

포아송 분포를 가정한 Wafer 수준 Statistical Bin Limits 결정방법과 표본크기 효과에 대한 평가

박성민^{1*} · 김영식²

삼성전자, System LSI

Methods and Sample Size Effect Evaluation for Wafer Level Statistical Bin Limits Determination with Poisson Distributions

Sungmin Park · Youngsig Kim

¹Senior engineer / ²Principal engineer, System LSI, Samsung Electronics Co., Ltd.

In a modern semiconductor device manufacturing industry, statistical bin limits on wafer level test bin data are used for minimizing value added to defective product as well as protecting end customers from potential quality and reliability excursion. Most wafer level test bin data show skewed distributions. By Monte Carlo simulation, this paper evaluates methods and sample size effect regarding determination of statistical bin limits. In the simulation, it is assumed that wafer level test bin data follow the Poisson distribution. Hence, typical shapes of the data distribution can be specified in terms of the distribution's parameter. This study examines three different methods; 1) percentile based methodology; 2) data transformation; and 3) Poisson model fitting. The mean square error is adopted as a performance measure for each simulation scenario. Then, a case study is presented. Results show that the percentile and transformation based methods give more stable statistical bin limits associated with the real dataset. However, with highly skewed distributions, the transformation based method should be used with caution in determining statistical bin limits. When the data are well fitted to a certain probability distribution, the model fitting approach can be used in the determination. As for the sample size effect, the mean square error seems to reduce exponentially according to the sample size.

Keywords: statistical bin limits, Poisson distribution, sample size, percentile, data transformation, wafer level test bin

1. 서론

1.1 연구배경

반도체 제품군에서, ASIC(Application Specific Integrated Circuit) 및 Soc(System On Chip) 제품 등을 포함한 비메모리 LSI(Large

Scale Integration) 제품군은 '다품종 변량' 생산특성을 갖는다. 한편, 제품의 고기능화가 가속화되고, 또한 life cycle이 단축되는 상황에서, 통계적 품질관리·보증(statistical quality control and assurance) 체계인 Statistical Bin Limits(SBL) 시스템이, 최근 반도체 산업계에서 활용되고 있다(Park *et al.*, 2000). 반도체 wafer 제조공정(semiconductor wafer fabrication process,

*연락처 : 박성민, 449-711 경기도 용인시 기흥읍 농서리 산24 삼성전자, 반도체총괄 DSN, System LSI, SOC 제품기술팀

Fax : 031-209-6229, E-mail : sungmin.park@samsung.com

2003년 9월 접수, 2회 수정 후 2003년 12월 게재 확정.

fab) 안정화 및 IC(Integrated Circuit) 제품특성을 대변하는 총체적 지표로서, 일반적으로 wafer 수율(즉, net-die-per-wafer-yield)이 활용될 수 있다. 하지만, wafer 수율만으로는, 공정의 미세변화를 관리하거나 제품의 세부특성을 파악하기에는 미흡하며, 이런 경우 SBL 시스템이 더욱 요구된다(Cheong *et al.*, 1996).

SBL 시스템의 목적은, 다음과 같이 크게 세 가지로 정리된다; 1) (품질 · 신뢰성 측면) 잠재적인 품질 · 신뢰성 문제가 내재되었다고 의심되는 제품을 선별해, 최종적으로 고객에 대한 품질을 보증; 2) (자원활용 측면) 비정상(abnormal) wafer가 반도체 backend 공정으로 인도됨으로써, 추가적으로 발생할 불필요한 부가가치 손실을 최소화; 그리고 3) (생산성 측면) wafer 수준 test bin 데이터 분석으로 도출된 유의한 결과는 관련 frontend 공정으로 신속히 feedback하여, wafer 수율향상을 도모하고, 제품 R&D 기간 · 양산제품 납기 등을 단축하는 것 등이다.

SBL 시스템을 활용해; 1) 비정상 wafer는 폐기(scrap); 그리고 2) 비정상 run은 정지(hold)된 후, re-test 혹은 불량분석(failure analysis; FA)된다. 한편, 협의의 SBL은, 정상 및 비정상 wafer를 구분짓는 wafer 수준 test bin 데이터의 통계적 관리한계를 의미한다. SBL은 비정상 wafer · run 판정을 위한 주요 관리지표이므로, 적절한 SBL 결정방법이 요구된다. 만약, 품질 · 신뢰성 향상을 위해 SBL이 엄격하게(tight) 결정되면, 요구되는 일정 수준을 초과한 wafer scrap 및 run hold로 인해, 과도한 생산물량 손실 및 제품납기 지연 등을 초래할 수 있다. 반면, 생산물량 확보가 강조돼 SBL이 느슨하게(loose) 결정되면, 품질 · 신뢰성 의심제품에 대한 선별능력이 저하될 수 있다.

1.2 Wafer 수준 test bin 데이터

Electrical Die Sorting(EDS) 공정에서, wafer 상 모든 die(즉, chip)는 ‘bin’이라는 전기적 특성항목으로 분류된다. 각 die는 미리 정해진 순차적인(sequential) test bin 항목에 의해 검사 · 분류되며, 모든 bin 항목에서 합격되어 검사완료된 die는 최종적으로 양품으로 판정된다. 반면, 임의의 특정 bin 한 개 항목에서라도 불량이면, 이 때 해당 bin에서의 불량 die 수는 ‘1’ 증가하고, 검사된 die는 추후 inking되어 불량품으로 폐기된다.

특정불량 bin으로 분류된 die의 고장형태(failure mode)는, 크게 두 가지 경우로 분류할 수 있다; 1) 직접적인 수율저하를 초래하는 경성고장(hard failure, 예: metal circuit patterning short or open); 또는 2) 제품검사규격 한계(marginal) 영역에서만 동작됨으로써, 품질 · 신뢰성에 문제가 있다고 판단된 연성고장(soft failure, 예: circuit current leakage).

식 (1)과 같이, 총 I 개 항목으로 구성된 wafer 수준 test bin 체계에서, EDS 종료된 wafer 1매의 각 bin별 die 수의 합계는 wafer 상 ‘net-die’ 수 (즉, wafer edge 부근의 일정 영역을 제외한, 중심부에 위치한 제품화 가능한 사각형 die 수) d_n 과 동일하다.

$$\sum_{i=1}^I x_i = d_n \quad (1)$$

$$y_w = \frac{x_1}{d_n} = \frac{1}{d_n} (d_n - \sum_{i=2}^I x_i) \quad (2)$$

식 (1)에서 x_i 는 i 번째 bin 항목의 불량 die 수, 즉, i 번째 bin 항목의 wafer 수준 test bin 데이터를 위한 확률변수(random variable) X_i 가 취할 수 있는 값이다. 한편, 일반적인 wafer 수준 test bin 체계에서는, 예외적으로 ‘good-die’ (즉, 양품 die) 수를 X_1 으로 관리하며, 나머지는 그 일련번호에 상응하는 bin 항목에서의 불량 die 수를 표시한다. 그러므로, wafer 수율 y_w 는, 식 (2)와 같이 계산될 수 있다. 이후, 본 논문에서는 특별한 언급이 없으면; 1) ‘수율’은 wafer 수율; 그리고 2) ‘bin 데이터’는, ‘Bin1’을 제외한 나머지 불량 bin에서, EDS 종료된 wafer 수준 test bin 데이터를 지칭한다. <그림 1>은 EDS 종료된 wafer 수준 test bin 데이터를 이용한 ‘EDS bin map’ 실례이다. 거의 대부분의 die가 Bin1, 즉, 양품이며, 대체로 wafer edge 영역에서 불량 die가 존재함을 확인할 수 있다.

