KIPS Tr. Software and Data Eng.
Vol.4, No.2 pp.65~70 pISSN: 2287-5905

http://dx.doi.org/10.3745/KTSDE.2015.4.2.65

Fault-Causing Process and Equipment Analysis of PCB Manufacturing Lines Using Data Mining Techniques

Hyun Sik Sim† · Chang Ouk Kim††

ABSTRACT

In the PCB(Printed Circuit Board) manufacturing industry, the yield is an important management factor because it affects the product cost and quality significantly. In real situation, it is very hard to ensure a high yield in a manufacturing shop because products called chips are made through hundreds of nano-scale manufacturing processes. Therefore, in order to improve the yield, it is necessary to analyze main fault process and equipment that cause low PCB yield. This paper proposes a systematic approach to discover fault-causing processes and equipment by using a logistic regression and a stepwise variable selection procedure. We tested our approach with lot trace records of real work-site. A lot trace record consists of the equipment sequence that the lot passed through and the number of faults for each fault type in the lot. We demonstrated that the test results reflected the real situation of a PCB manufacturing line.

Keywords : PCB Manufacturing Process, Fault-Causing Process and Equipment Analysis, Logistic Regression, Nano-scale Process

데이터마이닝 기법을 이용한 PCB 제조라인의 불량 혐의 공정 및 설비 분석

심 현 식† · 김 창 욱††

요 약

PCB(Printed Circuit Board) 제조공정에서의 수율은 제품의 원가와 품질을 결정하는 중요한 관리 요인이다. PCB 제조공정은 일반적으로 많은 단계의 미세공정을 거쳐서 제품인 칩(Chip)이 생산되기 때문에 높은 수율을 보장하기가 현실적으로 어렵다. 제품의 수율을 향상시키기 위해서는 저수율의 원인이 되는 불량요인을 분석하고, 불량요인에 영향을 미치는 중요공정 및 설비를 찾아서 관리해야 한다. 본 연구는 로지스틱 회귀분석 및 변수선택법을 이용하여 혐의공정 및 설비를 찾는 방법을 제안하였다. 데이터는 실제 현장의 로트 데이터를 사용하였고, 각 로트는 진행한 설비 및 불량유형별 불량수를 갖고 있다. 또한 분석 결과는 실제 현장 확인을 통하여 수율에 미치는 영향을 확인하였다.

키워드 : PCB 제조공정, 혐의공정 및 설비 분석, 로지스틱 회귀분석, 미세공정

1. 서 론

PCB는 전자부품을 탑재하여 부품과 부품 사이 혹은 신호선과 신호선 사이를 구리로 된 회로를 통해 전기적으로 접속시켜주는 회로 기판이다. PCB는 TV와 같은 일반 전자제품 뿐 아니라 핸드폰이나 태블릿 컴퓨터와 같은 정밀 기기에까지 광범위하게 사용되고 있다. 단순한 기능을 요구하는 일반 전자제품은 단순한 구조의 PCB만으로 생산이 가능하지만 최근 생산되는 스마트폰 등 정밀 기기는 고기능화 및 미세화에 따라 회로 구조가 더 복잡한 PCB의 생산을 위해 다층, 고기능 PCB 기판 구조의 사용이 불가피하다. 이러한 PCB는 결합(Package) 방식에 따라 Leaded Pack-age, BGA(Ball Grid Array), FCBGA(Flip Chip BGA) 방식 등으로 나눠진다.

복잡해진 기판 구조는 제품 생산 비용을 증가시킬 뿐만 아니라 제품의 수율을 감소시켜 기업의 제조경쟁력 확보에 어려움을 준다. 높은 수율과 품질을 유지하고 불량률을 줄이기 위해서는 설비의 안정적인 운영과 보전, 그리고 제한된 비용과 인력 안에서 최선의 관리 효율을 이끌어낼 수 있는 주요 공정 및 설비 파악이 매우 중요하다.

