정하는가에 대한 평가를 진행하였다. 그러나 클러스터링 파라미터를 추정하고자 하는 목적은 이미 생산된 wafer의 정확한 파라미터 추정이 아니다. 이미 생산된 wafer는 이미 실제수율이 측정되었으므로 파라미터 추정이 아무런 의미가 없기 때문이다. 본래의 목적은 이미 생산된 wafer를 통하여 클러스터링 파라미터를 추정하고 이를 미래에 생산할 wafer에 반영하여 얼마나 정확하게 수율을 예측할 수 있는가이다. 이러한 목적에 적합한 분석을 위하여 다음의 적용 프로세스를 고려해 보자.

- ① 6개의 wafer를 생산한다.
- ② 생산된 6개의 wafer에서 파라미터를 추정한다.
- ③ 추정한 파라미터의 평균값을 계산한다.
- ④ 새로운 6개의 wafer를 생산하는 과정에서 wafer의 결함을 관측한다(결함으로부터 결함크기분포와 critical area를 이용하여 결점의 수를 예측할 수 있다. 본 논문에서는 결점에 관한 것만 고려하므로 이 과정은 생략하였다).
- ⑤ wafer 당 결점수와 이미 추정한 파라미터 평균값을 이용하여 새로운 wafer의 수율을 예측한다.
- ⑥ 새로운 wafer의 실제수율과 예측수율을 비교 평가한다(미래에 생산되는 wafer는 기존과 동일한 공정을 거치므로 기생산된 wafer와 유사한 형태의 wafer map을 가지게 된다. 이를 반영하기 위해서 새로운 wafer는 기존과 동일한 wafer map을 갖는다고 가정한다).

이러한 프로세스를 가지고 분석한 결과를 <표 4>에 요약하였다. 실험결과는 자기 데이터를 이용한 단순 추정능력을 비교한 이전의 결과와 다소 다르게 나타났다. 사후최빈값을 이용한 베이지안 클러스터링 파라미터 추정치가 가장 정확도가 높았으며(평균 1.6%의 수율오차), 다음이 사후평균을 이용한 베이지안 추정치(평균 2.1%의 수율오차), 그리고 MLE(평균 2.2%의 수율오차) 순으

로 성능이 우수한 것으로 나타났다.

추정된 파라미터는 정확도도 높아야 하지만 외부의 영향에 대해 강건성(robustness)을 가져야 한다. 파라미터의 강건성을 측정하기 위해서 본 논문에서는 6개 wafer의 MAD에 대한 표준편차를 사용하였다. 강건성 측면에서는 사후평균을 이용한 베이지안 클러스터링 파라미터 추정치가 가장 정확도가 높았으며(평균 0.011), 다음이 사후최빈값을 이용한 베이지안 추정치(평균 0.014), 그리고 MLE(평균 0.016) 순으로 성능이 우수한 것으로 나타났다.

지금까지의 결과를 바탕으로 NB 수율모델을 위한 목적에 따른 파라미터 추정법의 순위를 정리하면 <표 5>와 같다. 자기데이터를 이용한 파라미터 추정은 MLE가 가장 우수하며, 베이지안 방법도 비교적 우수한 성능을 가지고 있다. MME는 음수의 파라미터를 도출하여 반도체 수율에는 적용이 어려운 것으로 판단된다. 미래의 예측력에 있어서는 베이지안 추정법이 기존의 방법에 비하여 좀 더 안정적이며 우수한 성능을 가지고 있다고 판단할 수 있다.

## 6. 결 론

기존에 결함을 이용한 반도체 수율예측 방법은 많이 이루어 졌으나 이를 실제로 적용하기 위해서는 파라미터의 추정이 필수적이다. 본 논문에서는 반도체 wafer 상에서 클러스터링 파라미터를 추정함에 있어 MCMC를 이용한 베이지안 추정법을 사용하여 파라미터를 도출하는 방법을 새롭게 제시하였다. 또한, 기존의 추정법과의 비교를 통하여 반도체 수율예측에 적합한 추정법을 도출하고자 하였다. 사용목적에 따라 MLE와 베이지안 추정법이 우수한 성능을 보였으나 본 논문의 데이터로는 모든 목적에 확실한 비교우위를 갖는 추정법은 존재하지 않았다. 클러스터링 파라미터는 결함의 특성 파악 또는 수율예측을 위한 기본 데이터 등 다양한 목적의 사용이 가능하다. MLE나 MCMC가 모든 목적에 대해 확실한 비교우위를 갖지 않으므로 다양한 목적에 따라 계층화분석과정(Analytic Hierarchy Process : AHP)와 같은 다기준 의사결정방법을 사용하여 필요 목적 별로 적절한 방법을 선택하는 것도 고려할 수 있다.