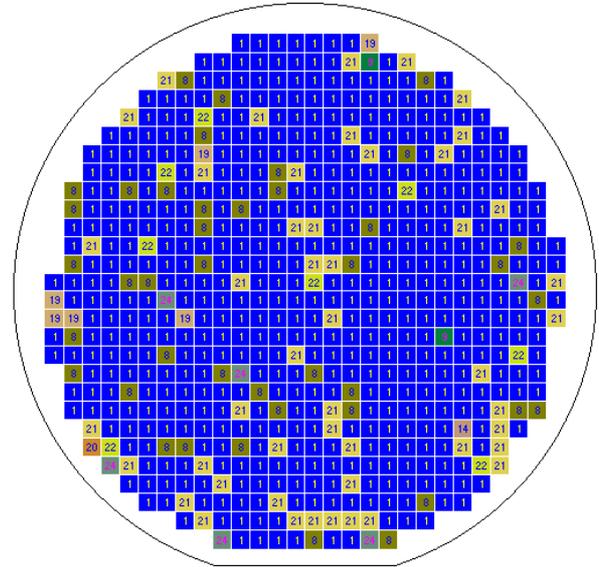


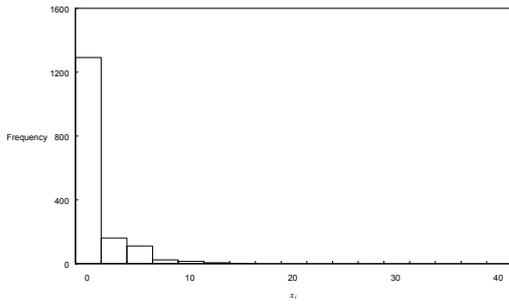
그림 1. An illustration: EDS bin map.

1.3 연구주제

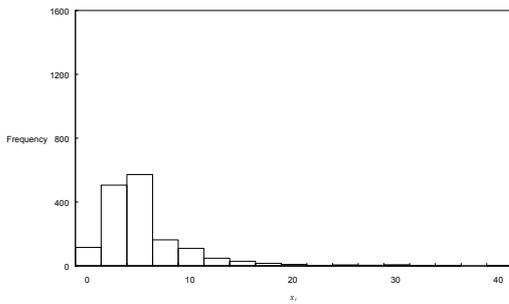
대부분의 bin 데이터가 정규성(normality)을 만족하지 못하기 때문에, 정규성에 근거한 bin 데이터의 통계적 추론은 타당한 결과를 도출하기 어렵다. <그림 2>는, wafer 직경 8inch(즉, 20cm), chip 크기 $0.4235cm^2$, net-die 603개인 임의의 ASIC 제품의, 4개월 간 생산된 125개 run, 1609매 wafer를 이용한, 세 가지 bin 데이터 히스토그램이다.

<그림 2.a>는 불량률이 0이 아닌 불량 bin 중에서, 불량률 하위 50%에 속하며 상대적으로 불량률이 낮은 SRAM BIST/BIRA 고장과 관련이 있는 bin 데이터 분포를 보여준다. IC chip 내부에 포함된 SRAM, DRAM 등과 같은 memory block test를 위해, Direct Access(DA) test가 일반적으로 실시된다. 반면, memory test

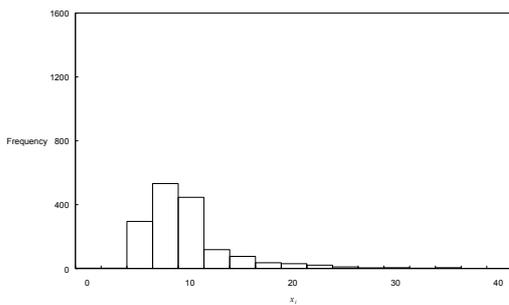
speed 혹은 coverage를 높이기 위해, 자체 test logic을 chip 내부에 설계·삽입하여 memory test를 실시할 수 있도록 고안된 것이, Built In Self Test(BIST) 방식이다. 한편, repairable redundancy dummy block이 있을 경우에는, failure memory cell address까지 확인할 수 있는 test 방식인 Built In Redundancy Analysis(BIRA)를 활용할 수 있다. 단, BIST 방식은 logic block test에도 사용될 수 있다.



(a)



(b)



(c)

그림 2. Bin data histograms; (a) SRAM BIST/BIRA; (b) short; and (c) open.

반면, <그림 2.b>와 <그림 2.c>는 순서대로 불량률이 상위 15, 10% 내에 속하는 회로단락(short) 및 개방(open)과 관련된 bin 데이터 분포다. 한편, fab 공정이 안정화될수록, ‘critical bin’으로 지칭되는 소수의 특정 bin에서만 불량이 발생되고, 나머지 대부분 bin에서는 불량이 발생되지 않는다. <그림 2>의 분석

대상 제품은; 1) 불량 die 중 90%가 전체 (I-1)개 bin의 약 4% bin만으로 분류되고; 2) 70%의 bin에서는 불량이 발생하지 않는다.

<그림 2>에서도 볼 수 있듯이, 심한 우변사향(skewed to the right) 분포라 하더라도, 우측 꼬리부분으로 길게 늘어진 극한의 일정영역에 위치한 bin 데이터(즉, 이런 bin 데이터를 갖는 비정상 wafer)를, 나머지 bin 데이터(즉, 불량이 상대적으로 적은 정상 wafer) 군과 구분짓는 적절한 SBL이 필요하다. 만약, 비정상 bin 데이터 판정비율이 일정 수준으로 주어지면, 비록 시간경과에 따른 fab의 동적 변동(dynamic variation) 등과 같은 주요인을 감안하더라도, 주어진 수준에 근접한 비정상 bin 데이터 판정비율을 실현할 수 있는 적절한 SBL 결정방법이 필요하다.

현장에서는; 1) 생산물량 관리에 초점을 둔 fab, EDS 관련 부서와; 2) 품질지향의 제품기술(product engineering, PE), 품질보증(quality assurance, QA) 관련 부서의 협의하에, wafer scrap 및 run hold을 수준에 대한 지침이 주기적으로 검토될 수 있으며, 이에 근거해 요구되는 일정 수준의 비정상 bin 데이터 판정비율이 제공될 수 있다.

본 연구에서는, <그림 2>에서 검토된 바와 같이, 몇 가지 전형적인 비대칭도(skewness)를 갖는 bin 데이터 분포에 대해, 적절한 SBL 결정방법 및 표본크기(즉, wafer 매수)를, Monte Carlo 시뮬레이션을 통해 검토하고자 한다. 본 논문, 2장에서는 시뮬레이션 실험설계에 대해 설명하고, 3장에서는 시뮬레이션 데이터 분석결과를 설명한다. 또한, 4장의 사례분석을 통해, 잡음(noise)이 상대적으로 심한 실제 bin 데이터의 경우에 대한 검토 결과를 보완한다. 한편, 부록에 반도체 frontend 및 backend 공정에 대한 독자의 이해를 돕고자, 반도체 IC 제품 제조공정을 요약·설명한다.

2. 실험설계

반도체 IC 제품 제조공정과 관련된 전형적인 Poisson 분포의 응용 예로서, 생산단위당 결점수 모형화 연구가 있다(예: wafer당 bin 불량수 혹은 die당 defect 개수). Illyes and Baglee(1992)는, Intel Corporation의 Poisson 분포를 이용한 SBL 시스템을 소개한 바 있다. Poisson 분포를 가정한 net-die-per-wafer-yield 모형, 이른바 ‘Poisson 수율모형’을 이용한 수율 및 defect density(D_0) 관리활동도 반도체 산업에서 널리 활용되고 있다(Cunningham, 1990).