† 정 회 원 : 연세대학교 정경대학 외래교수
†† 비 회 원 : 연세대학교 정보산업공학과 교수
Manuscript Received : May 28, 2014
First Revision : September 16, 2014
Accepted : September 25, 2014
* Corresponding Author : Hyun Sik Sim(hssim@yonsei.ac.kr)

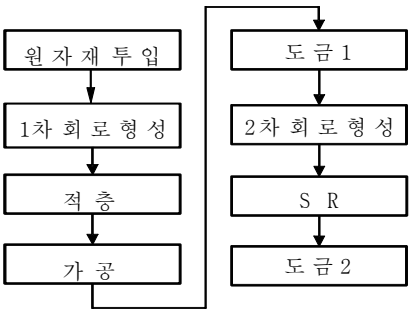


Fig. 1. PCB Process Flow Chart

PCB 공정은 Fig. 1과 같이 원자재 투입, 1차 회로 형성, 적층, 가공, 도금1, 2차 회로 형성, SR(Solder Resist), 도금2 공정으로 구성된다[1]. 가장 먼저 원자재인 기판을 제품 종류별로 재단하여 공정에 투입한다. 1차 회로 형성 과정에서 원자재 기판에 전도층을 위한 표면도금을 형성하고, 감광 레지스트를 밀착 후 자외선 노광 및 필요한 패턴을 현상하고 에칭, 박리하여 필요한 회로를 형성한다. 적층은 회로형성이 완료된 기판에 동박을 추가해 다층 기판을 형성하는 과정이다. 가공은 적층이 완료된 다층 기판에 상하, 내외로 전기적 연결을 위한 구멍을 만드는 공정이다. 가공이 끝난 후 회로 연결을 위한 도금1(동도금) 공정을 진행하여 도금을 증착한다. 2차 회로 형성은 현상, 에칭, 박리 공정을 거쳐 적층 공정에서 새로 형성된 층에 회로를 구성한다. SR 공정은 제품의 표면 회로를 보호하기 위해 절연 잉크를 도포하는 과정이다. 마지막으로 솔더볼의 전기 전도율 향상과 내부식성 증가를 위한 도금2(금도금) 공정을 거친다.

제조공정에서는 오래전부터 제품의 가공이 끝난 후 기판의 회로검사 또는 도금두께나 선폭 등을 계측하여 불량 여부를 확인하는 통계적 공정관리(Statistical Process Control) 기법을 도입하여 품질관리를 수행해왔다. 특히 1차 회로 형성 후 회로선의 연결 여부와 회로의 상태 등을 검사하여 수율을 산출한다. 1차 회로공정의 불량은 불량품의 후공정 진행 시 더 큰 공정 원가가 발생하여 제품의 원가를 가중시키는 중요한 요인으로 작용하기 때문에, 불량을 최소화하고 수율을 극대화하는 것이 절대적으로 필요하다. 특히 검사공정에서 저수율의 원인이 되는 불량이 어떤 공정의 어떤 설비에서 많이 발생하였는지를 정확히 찾아서 관리하는 것이 중요하다.

그러나 검사공정의 주 불량요인이 어느 공정, 어느 설비에서 발생했는지 찾는 것은 매우 어려운 일이다. PCB 제조라인에서 발생하는 설비변수는 수백 개에 달하지만 설비성능, 분석자원, 시간을 고려하면 모든 설비변수를 추출하여 분석하는 것은 불가능해진다[2]. 제조경쟁력을 높이기 위해서는 이러한 설비변수 데이터가 없는 상태에서 제품이 설비를 거쳐간 경로를 분석하여 혐의공정 및 설비를 찾아서 관리하는 것이 가장 효과적이다[3].

지금까지 로지스틱 회귀분석을 이용한 연구로는 단변양 및 다변양 분석기법을 이용하여 금융, 마케팅, 고객 만족도 조사 등 여러 분야에 대한 연구들이 있다[4-6].