추후 연구로는 결함크기분포와 critical area를 포함한 파라미터 추정에 관한 연구를 고려하고 있다. 본 논문에서는 클러스터링 파라미터의 추정이라는 주제에 집중하기 위하여 결함으로부터 결점을 도출하는 과정을 단순화하여 적용하였으나, MCMC를 이용한 베이지안 추정법은 복잡한 결함분포를 가진 파라미터 추정에 장점이 있으므로 제시한 연구에 장점이 있을 것으로 판단된다.

## 참고문헌

- [1] Andrieu, C., Freitas, N. D., Doucet, A., and Jordan, M. I.; "An introduction to MCMC for machine learning," *Machine Learning*, 50 : 5-43, 2003.
- [2] Bae, S. J., Hwang, J. Y., and Kuo, Way.; "Yield prediction via spatial modeling of clustered defect counts across a wafer map," *IIE Transactions*, 39 : 1073-1083, 2007.
- [3] Congdon, P.; *Bayesian Models for Categorical Data*, Wiley, Chichester, 2005.
- [4] Cunningham, J. A.; "The use of evaluation of yield models in integrated circuit manufacturing," *IEEE Transactions on Semiconductor Engineering*, 3(2) : 60-71, 1990.
- [5] Geman, S. and Geman, D.; "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6) : 721-741, 1984.
- [6] Ghosh, S. K., Mukhopadhyay, and P., Lu, J. C.; "Bayesian analysis of zero-inflated regression models," *Journal of Statistical Planning and Inference*, 136 : 1360-1375, 2006.
- [7] Gilks, W. and Wild, P.; "Adaptive rejection sampling for Gibbs sampling," *Applied Statistics*, 41 : 337-348, 1992.
- [8] Ha, C.; "Effective construction method of defect size distribution using AOI data: application for semiconductor and LCD manufacturing," *IE Interfaces*, 21(2) : 151-160, 2008.
- [9] Hastings, W. K.; "Monte Carlo sampling methods using Markov chains and their application," *Biometrika*, 57 : 97-109, 1970.
- [10] ICE; *Yield and Yield Management*, Chapter 3 in *Cost Effective IC Manufacturing 1998~1999*, Integrated Circuit Engineering Corporation, 1997.
- [11] Milor, L. S.; "Yield modeling based on in-line scanner defect sizing and a circuit's critical area," *IEEE Transactions on Semiconductor Manufacturing*, 12(1) : 26-35, 1999.
- [12] Koren, I. and Koren, Z.; "Defect tolerance in VLSI circuits: Techniques and yield analysis," *Proceedings of the IEEE*, 86(9) : 1819-1837, 1998.
- [13] Kuo Way, Chien, W. K., and Kim, T.; *Reliability, Yield, and Stress Burn-in : An Unified Approach for Microelectronics Systems Manufacturing and Software Development*, Norwell, Kluwer Academic Publishers, 1998.
- [14] Spiegelhalter, D. J., Thomas, A., Best, N. G., and Lunn, D.; *WinBUGS Version 1.4 Users Manual*, MRC Biostatistics Unit, Cambridge, URL <http://www.mrc-bsu.cam.ac.uk/bugs/>, 2003.
- [15] Stapper, C. H.; "Modeling of defects in integrated circuit photolithographic patterns," *IBM Journal of Research and Development*, 28(4) : 461-475, 1984.
- [16] Stapper, C. H.; "Improved yield models for fault-tolerant memory chips," *IEEE Transactions on Computers*, 42(7) : 872-881, 1993.
- [17] Sturtz, S., Ligges, U., and Gelman, A.; "R2WinBUGS : a package for running WinBUGS from R," *Journal of Statistical Software*, 12(3) : 1-17, 2005.