그러므로; 1) 제품별 wafer 상 d_n 이 일정; 그리고 2) X_i 분포의 비대칭도를 고려할 때, X_i 는 Poisson 분포를 따른다고 가정한다. Poisson 분포의 이산(discrete) 확률변수 X_i 의 확률질량함수(probability mass function)는, 모수(parameter) $\lambda_i > 0$ 을 갖는 식 (3)과 같이 정의된다.

$$p(x_i) = \frac{e^{-\lambda_i} \lambda_i^{x_i}}{x_i!} \quad x_i = 0, 1, \dots \quad (3)$$

2.1 실험조건

- 비대칭도(λ_i): 식 (3)에서, 모수 λ_i 가 취하는 값에 따라 비대칭도가 변하는데, λ_i 가 커질수록 분포는 대칭된다. <그림 3>은 $\lambda_i=0.5, 5.0, 10.0$ 인 Poisson 분포이며, 분포형태를 명확히 보이기 위해 편의상 점선으로 연결되었다. 본 연구에서는, bin 데이터의 비대칭도를 λ_i 로 정의하고, <그림 3>의 세 가지 bin 데이터 분포가, 나머지 연구대상 bin 데이터 분포에 대한 대표성이 있다고 가정한다. 일반적으로, λ_i 가 10 이상이면, Poisson 분포의 비대칭도가 현저히 감소됨을 확인할 수 있다. <그림 2>는, 순서대로 $\hat{\lambda}_i=0.6, 5.5, 10.3$ 인 bin 데이터의 histogram이다.

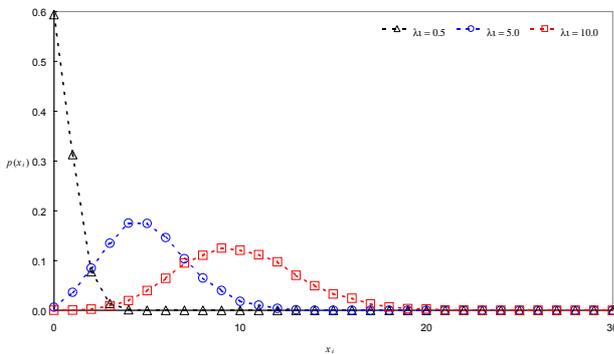


그림 3. Poisson probability distributions for selected values of λ_i .

- 표본크기(n): SBL 결정 시 사용될 표본크기 n 의 범위는 다음 fab-in 투입물량에 대한 가정을 근거로 설정한다; 1) 25wafers/run; 및 2) 1~10runs/주 × 4주/월. 그러므로, 월별로 수집된 최근 1개월 간 데이터를 이용해 SBL을 결정하고, 이렇게 결정된 SBL을 현재 해당 월 데이터에 적용한다면, 시뮬레이션에서 고려할 수 있는 n 의 범위는(min, max)=(100, 1000)이 되며, $n \leftarrow n+100$ 으로 증가되는 표본크기에 대해서 시뮬레이션을 실시할 수 있다.
- 비정상 bin 데이터 판정비율(p): 비정상 bin 데이터 판정비율 p 는, 실제 fab에서 wafer scrap 및 더 나아가 run hold율과 연계된다. 표준정규분포를 따르는 확률변수 Z 의 누적확

률을 이용해 p 를 설정하면, <표 1>과 같다. 즉, Z 가 취하는 값 z 를, Z 의 표준편차의 양의 정수배만큼 평균에서 떨어진 값(즉, 'sigma 수준')으로 정하고, 그 값을 초과하는 단측영역(one-sided) 확률을 p 로 정한다. 그러므로, p 와 Poisson 분포의 모수 λ_i 값을 이용해 이론적으로 계산된 SBL 모수 l_i^p 가 <표 1>에 함께 제시된다. SBL 모수 l_i^p 는, 모수 λ_i 를 갖는 Poisson 분포의 임계값(critical value)으로서, 이 값을 초과하는 단측영역 확률 p 를 갖도록 정의된다. 한편, SBL 모수 l_i^p 이산 추정치 계산과 관련된 내용은, 식 (7)에 대한 설명으로 대신한다.

2.2 SBL 결정방법

이론적으로는, Poisson 분포를 따르는 bin 데이터를 가정한 후, SBL을 구할 수 있다. 본 연구에서는, 비정규(nonnormal) bin 데이터의 SBL 결정 시, 현장 엔지니어에 의해 용이하게 활용되고 있는, 다음 세 가지 방법을 비교하고자 한다.

- 백분위수(Percentile; PT): 비모수적(nonparametric) 방법으로서, <표 1>에 정리된 p 값에 상응하는; 1) 97.725th; 2) 99.865th; 그리고 3) 99.997th 백분위수를 이용해 SBL을 결정한다.
- 데이터 변환(Data Transformation; DT): Box-Cox 데이터 변환을 통해, 데이터에 정규성과 분산안정화(variance stabilization)를 유도한다. 식 (4)처럼, Poisson 분포를 따르는 확률변수 X_i 의 평균 μ_{X_i} 와 분산 $\sigma_{X_i}^2$ 는 같기 때문에, 간단히, 분산안정화만을 위한, 식 (5)와 같은 제곱근 누승변환(power transformation)을 실시해, '변환 데이터' X_i' 을 얻을 수도 있다(Montgomery, 1997).

$$\mu_{X_i} = \sigma_{X_i}^2 \tag{4}$$

$$X_i' = (X_i + 1)^{\tau_i} = (X_i + 1)^{0.5} \tag{5}$$

반면, $\lambda_i=0.5, 5.0, 10.0$ 인 Poisson 분포를 따르는 $n=100,000$ 인 dataset을 각각 생성해, Box-Cox 데이터 변환시 사용될 τ_i 값

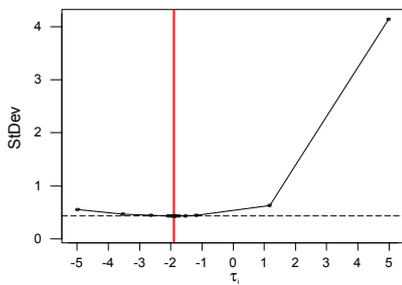
표 1. Disposition proportion of abnormal bin data & SBL parameter

Sigma level	Cumulative probability	Disposition proportion of abnormal bin data	SBL parameter		
			l_i^p		
z	$\Phi(z) = P(Z \leq z)$	$p = 1 - \Phi(z)$	$\lambda_i = 0.5$	$\lambda_i = 5.0$	$\lambda_i = 10.0$
+2	0.97725	0.02275	2	10	17
+3	0.99865	0.00135	4	13	21
+4	0.99997	0.00003	5	16	25

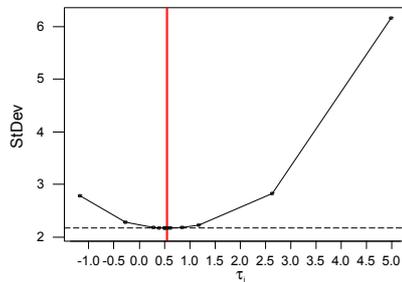
을 MINITAB™을 이용하여 추정된 결과가 <그림 4>에 제시된다. $\lambda_i=5.0, 10.0$ 인 경우, 대략 $\hat{\tau}_i=0.5$, 비대칭도가 심한 $\lambda_i=0.5$ 일 때, 대략 $\hat{\tau}_i=-2.0$ 이다. <그림 4>에서의 표준편차(StDev)는, Box-Cox 변환식으로 변환된 데이터를, I-MR (Individual Measurements-moving Range) 관리도에서 사용되는 식 (6)을 이용해 계산된 값과 일치한다(Montgomery, 2001).

$$\text{StDev} = \frac{\overline{\text{MR}}}{d_2} = \frac{\overline{\text{MR}}}{1.128} \quad (6)$$

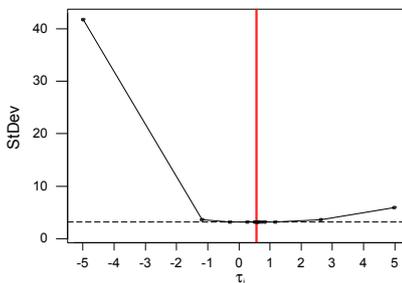
본 분석에서는, SBL 결정시 사용되는 dataset 고유의 τ_i 값을 Box-Cox 데이터 변환을 통해 매 번 추정한다. 각 dataset마다의 $\hat{\tau}_i$ 을 이용해, 변환데이터 X_i 의 SBL을 결정한 후, 이 값을 ‘역변환’(inverse transformation)하여 X_i 의 SBL을 구한다(Montgomery, 1997).