그러나 PCB 산업의 제조공정에서 수율에 영향을 미치는 혐의공정 및 불량설비를 찾기 위한 연구는 현재까지 없었다. 다만 주요설비에 대하여 실시간으로 센서 데이터를 모니터링하여 이상을 감지하는 FDC(Fault Detection and Classification)에 대한 연구는 다수 존재한다. Cherry *et al.*[7]는 주성분 분석을 활용한 다변량 분석 기법을 post-lithography 계측 공정에 적용하여 18개 중요 계측 파라미터를 관리하였다. Yan[8]은 36개의 파라미터로 구성된 aluminum Gate CMOS(Complementary Metal Oxide Semiconductor) 공정의 제어 데이터에 주성분 분석을 적용하였다. Goodlin *et al.*[9]은 어느 관리도에서든 불량이 감지됐을 때, 불량의 유형별로 이상 원인을 즉시 식별할 수 있는 특화된 관리도(Type-Specific Control Chart)를 제안하였다. 이를 6개 불량 유형, 19개 공정 파라미터를 가진 에칭 공정에 적용하였지만 여전히 더 많은 연구가 필요하다고 언급하였다. Ma *et al.*[10]는 반도체의 공정 중 하나인 CVD(Chemical Vapor Deposition)공정에 적용하여 품질에 영향을 미치는 중요 파라미터를 관리하고 수율을 증가시켰다. CVD의 공정을 이루는 3개의 챔버(Chamber) 데이터에 대해서 Wilcoxon rank sum test, Clustering, Regression 방법을 이용하여 31개의 중요 파라미터를 찾아내고, 새로운 규격 한계선을 설정하고 관리하여 수율을 개선하였다.

이와 같은 FDC는 설비 센서 데이터를 실시간으로 모니터링하여 이상을 감지하는 방법인 데 비하여, 본 연구는 제조 라인의 전체 공정으로부터 축적되는 공정 데이터를 이용하여 최종 수율에 영향을 가장 많이 미치는 공정 및 설비를 찾는 연구이다.

본 연구의 목적은 PCB 제조공정의 수율에 영향을 미치는 주 불량요인을 분석하고, 공정별로 설비 분석을 진행하여 주요 불량요인별로 가장 큰 영향을 미치는 혐의공정 및 혐의설비를 찾아 선택된 설비의 집중관리를 통해 불량요인을 개선하고 최종적으로 제품의 수율을 향상시키는 데 있다. 이를 위해 본 연구에서는 중요 공정 및 설비의 선택을 위해 1차 회로 공정의 수율에 영향을 미치는 불량요인을 선별하고, 선택된 인자별 경로설비(범주형 데이터) 분석을 로지스틱 회귀분석[11]을 적용해 적합도를 측정하고 선택된 중요 공정 및 설비가 주 불량요인에 미치는 영향을 확인했다.

본 논문은 다음과 같이 구성되어있다. 2절에서는 대상제품 및 공정을 선정한다. 3절에서는 불량요인별 중요 공정 및 설비를 선택하기 위해 사용된 분석 모델을 설명한다. 4절에서는 실험 및 데이터 분석 과정을 설명하고, 5절에서는 제안된 분석 모델로 선택된 중요 공정, 설비를 확인한다. 6절에서 연구의 결론을 정리하고 향후 공정 및 설비 데이터와 다양한 분석 가능성에 대해 논한다.

2. 분석 대상

2.1 제품 및 공정 선정

PCB 공정에서의 수율은 제품의 품질과 원가에 영향을

미치는 중요한 인자로 제조경쟁력을 평가하는 수단으로 사용되고 있다. 또한 한 개의 라인에서 여러 종류의 제품을 생산해야 하고 수백 개의 단위공정 설비를 거치면서 수많은 변수에 따라 영향을 받게 된다. 해당 라인의 관리 상태, 즉 공정능력지수에 따라 균일한 중심값을 갖는 제품을 생산할 수 있고, 제품의 수율을 극대화할 수 있다. 개별 제품별 수율은 제품의 설계사양 및 공정의 난이도와 공정 및 설비변수에 따라 결정된다.

본 연구에서는 월별 생산량이 많은 제품 중에서 수율이 평균치보다 낮은 제품을 선정하였고, 선정된 제품의 해당 기간 중 생산된 로트들의 회로공정 검사 결과를 분석하여 수율을 산출하였다. 수율은 각 공정이 끝난 후 수율을 계산하여 관리하고 있으나, 1차 회로 형성 공정에서 불량이 발생할 시, 후공정에서 공정원가가 증가하고 불량원인을 찾아서 개선하는 것이 어려워지므로 1차 회로 형성 공정의 수율 관리가 중요하다. 따라서 개선목표를 1차 회로 형성 공정의 수율을 향상시키는 것으로 선정하였다.

2.2 주 불량요인 분석

회로 공정 완료 후 검사공정에서 수백 개의 검사항목을 통하여 양품을 선별해낸다. 저수율 발생 시 각 항목별로 불량수를 집계하여 주 불량요인을 선정한다. 선정된 제품의 해당기간 내 생산된 로트들의 불량요인을 분석한 결과, 주 불량요인 및 불량 비율은 Fig. 2와 같다[1]. 수율에 가장 많은 영향을 미치는 주 불량요인은 Short(B), Defect(D), Open(F) 등 6개를 선정하였다. 수십 개의 불량요인 중 A~F를 제외한 불량요인은 전체 불량 중 1% 미만이므로 본 연구에서는 제외한다.

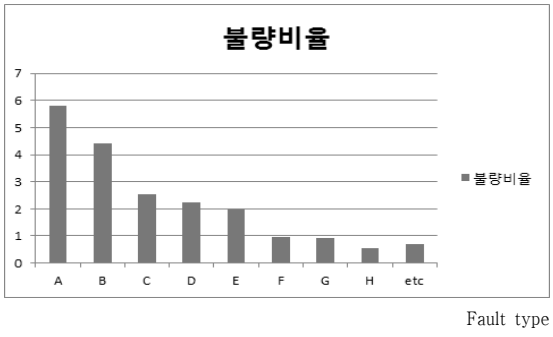


Fig. 2. Major Fault Type Graph

본 연구에서는 1차 회로 형성 공정을 대상으로 회로 형성 후 검사공정의 수율 및 수율에 영향을 미치는 불량요인을 분석하고, 불량요인별로 가장 영향을 많이 주는 혐의공정 및 설비를 분석하였다. 불량요인은 앞에서 선정한 6개이다. Table 1에서 보듯이 대상 공정은 1차 회로 형성에 영향을 미치는 10개 공정을 선정하였고, 공정별로는 동일한 가공을 수행하는 설비가 최소 2개에서 최대 5개 배치되어있다. 이 표에서 P는 공정(Process)의 약자이고 E는 설비(Equipment)의 약자이다.

Table 1. Processes and Equipment

공정 및 진행설비									
P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
E11	E21	E31	E41	E51	E61	E71	E81	E91	E101
E12	E22	E32	E42	E52	E62	E72	E82	E92	E102
E13		E33	E43	E53	E63	E73	E83	E93	
		E34	E44			E74			
		E35				E75			

* P3(화학동), P5(노광), P6(현상), P7(도금), P9(박리).

PCB 제조공정의 작업은 로트 단위로 진행된다. 하나의 로트는 수십 개의 패널(Panel)로 구성되며, 하나의 패널에는 수백 개의 칩이 있다. 각 로트는 설비 경로를 따라 진행되며 1차 회로 형성 공정이 끝나면 로트에서 몇 개의 패널을 임의로 선정하여 칩을 검사한다. 따라서 검사 공정이 완료되면 로트의 설비 경로와 불량요인별 불량 칩의 개수가 산출된다. 본 연구에서는 불량요인별로 총 검사 수량분의 불량 개수를 계산하여 불량률로 변환하였고, 설비경로는 하나의 로트가 특정 설비를 거쳐 갔으면 1, 아니면 0값을 가지도록 하였다.