(a)



(b)



(c)

그림 4. Plots of StDev versus τ_i ; (a) $\lambda_i=0.5$; (b) $\lambda_i=5.0$; and (c) $\lambda_i=10.0$.

- 모형적합(Model Fitting; MF): 주어진 표본크기 n 을 이용해,

λ_i 를 추정한 후, 식 (7)처럼 Poisson 분포의 누적확률이 p 값 이상을 만족하는 가장 작은 정수 c_i^{min} 을 SBL로 결정한다.

$$P\{x_i \leq c_i\} = \sum_{x_i=0}^{c_i} \frac{e^{-\hat{\lambda}_i} \hat{\lambda}_i^{x_i}}{x_i!} > p \quad x_i = 0, 1, \dots \quad (7)$$

2.3 시나리오

2.1절 및 2.2절에서 설명된 실험조건 및 SBL 결정방법을 조합해 시뮬레이션 실험 시나리오를 수립한다. 즉, 주어진 비대칭도에 대응하는 모수의 Poisson 분포로 확률표집(random sampling)된 동일한 dataset을 대상으로; 1) 방법; 및 2) 표본크기를 달리하여 주어진 sigma 수준에 대응하는 SBL을 추정한다. 이와 같은, 시뮬레이션을 $r=100$ 회 반복(replication)한다.

2.4 성능측도

각 시뮬레이션 조합별 성능측도(performance measure)로서, 식 (8)과 같이 정의된 평균제곱오차(Mean Square Error; MSE)를 이용한다. 식 (8)은, X_i 의 SBL 모수 l_i^p 에 대한 추정치 \hat{l}_i^p 의 정밀도(즉, 분산, variance)와 정확도(즉, 편의, bias)를 동시에 평가한다(Myers and Montgomery, 1995).

$$\begin{aligned} \text{MSE} &= E [\hat{l}_i^p - l_i^p]^2 \\ &= \text{var} (\hat{l}_i^p) + [E (\hat{l}_i^p) - l_i^p]^2 \\ &= \text{variance} + \text{bias}^2 \end{aligned} \quad (8)$$

3. 실험 및 분석

각 시뮬레이션 조합별 SBL 추정치의 MSE가 <표 2>에 정리되어 있다. 단, $\lambda_i=0.5$ 일 때, 데이터 변환에 의한 SBL 추정치의 MSE는 도출할 수 없다(not available, n/a). <표 3>에, 데이터 변환시 사용된 $\hat{\tau}_i$ 의 평균(Avg) 및 표준편차의 통계치가 제시되어 있다. $\lambda_i=0.5$ 의 경우도, 모든 dataset에 대해, 데이터 변환은 가능하지만, <표 4>에 제시된 것처럼 SBL이 양의 정수로 추정된 성공횟수가 없거나 매우 드문 것으로 확인되었다. 즉, 변환 데이터의 SBL이 음수로 추정되어, 역변환이 불가능하기 때문이다.

<표 2>의 시뮬레이션 결과를 그래프화 하여, 비대칭도 및 sigma 수준에 따른 SBL 결정방법 및 표본크기 효과를 검토한다. <그림 5~7>은, $\lambda_i=0.5, 5.0, 10.0$ 에 해당하는 <표 2>의 결과를 그래프화한 것이다. 이 중, <그림 6>을 중심으로, 시뮬레이션 결과에 대한 분석을 설명한다.

<그림 6>은 $\lambda_i=5.0$ 인 MSE 데이터를 이용한 그래프이며, 단순선형회귀(simple linear regression) 모형으로, 시뮬레이션 결과에 대한 metamodel 결과가 추가되어 있다(Law and Kelton, 1991). 즉, <그림 6>의 각 panel마다, MSE 데이터를 잇는 꺾은

표 2. MSE's

Method	n	$\lambda_i=0.5$			$\lambda_i=5.0$			$\lambda_i=10.0$		
		$z=+2$	$z=+3$	$z=+4$	$z=+2$	$z=+3$	$z=+4$	$z=+2$	$z=+3$	$z=+4$
Percentile	100	0.16136	1.44362	4.52362	0.61553	4.76775	24.14775	1.35175	7.48872	42.44872
	200	0.20162	0.89331	3.39331	0.45432	2.53575	17.29774	0.82770	3.88031	30.44354
	300	0.15129	0.64330	2.76330	0.31313	1.38197	12.78427	0.65652	2.17289	24.63789
	400	0.09083	0.50272	2.43331	0.18156	1.04818	11.59354	0.40364	1.69082	21.67786
	500	0.07066	0.44283	2.20354	0.16147	0.86765	10.50444	0.34318	1.50078	19.58694
	600	0.05048	0.40287	2.04354	0.10099	0.78747	9.65475	0.25240	1.29040	17.91594
	700	0.06057	0.33280	1.76323	0.10097	0.67669	8.82371	0.16158	1.14003	16.59541
	800	0.04039	0.57248	1.48273	0.12111	1.02642	7.98170	0.23210	1.62889	15.66604
	900	0.05048	0.50253	1.36281	0.10097	0.88693	7.57424	0.19175	1.45955	14.37559
	1000	0.04039	0.44249	1.28283	0.08077	0.73627	7.18376	0.20187	1.20872	13.89928
Data transformation	100				0.36336	1.77440	7.74879	0.84014	3.37741	12.26922
	200				0.19470	0.88888	3.56289	0.41256	1.54366	5.00121
	300				0.13054	0.68002	3.00697	0.25102	0.91032	3.00513
	400				0.09030	0.44988	2.10621	0.15893	0.62010	2.28497
	500	n/a	n/a	n/a	0.07000	0.34469	1.67660	0.13377	0.47191	1.71171
	600	n/a	n/a	n/a	0.06309	0.29775	1.45708	0.11130	0.40873	1.50445
	700				0.05402	0.25040	1.28350	0.09513	0.33094	1.23498
	800				0.05220	0.25210	1.27589	0.07824	0.28451	1.13082
	900				0.04701	0.22930	1.19038	0.06853	0.25260	1.06481
	1000				0.03871	0.20292	1.12243	0.06340	0.25090	1.12192
Poisson model fitting	100	0.09083	0.32220	0.16147	0.15134	0.23196	0.26259	0.33289	0.40325	0.33311
	200	0.01010	0.28204	0.04039	0.05048	0.11103	0.14138	0.20169	0.29219	0.19180
	300	0.00000	0.23179	0.04039	0.02020	0.06059	0.07070	0.10095	0.26204	0.09091
	400	0.01010	0.15129	0.01010	0.03029	0.05048	0.06059	0.08074	0.19155	0.08080
	500	0.00000	0.09083	0.01010	0.01010	0.03029	0.02020	0.04039	0.17143	0.04040
	600	0.00000	0.11099	0.01010	0.00000	0.01010	0.00000	0.02020	0.14122	0.02020
	700	0.00000	0.09083	0.00000	0.00000	0.00000	0.00000	0.04039	0.12107	0.01010
	800	0.00000	0.06057	0.00000	0.00000	0.00000	0.00000	0.01010	0.10091	0.00000
	900	0.00000	0.05048	0.00000	0.00000	0.00000	0.00000	0.02020	0.08074	0.02020
	1000	0.00000	0.04039	0.00000	0.00000	0.00000	0.00000	0.00000	0.06057	0.00000