3. 로지스틱 회귀분석

한 개의 종속변수와 여러 개의 독립변수 간의 상호관련성을 분석할 때 선형 회귀모형을 적용한다. 일반적으로 선형 회귀모형에서 종속변수는 연속적인 값을 가지는 변수로 가정하지만, 독립변수에 따른 양품, 불량 여부를 판단하거나, 종속변수가 0 혹은 1로 표시되는 이항적인 사건을 다루는 경우에는 기존의 선형 회귀분석으로는 독립변수와 종속변수의 관계를 충분히 설명해주지 못한다[12]. 이러한 경우 종속변수를 양적으로 설명하는 것이 아니라 종속변수가 발생하게 되는 확률을 설명하도록 모델을 재구성해야 한다, 일반적으로 회귀분석에서의 모형은 독립변수$(x_1,x_2,\cdots,x_k)$와 종속변수 y의 선형식으로 표현 가능하다.

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_mx_m \quad (1)$$

그러나 불량의 발생 유무와 같은 이분형 종속변수인 경우를 다루기 위해서는 y값이 확률값(P)이고 0에서 1 사이의 값을 가져야 한다. Equation (1)의 선형 회귀식에서는 좌변인 y나 우변 모두 $(-\infty,+\infty)$값을 가질 수 있어서 범위가 동일한데 y 대신에 단순히 확률값 P를 대입하면 독립변수와 종속변수 사이의 범위가 서로 일치하지 못하는 경우가 발생한다. 따라서 이를 해결하고자 아래와 같은 로지스틱 함수를 사용하여 로짓 변환(Logit Transformation)을 통해 범위를 통일시킨 선형 회귀식을 만들 수 있다. 참고로 로짓 변

1) 정식 불량항목 이름 및 설명은 보안상 본 논문에서는 밝히지 못함.

환뿐만 아니라 Gompertz, Probit 등의 모델도 많이 사용되고 있다[13].

$$P=\frac{\exp(\beta_0+\beta_1x_1+\cdots+\beta_mx_m)}{1+\exp(\beta_0+\beta_1x_1+\cdots+\beta_mx_m)} \quad (2)$$

Equation (2)의 로지스틱 함수는 사건이 일어날 확률과 일어나지 않을 확률 사이의 비율을 나타내는 오즈(Odds)의 개념을 활용하여 Equation (3)처럼 선형식으로 표현할 수 있다[14].

$$\ln\left(\frac{P}{1-P}\right)=\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_mx_m \quad (3)$$

오즈(Odds)란 사건이 발생하는 상대적인 척도이며, 오즈 비율(Odds ratio)은 독립변수가 한 단위 증가할 때 종속변수의 오즈가 변하는 정도를 의미한다[15].

계수 β를 추정하는 방법은 기존 선형 회귀분석에서는 최소자승법(Least Sqaures Method)을 통하여 잔차의 제곱합을 최소화하는 회귀계수를 사용하지만, 로지스틱 회귀분석에서는 사건의 발생 가능성을 최대로 하는 최대우도 추정법(Maximum Likelihood Method)을 이용하여 계수를 추정한다[16].

본 연구에서는 PCB 공정에서 높은 수율 확보를 위해 로지스틱 회귀분석을 실시하였는데, 종속변수는 불량 유무를, 독립변수는 공정의 설비를 의미하고 로트가 해당 설비를 거쳐 가면 1 아니면 0을 갖는 이산형 변수로 변환하여 종속변수와 독립변수 사이의 상관성을 측정하였다. 이진변수의 로지스틱 회귀모형에서는 불량을 나타낼 확률 P가 0.5보다 크면 집단 1(불량)로 분류하고, 반면에 P가 0.5보다 작으면 집단 0(정상)으로 분류한다. 그리고 마지막으로 추정한 로지스틱 회귀모형을 바탕으로 단계적 변수선택법(Stepwise Selection Method)을 사용하여 불량에 영향을 미치는 핵심 설비 요인을 발견한다[10, 17].

단계적 변수 선택법은 종속변수에 크게 영향을 주는 변수부터 하나씩 추가하고, 새로운 변수가 추가될 때마다 기 추가된 변수를 제거하거나, 또는 제거 시 이미 제거된 변수의 중요도가 높아져 추가될 수 있는지를 단계별로 검토하여 선택하는 방법이다[2, 18]. 본 실험에서는 변수별 p-value가 0.05 이하면서 Chi-square값이 큰 순으로 선택하였다.

4. 실험 및 데이터 분석

4.1 혐의공정 분석

본 연구에서는 상기모델에 따라 최적의 모형식을 통계적으로 만들고 독립변수별 유의성 검정을 통하여 p-value가 0.05 이하이면서, Chi-square값을 기준으로 값이 큰 변수를 선정하였다. 여기서 Chi-square값은 종속변수(불량률)에 독립변수(공정 또는 설비)가 미치는 영향 정도를 설명한다. Table 2는 불량요인 B(Short)에 영향을 미치는 공정을 내림차순으로 보여준다. 이 표의 결과에서 보듯이 불량요인 B(Short)에 영향을 많이 미치는 공정은 영향도가 큰 기준으로 볼 때 노광(P5), 현상(P6), 박리(P9) 공정의 순으로 중요한 공정이 선정되었다. 본 연구에서는 공정별 누적 영향도를 계산하여 전체 중 약 70%가 되는 3개 공정을 선정하였다(P8, P10은 P-Value가 0.05 이상임).

Table 2. Results of Fault Type B(Short)

공정	설비 대수	Chi-square	P-value	영향도 (%)	누적 영향도
P5	**3**	**151.2**	**<0.0001**	**29.1**	**29.1%**
P6	**3**	**103.8**	**<0.0001**	**20.0**	**49.1%**
P9	**3**	**100.3**	**<0.0001**	**19.3**	**68.4%**
P4	4	48.7	<0.0001	9.4	77.8%
P3	5	37.9	<0.0001	7.3	85.1%
P7	5	36.9	<0.0001	7.1	92.2%
P1	3	34.8	<0.0001	6.7	98.9%
P2	2	4.59	<0.0320	0.9	99.8%

4.2 불량요인별 혐의공정

다른 불량요인에 대하여도 상기와 같은 방법으로 분석하여 영향도를 계산하고 혐의공정을 영향도가 큰 순서로 선정하였다. 분석 결과 각 불량요인별로 영향을 많이 미치는 공정은 Table 3과 같이 추출되었다.

Table 3. Results of Fault Process

	A	C	D	E	B	F
P1						
P2	○3					
P3	●1	●3	●1			
P4						
P5					○1	
P6		●2		●2	●2	●3
P7	●2		●3	●1		●2
P8						
P9			●2	●3	●3	●1
P10		○1		○3		

* ●: 불량요인에(3개 이상) 공통으로 영향을 미치는 공정
○: 특정 불량요인에만 영향을 미치는 공정
1, 2, 3: 불량 Type별로 누적 영향도가 큰 공정(순위)

각 변수별로 Chi-square값을 계산하여 p-value가 0.05 이하를 만족하면서 Chi-square값이 큰 값을 선정하였다(누적 영향도를 계산하여 70%가 되는 3개 공정을 선정함). 이 중에서 전체 불량요인에 공통으로 영향을 많이 미치는 공정은 화학동(P3), 현상(P6), 도금(P7), 박리(P9) 공정이 선정되었다. P1, P2, P4, P5, P8, P10은 거의 영향이 없거나 특정 불량인자에게만 영향을 미친 것으로 분석되었다. 따라서 전체 불량에 공통적으로 영향을 미치는 4개의 공정(P3, P6, P7,

P9)을 우선적으로 관리하는 것이 효과가 클 것으로 예상된다. 다음으로는 각 불량요인별로 혐의가 큰 공정 내에서 어떤 설비가 가장 영향을 크게 미쳤는지 혐의설비를 찾는 분석 작업을 진행했다.

4.3 혐의설비 분석

공정별로 혐의설비를 찾는 분석은 각 공정별로 2대 이상 설비로 작업된 공정을 대상으로 분석하였다. 4.2절에서 단계적 변수 선택법을 이용하여 선정된 공정에 대하여 최대우도 추정법을 이용하여 각 독립변수(설비)의 로지스틱 회귀식의 계수를 추정하였다. 이때 추정값이 작은 설비가 불량이 발생할 확률이 작다. 즉 같은 공정 내 여러 대의 설비가 있을 때 독립변수(설비)의 추정값이 작은 설비가 양호설비, 추정값이 큰 설비가 불량설비라고 할 수 있다. 4.2절에서 분석한 불량요인 B(Short)의 경우 혐의설비 분석 결과는 Table 4와 같다(각 공정별로 Reference 설비는 표에서 제외함).