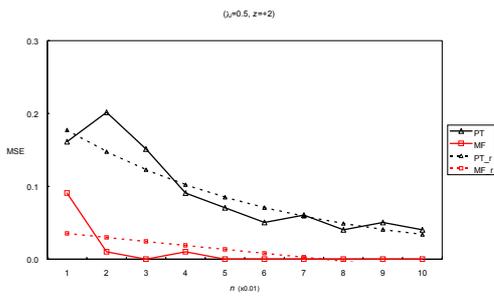
표 3. Statistics for $\hat{\tau}_i$

n	$\lambda_i=0.5$		$\lambda_i=5.0$		$\lambda_i=10.0$	
	Avg	StDev	Avg	StDev	Avg	StDev
100	-1.96	0.69	0.55	0.20	0.58	0.29
200	-1.96	0.50	0.55	0.14	0.60	0.19
300	-1.94	0.40	0.54	0.13	0.59	0.15
400	-1.90	0.32	0.53	0.10	0.58	0.13
500	-1.90	0.28	0.53	0.08	0.58	0.12
600	-1.92	0.26	0.53	0.07	0.59	0.11
700	-1.91	0.22	0.53	0.07	0.59	0.10
800	-1.89	0.21	0.54	0.06	0.59	0.09
900	-1.89	0.19	0.54	0.06	0.58	0.09
1000	-1.90	0.19	0.54	0.06	0.57	0.09

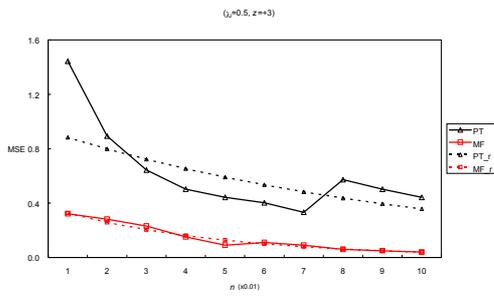
선과, MSE 데이터로 추정된 회귀모형의 평활곡선(smooth curve)이 함께 제시되어 있다. 추가된 평활곡선으로, 표본크기에 대한 MSE 감소효과를 보다 선명하게 파악할 수 있다.

표 4. Number of successes for SBL determination by data transformation

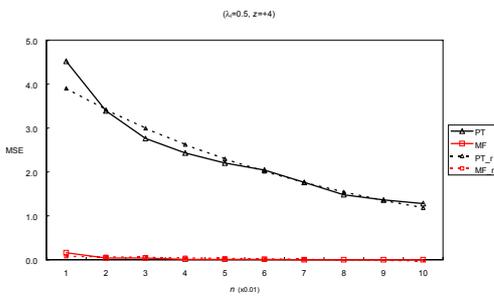
n	$\lambda_i = 0.5$		
	z = +2	z = +3	z = +4
100	28	3	0
200	22	1	0
300	18	0	0
400	14	0	0
500	13	0	0
600	8	0	0
700	5	0	0
800	4	0	0
900	4	0	0
1000	3	0	0



(a)

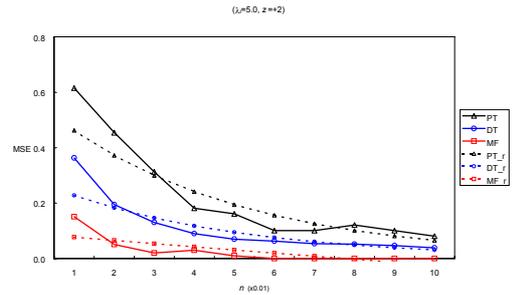


(b)

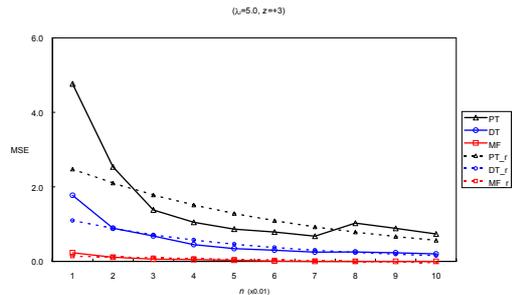


(c)

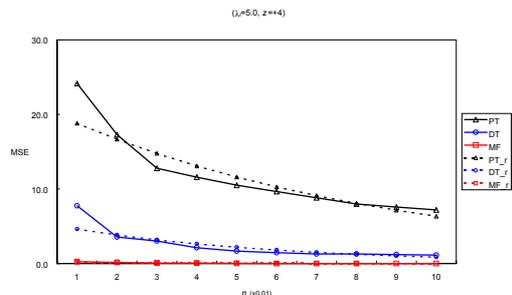
그림 5. MSE graphs ($\lambda_i = 0.5$); (a) z = +2; (b) z = +3; (c) z = +4.



(a)

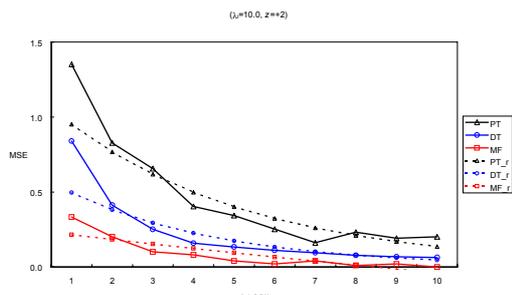


(b)

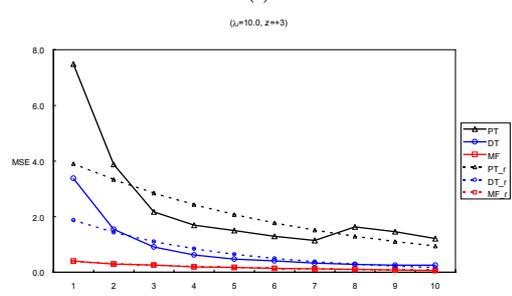


(c)

그림 6. MSE graphs ($\lambda_i = 5.0$); (a) z = +2; (b) z = +3; (c) z = +4.



(a)



(b)

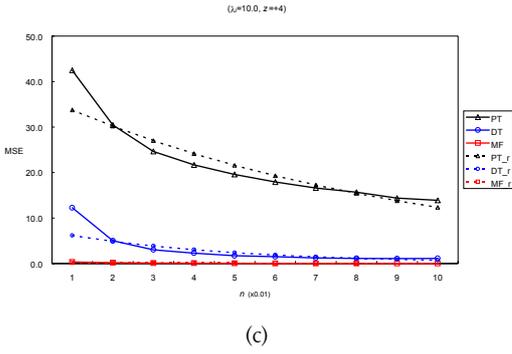


그림 7. MSE graphs ($\lambda_i = 10.0$); (a) $z = +2$; (b) $z = +3$; (c) $z = +4$.