Table 4. Fault Equipment Analysis (Fault type:B)

Analysis of Miximum Likelihood Estimate						
Para-meter		DF	Estimate	STD Error	Chi-Square	Pr> ChiSq
Inter-cept		1	-4.430	0.429	106.61	<.0001
P1	E11	1	-0.065	0.034	3.66	0.0557
	E12		0.113	0.022	25.94	<.0001
P2	E21	1	0.029	0.013	4.59	0.0320
P3	E31	1	0.046	0.041	1.26	0.2603
	E32	1	-0.245	0.059	16.95	<.0001
	E33	1	0.056	0.145	0.14	0.7009
	E34	1	0.155	0.048	10.42	0.0012
P4	E41	1	0.311	0.067	21.50	<.0001
	E42	1	0.192	0.078	6.28	0.0121
	E43	1	-0.119	0.044	7.20	0.0073
P5	E51	1	-0.178	0.019	80.67	<.0001
	E52	1	-0.090	0.030	8.98	0.0027
P6	E61	1	-0.351	0.034	100.40	<.0001
	E62	1	0.029	0.024	1.53	0.2153
P7	E71	1	0.029	0.044	0.45	0.5018
	E72	1	0.124	0.030	16.98	<.0001
	E73	1	-0.097	0.030	10.52	0.0012
	E74	1	-0.084	0.023	13.14	0.0003
P9	E91	1	0.017	0.086	0.03	0.8431
	E92	1	-0.162	0.047	11.84	0.0006

Table 4에서 구한 각 독립변수(설비)의 계수 추정값을 Equation (3)의 회귀방정식에 대입하면 Equation (4)를 얻을 수 있다.

$$\log\left(\frac{p}{1-p}\right) = -4.43 + 0.113 \cdot E12 + 0.029 \cdot E21 - 0.245 \cdot E32 + 0.155 \cdot E34 + 0.311 \cdot E41 + 0.192 \cdot E42 - 0.119 \cdot E43 - 0.178 \cdot E51 - 0.090 \cdot E52 - 0.351 \cdot E61 + 0.124 \cdot E72 - 0.097 \cdot E73 - 0.084 \cdot E74 - 0.162 \cdot E92 \quad (4)$$

4.4 검 증

본 절에서는 4.3절에서 분석한 결과를 토대로 불량요인별 영향이 큰 공정별로 양호설비, 불량설비를 구분하였다. 구체적으로 불량요인별로 4.2절에서 선정된 누적 영향도가 큰 3개 공정을 선정하고, 다시 공정별로 양호설비와 불량설비를 구분하였다. Table 5는 각 불량요인별로 혐의공정 및 설비를 찾은 결과(Table 3, Table 4)를 정리하여, 공정별로 영향도가 큰 설비를 추출하였다. 전체 불량요인에 공통적으로 영향을 미치는 공정은 화학동(P3), 현상(P6), 도금(P7), 박리(P9) 공정이고 각 공정별 양호설비와 불량설비는 Table 5의 내용과 같다. 이 표에서 화학동(P3) 공정의 경우 불량요인 A와 D(Defect)에서 설비 E32는 양호한 것으로 판명 났다. 즉 이 설비는 A와 D(Defect)에 영향을 미치지 않았다는 뜻이다. 같은 맥락으로 도금(P7) 공정에서 설비 E72는 불량요인 D(Defect), E, F(Open)에 영향을 미치지 않는다는 결과가 나왔다. 그러나 몇 가지 설비는 상충되는 결과가 나왔다. 예를 들어 박리(P9) 공정의 경우에 설비 E91은 불량요인 D(Defect), F(Open)에는 영향을 안 주었는데 요인 E에는 큰 영향을 주었다. 이런 현상은 불량요인의 특성이 서로 상반되기 때문이다. 예를 들어 PCB의 연결 두께를 크게 하면 단선 불량은 줄어들지만 합선 불량은 증가한다.