Metamodel 수립시, 식 (9)와 같은 비선형 모형이 가정된다. 하지만, 식 (9)는 모형변환(model transformation)을 통해 직선화가 가능한 내재선형(intrinsically linear) 식으로서, 식 (9)는 식 (10)의 로그변환(logarithmic transformation)으로써 식 (11)과 같이 변환되어 추정될 수 있다(Montgomery and Peck, 1992).

$$y = \beta_0 e^{\beta_1 x} \epsilon \tag{9}$$

$$\ln y = \ln \beta_0 + \beta_1 x + \ln \epsilon \tag{10}$$

$$y' = \beta_0' + \beta_1 x + \epsilon' \tag{11}$$

식 (9)~(10)에서, (n , MSE) 데이터 쌍이(x , y)로 대입된다. 그리고, 식 (11)에서; 1) $y' = \ln y$; 2) $\beta_0' = \ln \beta_0$; 및 3) $\epsilon' = \ln \epsilon$ 과 같이 치환된 것이다. <그림 6.a>의 경우를 예로 들면, 아래와 같은 세 가지 회귀모형을 추정해 평활곡선이 생성되었다.

$$\begin{aligned} \text{PT_r: } \hat{y} &= 0.57447e^{-0.21687x}, R^2 = 86.6\% \\ \text{DT_r: } \hat{y} &= 0.28417e^{-0.22017x}, R^2 = 88.4\% \\ \text{MF_r: } \hat{y} &= 1.08919e^{-0.01099x} - 1, R^2 = 57.1\% \end{aligned}$$

표 5. Metamodels: summary

Skewness	Method	$z = +2$		$z = +3$		$z = +4$	
		Metamodel	R^2	Metamodel	R^2	Metamodel	R^2
$\lambda_i = 0.5$	PT_r	$0.21325e^{-0.18372x}$	0.867	$0.97716e^{-0.10051x}$	0.506	$4.46170e^{-0.13262x}$	0.971
	DT_r	n/a		n/a		n/a	
	MF_r	$1.04096e^{-0.00535x} - 1$	0.357	$0.41395e^{-0.23499x}$	0.964	$1.09161e^{-0.01123x} - 1$	0.543
$\lambda_i = 5.0$	PT_r	$0.57447e^{-0.21687x}$	0.866	$2.91888e^{-0.16401x}$	0.641	$21.22989e^{-0.12091x}$	0.901
	DT_r	$0.28417e^{-0.22017x}$	0.884	$1.36615e^{-0.21737x}$	0.875	$5.56502e^{-0.18773x}$	0.838
	MF_r	$1.08919e^{-0.01099x} - 1$	0.571	$1.16036e^{-0.01864x} - 1$	0.714	$1.18530e^{-0.02159x} - 1$	0.711
$\lambda_i = 10.0$	PT_r	$1.18187e^{-0.21578x}$	0.851	$4.57040e^{-0.15778x}$	0.645	$37.76943e^{-0.11185x}$	0.908
	DT_r	$0.64604e^{-0.26288x}$	0.898	$2.44734e^{-0.26613x}$	0.879	$7.82560e^{-0.23769x}$	0.821
	MF_r	$1.24686e^{-0.02604x} - 1$	0.739	$0.45957e^{-0.19640x}$	0.992	$1.24133e^{-0.02626x} - 1$	0.729

실제로, 회귀 추정식은; 1) $x = 0.01n$; 그리고 2) MF_r의 경우, MSE=0인 경우가 있어 식 (9)의 $\ln y$ 변환이 불가능하여, 식 (5)와 유사한 개념으로써, $y = \text{MSE} + 1$ 이 대입되어 추정되었음을 밝힌다. MF_r을 제외한 나머지 두 가지 경우, 결정계수(coefficient of determination) R^2 값이 약 90%이다. 단, 이러한 결정계수 값은, n 이 갖는 지수함수 형태의 MSE 감소추세를 파악하기 위한 지표일 뿐, SBL 결정방법의 성능측도와는 무관하다. <그림 5>~<그림 7>에, 분석·제시된 metamodel 결과가 <표 5>에 요약되어 있다.

Poisson 분포로써 확률표집된 dataset에 대한, 시뮬레이션 결과를 요약하면; 1) 표본크기에 상관없이, Poisson 모형을 가정한 후, 데이터를 모형에 fitting시켜 구한 MSE가 가장 작다; 2) 비대칭도가 심할 때, 데이터 변환을 이용한 SBL 결정이 불가능한 경우가 발생한다; 3) z 값이 클수록, 표본크기에 대한 MSE 감소 효과가 크다; 그리고 4) 전반적으로, 표본크기가 증가할수록 MSE는 지수함수 형태의 감소 추세를 보인다. 실험설계 조건에 따라 차이는 있으나, 약 $n=400$ 에서의 MSE는, $n=100$ 일 때의 MSE와 비교할 때, 상당부분 감소한 것을 시각적으로 확인할 수 있다.

4. 사례분석

실제 bin 데이터를 이용한 사례분석을 실시한다. 임의의 제품에 대해, 분석기간 2개월 간 생산물량 중 wafer 1909매를 수집해; 1) 전기 한달 간 생산된 400매 wafer의 bin 데이터를 이용해 SBL을 결정하고; 2) 후기 한달 간 생산된 나머지 1509매 wafer의 bin 데이터에 SBL을 적용하여 비정상 wafer 판정비율을 검토한다. 검토대상 bin 항목은 세 가지(즉, $i=25, 19, 20$)이다. <표 6>에 세 개 panel로 분류된 사례분석 결과가 요약되어 있다.

<표 6.a>에 제시된 값은, 각 bin별로 Box-plot의 $Q3 + 3.0IQR$

(interquartile range, 사분위 간 범위) 값을 초과하는 극단 이상치 (extreme outlier) 데이터를 제거한 후, 산정된 수치이다. $\hat{\tau}_i$ 을 이용한 변환 데이터의 SBL(‘Transformed’)와 이를 역변환(‘Inverse-transformed’)한 SBL이 함께 제시되어 있다. 단, $i=25$ 의 경우, 데이터 변환은 가능하나, 변환 데이터의 SBL이 음수여서, 역변환이 불가능하다. Poisson 모형으로 fitting시킨 경우에는, 제시된 $\hat{\lambda}_i$ 값을 이용해 SBL이 결정되었다.

<표 6.b>와 <표 6.c>의 결과처럼, 상대적으로 잡음이 심한 실제 bin 데이터의 경우, 단순히 Poisson 모형으로 데이터를 fitting시켜 도출한 SBL은, 나머지 방법에 비해 상대적으로 더 많은 비정상 bin 데이터 판정을 한다. 반면, 데이터 변환 및 백분위수에 근거한 방법은, 주어진 sigma 수준에 상응하는 이론적 비정상 bin 데이터 판정비율에 상대적으로 더 근접한 판정 비율을 갖는다. 백분위수에 근거한 방법은, 비대칭도가 다른 세 가지 bin 모두에서 SBL 결정이 수행될 수 있었다. 한편, <표 6.c>의 비정상 wafer 판정비율이 <표 1>의 이론치보다 상대적으로 큰 것은, 실제 bin 데이터가 갖는 잡음의 정도와 시간에 따

른 공정의 동적 변동 그리고 이상치 데이터에 의한 영향일 것으로 사료된다. 본 사례분석 제품의 경우, 30개 이상의 mask layer 각각에 대해, ‘frontend’ 기본 단위공정과 검사·측정이 반복되므로, 공정변동과 이상치 발생이 가능하다. 다른 연구사례에서도, SBL sigma 수준 $z=+4$ 적용 시, 보통 wafer scrap율 2% 이상인 경우가 보고된 바 있다(Illyes and Baglee, 1992).

카이제곱(chi square) 통계량을 이용한 사례분석 데이터의 Poisson 분포에 대한 적합도 검정(goodness of fit test) 결과가 <표 7>에 요약되어 있다. $i=25$ 인 bin 데이터만이 유의수준 $\alpha=0.05$ 로서 Poisson 분포를 따른다고 판단할 수 있다. 이 경우, 모형으로 fitting시켜 구한 SBL의 비정상 wafer 판정비율과 백분위수에 근거해 도출된 비율이, 상대적으로 나머지 두 개 bin 항목과 비교할 때, 근사한 수준임을 <표 6.c>에서 확인할 수 있다. <표 7>의 각 bin별 분석 데이터는, <표 6>의 극단 이상치 데이터 제거 후, SBL 결정 시에 사용된 표본과 동일하다. 한편, 카이제곱 통계량 산정 시, 기대빈도수(expected frequency)가 5 이상이 되도록 계급(class) 수가 설정되었다.

표 6. A case study: SBL and disposition summary

(a)

SBL (based on 400 wafers)

i	Q3+3.0IQR	$\hat{\lambda}_i$	$\hat{\tau}_i$	PT			DT			DT			MF		
				z=+2	z=+3	z=+4	Inverse-transformed			Transformed			z=+2	z=+3	z=+4
							z=+2	z=+3	z=+4	z=+2	z=+3	z=+4			
25	4	0.519	-3.034	3	4	4	n/a	n/a	n/a	-0.21906	-0.65804	-1.09703	2	4	5
19	50	19.228	0.225	44	49	50	38	52	71	2.26318	2.43517	2.60715	28	34	39
20	30	8.970	0.113	18	25	27	22	34	53	1.41559	1.49090	1.56622	15	19	23

(b)

Disposition counts of abnormal wafers (among 1509 wafers)

i	PT			DT			MF		
	z=+2	z=+3	z=+4	z=+2	z=+3	z=+4	z=+2	z=+3	z=+4
25	5	3	3	n/a	n/a	n/a	18	3	1
19	31	26	25	52	24	17	121	70	48
20	231	105	102	121	86	55	378	180	113

(c)

Disposition proportions of abnormal wafers (among 1509 wafers)

i	PT			DT			MF		
	z=+2	z=+3	z=+4	z=+2	z=+3	z=+4	z=+2	z=+3	z=+4
25	0.00331	0.00199	0.00199	n/a	n/a	n/a	0.01193	0.00199	0.00066
19	0.02054	0.01723	0.01657	0.03446	0.01590	0.01127	0.08019	0.04639	0.03181
20	0.15308	0.06958	0.06759	0.08019	0.05699	0.03645	0.25050	0.11928	0.07488

표 7. A case study: testing for goodness of fit

<i>i</i>	No. of data	No. of classes	Degrees of freedom	Significance level	Critical value	Chi square test statistic	p-value
		<i>k</i>	d.f.	α	$\chi^2_{\alpha, d.f.}$	χ^2_0	
25	397	4	2	0.01	9.210	5.543	0.063
				0.05	5.991		
				0.10	4.605		
19	390	20	18	0.01	34.805	320.892	0.000
				0.05	28.869		
				0.10	25.989		
20	399	14	12	0.01	26.217	211.106	0.000
				0.05	21.026		
				0.10	18.549		

5. 결론

3장의 시뮬레이션 분석에서는, Poisson 분포로 확률표집된 데이터를 이용하여, Poisson 모형에 데이터를 fitting하는 방법을 기준으로, 백분위수 및 데이터 변환에 근거한 방법을 비교·평가하였다. 인공적으로 생성된 데이터처럼, 잡음이 상대적으로 작을 경우, 데이터를 선택된 확률분포 모형에 fitting시켜 SBL이 결정될 수 있다. SBL 결정을 위한 효과적인 표본크기는, 대략 최소 wafer 400매 정도가 적정한 것으로 판단된다.

반면, 4장의 사례분석처럼, 잡음이 상대적으로 심한 실제 bin 데이터의 경우에는, 단순히 임의의 확률분포 모형을 가정한 후, 모형 fitting에 의해 SBL을 결정하기보다는, 백분위수 혹은 데이터 변환에 근거한 SBL 결정이 선호될 수 있다. 단, 비대칭도가 심한 bin 데이터의 경우, 데이터 변환에 의한 SBL 결정은 주의가 필요하다. 또는, 분석대상 bin 데이터에 대한 선택된 분포로의 적합도 검정을 선행한 후, 그 결과에 따라 SBL 결정방법을 취사선택할 수도 있다고 판단된다. 본 연구내용에 기초해, <그림 8>과 같이 SBL 결정방법 선정절차가 제안된다. 즉; 1) $n \geq 400$ 인 bin 데이터의 적합도 검정결과를 확인한 후; 2) 데이터가 특정 확률분포에 적합하면, MF 방법으로 SBL을 결정한다; 반면 3) MF 방법에 의한 SBL 결정이 적절치 않을 경우, 데이터 분포의 비대칭도를 고려해 PT 혹은 DT 방법을 선택할 수 있다. 한편, 현재 구축되어 운용중인 SBL 결정절차를 요약하면 아래와 같다.

- 첫째, 특정 확률분포를 가정한 MF 방법으로 SBL을 결정한다(SBL^1).
- 둘째, 제품별 담당 엔지니어가 SBL^1 과 bin 데이터 분포를 확인한다.

- 셋째, SBL^1 변경 요청 시, 아래 사항을 검토해 SBL이 재결정될 수 있다(SBL^2).

- 1) bin 데이터 이상치 선별·제거
- 2) PT 혹은 DT 방법에 의한 SBL 결정

이와 같은 SBL 결정절차를 통해, 엔지니어 변경 요청 없이 SBL^1 이 그대로 받아들여진 경우는, 총 4884개 SBL 중 53%인 2605개인 것으로 사례분석된 바 있다. 한편, 품질·신뢰성 강화를 위해 SBL^2 로 변경된 경우도 확인되었는데; 1) sigma 수준 조정; 2) 고정 SBL 적용; 그리고 3) 최근 SBL 유지 등으로 변경된 것으로 조사되었다.

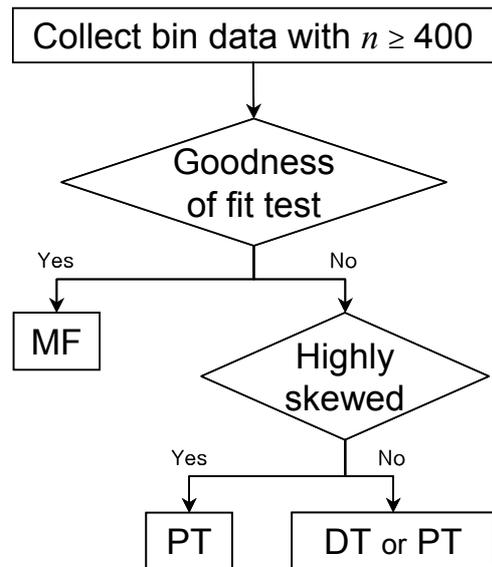


그림 8. A flow chart: the selection procedure for SBL determination methods.

본 연구와 관련하여, 향후; 1) bin 데이터의 이상치를 선별·제거하는 방법; 2) 좌변사향 수율 및 우변사향 bin 데이터를 동시에 분석·관리할 수 있는 SBL 모형 및 시스템 개발; 그리고 3) 비정상 wafer 판정비율과 이에 연계되는 run hold율을 동시에 적정 수준으로 유지할 수 있는 SBL 시스템 운영 algorithm 개발 등이 필요하다.

부록

- Frontend vs. backend: 반도체 IC 제품 제조공정은 크게, frontend와 backend 공정으로 구분할 수 있다. 단결정 실리콘 wafer가 fab-in되어 fab-out될 때까지의 fab 공정이 frontend 공정이고, backend 공정은 EDS부터 조립(assembly) 및 환경·전기적 package test(즉, final test) 그리고 신뢰성 검사를 포함한 나머지 제공정을 지칭한다. <그림 A.1>은 반도체 IC 제품 제조공정 흐름도이다.

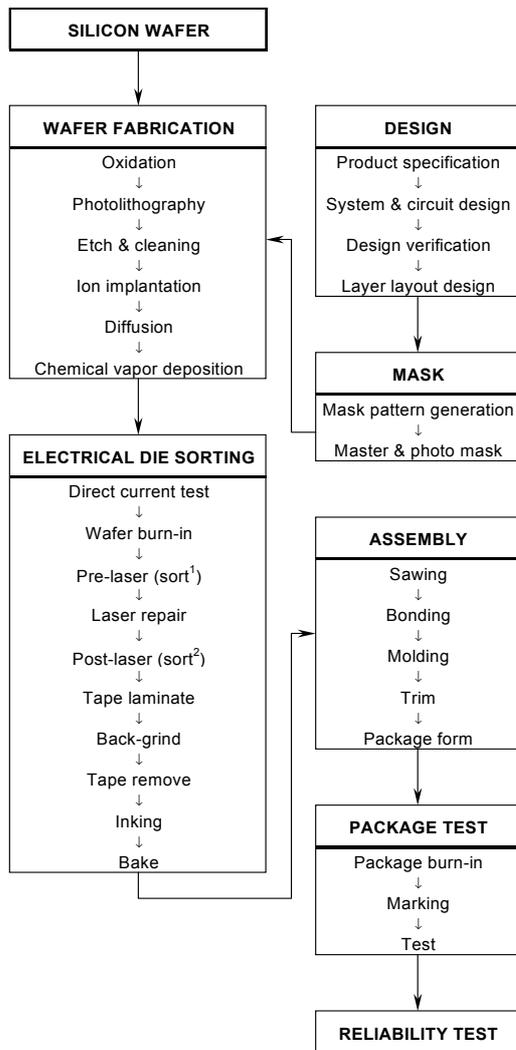


그림 A.1. A flow chart: a semiconductor IC device manufacturing process.

제품회로를 사양에 맞게 설계한 후, 보통 컴퓨터 시뮬레이션을 통해 설계가 검증되면, 설계회로를 다수의 ‘layer’로 분리한 후, layer별 원·사본 mask를 제작한다. 소위, ‘reentrant’되는 frontend 공정에서는, 단결정 실리콘 wafer 상에 반도체 IC 제품 소자가 형성되며; 1) 산화(oxidation); 2) 사진(photolithography); 3) 식각 및 세정(etch and cleaning); 4) 이온주입(ion implantation); 5) 확산(diffusion); 및 6) 화학증착(chemical vapor deposition) 등이 실시된다. 단, 여섯 가지 기본 단위공정이 항상 동일순서로 반복되는 것은 아니며, 형성되는 소자에 따라 취소선택된다.

EDS를 포함한 backend 공정에서는 주로, 제품검사와 package 조립 등이 실시된다. EDS에서는, wafer 상 전기적 특성검사가 실시되는데; 1) 소자에 대한 기본적인 작동 여부만을 검사하는 Test Element Group; TEG)의 direct current test; 및 2) wafer 상 die에 과도한 부하를 가해, 초기 불량제품 제거를 위한 wafer burn-in 이 실시된다. 이후, 순차적으로 설정된 bin 체계에 의거하여 보통 두 차례 분류(sorting)된다. ‘sort¹’ 결과 memory array 소자 등이 repairable한 die는, laser repair를 거쳐 ‘sort²’된 후, 최종 EDS wafer 수율이 산출된다. 다음, sawing을 위해 wafer 뒷면을 연마해야 하는데, 이 때에는 wafer 보호를 위하여 tape을 부·탈착한다. 검사결과, wafer 상 양·불량 die를 육안으로 식별하기 위해 불량 die를 inking하고, ink 및 wafer 내 수분을 oven에서 bake한다. 조립공정은 die와 package 보조부품을 결합하여, die를 외부 환경으로부터 보호함과 동시에 die가 외부장비와 연결될 수 있게 한다. 조립공정부터는, frontend 공정에서 사용되는 물류의 기본단위인 ‘run’ 대신, ‘lot’으로 물류를 관리하게 되는데, 그 이유는 조립후 die단위로 package되면, 무게·부피 증가로 인한 운반 및 설비 취급의 어려움으로 run 1개(보통, fab에서 최고 wafer 25매 용량의 run 또는 carrier box)가 다수의 lot으로 나누어지기 때문이다.

Wafer 상 die를 절단선을 따라 분리하는 sawing 후, die pad와 leadframe tip을 금 혹은 알루미늄 세션으로 접합시켜 주는 과정이 bonding이다. 다음, molding을 통해 외부의 물리적·화학적 변화로부터 조립 package를 보호하고, trim한다. 마지막으로, leadframe에서 필요한 만큼 lead를 절단하여 package form한다. Package된 제품을 다시 한번 burn-in하고, 동작제품만을 선별하여 제품정보를 marking한 후, 기능별·속도별 package test를 실시한다. 제품출하를 위한 입고에 앞서, 제품의 수명·사용환경·기초특성 항목에 대한 sampling 신뢰성 검사가 실시된다.

참고문헌

Cheong, K., Kim, J., Son, H., Kim, K., Park, Y. and Yun, J. (1996), Studies on Construction of SBL System and Improvement for Process of Testing Semiconductor, *Korean Semiconductor Technical Journal*, 34, 1405-1418.
 Cunningham, J. A. (1990), The Use and Evaluation of Yield Models in Integrated Circuit Manufacturing, *IEEE Transactions on Semiconductor Manufacturing*, 3(2), 60-71.

- Illyes, S. and Baglee, D. A. G. (1992), Statistical Bin Limits-an Approach to Wafer Disposition in IC Fabrication, *IEEE Transactions on Semiconductor Manufacturing*, 5(1), 59-61.
- Law, A. M. and Kelton, W. D. (1991), *Simulation Modeling and Analysis*, 2nd edn, McGraw-Hill, New York.
- Montgomery, D. C. (1997), *Design and Analysis of Experiments*, 4th edn, Wiley, New York.
- Montgomery, D. C. (2001), *Introduction to Statistical Quality Control*, 4th edn, Wiley, New York.
- Montgomery, D. C. and Peck, E. A. (1992), *Introduction to Linear Regression Analysis*, 2nd edn, Wiley, New York.
- Myers, R. H. and Montgomery, D. C. (1995), *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, Wiley, New York.
- Park, K., Lee, J., Kim, K., Kim, B., Lee, Y. and Ban, K. (2000), Development of SBL System for Effectively Detecting Process Variation and Continuous Process Improvement, *Technical Report*, No. GN-00101442, Device Solution Network, Samsung Electronics Co., Ltd.



박성민

건국대학교 산업공학과 학사
 고려대학교 산업공학과 석사
 미국 Arizona State University 산업공학과 박사
 현재: 삼성전자 반도체총괄 DSN, System LSI,
 책임연구원
 관심분야: 시뮬레이션, 응용통계



김영식

인하대학교 전자공학과 학사
 현재: 삼성전자 반도체총괄 DSN, System LSI,
 수석연구원
 관심분야: 시뮬레이션, 스케줄링