Table 5. Fault Processes and Equipment

불량요인	설비	혐의공정/설비					
		P2	P3	P5	P6	P7	P9
A	양호	E22	E32			E75	
	불량	E21	E33			E71	
B	양호			E51	E61		E92
	불량			E53	E63		E93
C	양호		E35		E62		
	불량		E31		E63		
D	양호		E32			E72	E91
	불량		E34			E75	E93
E	양호				E63	E72	E93
	불량				E61	E73	E91
F	양호				E63	E72	E91
	불량				E61	E73	E92

불량요인에 가장 많은 영향을 미치는 공정은 실제 제조현장에서 관리하는 공정인자(선폭, 도금두께 등) 데이터를 분석하여 실제로 해당공정 불량요인에 영향을 미쳤다는 사실을 확인하였다. 실제로 도금 공정에서 설비조건에 따라 공정인자값이 변하고 불량요인에 영향을 미치는 것을 이론적으로 확인하였다. 또한 각 공정별 양호설비와 불량설비는 직접 실험을 할 수는 없지만, 설비 조건이나 상태에 따라서 공정인자가 변하고 불량요인에 영향을 준다는 사실이 확인된 만큼 설비별로 설비변수 분석을 통한 검증 작업이 추가로 필요하다.

5. 결 론

본 연구의 목적은 현재 폭넓게 보급된 PCB 제품을 생산하는 공정의 수율에 영향을 미치는 공정 및 설비를 분석함으로써 PCB 제조라인의 수율 향상과 생산성 향상에 따른 기업 경쟁력을 확보하는 데 있다. 이를 위해 본 논문에서는 PCB 제조라인의 불량 데이터 및 설비변수 데이터를 분석하여 어떤 공정이 수율에 영향을 미치는지를 분석하고, 공정별로 제품을 가공한 설비를 분석하여 불량에 가장 많은 영향을 미치는 혐의설비가 무엇인지 발견하는 분석기법을 제안했다. 중요 인자로 분류된 공정 및 설비들은 현장 엔지니어의 의견을 수렴해 집중적으로 관리될 필요가 있으며, 추가로 설비변수와의 연계분석을 통하여 차후 더 발전된 요인 선택 연구를 진행할 가능성도 존재한다. 마지막으로 아직까지 PCB 생산 공정에서 설비 데이터를 수집하고 생산성 향상에 활용하는 연구 사례가 부족한 상황에서 실제 공정 환경을 고려한 다양하고 새로운 데이터 분석 기법의 개발이 요구된다.

References

[1] 김강희, 이병엽, “The PCB,” 북두, 2013.

[2] I. Guyon, A. Elisseeff, “An introduction to variable and feature selection,” *The Journal of Machine Learning Research3*, pp.1157-1182, 2003.

[3] P. A. Estevez, M. Tesmer, C. A. Perez, and J. M. Zurada, “Normalized mutual information feature selection,” *IEEE Transactions on Neural Networks,* Vol.20, No.2, pp.189-201, 2009.

[4] Z. S. Hua, Y. Wang, X. Xu, B. Zhang, and L. Liang, “Predicting corporate financial distress based on integration of support vector machine and logistic regression,” *Expert Systems with Applications,* Vol.33, No.2, pp.434-440, 2007.

[5] C. Lawson, D. C. Montgomery, “Logistic Regression Analysis of Customer Satisfication Data,” *Quality and Reliability Engineering International,* Vol.22, pp.971-984, 2006.

[6] D. Hosmer, S. Lameshow, “Applied Logistic Regression,” Wiley, 2000.

[7] G. A. Cherry, S. J. Qin, “Multiblock Principal Component Analysis Based on a Combined Index for Semiconductor Fault Detection and Diagnosis,” *IEEE Transactions on Semiconductor Manufacturing,* Vol.19, No.2, pp.159-172, 2006.

[8] L. Yan, “A PCA-based PCM Data Analyzing Method for Diagnosing Process Failures,” *IEEE Transactions on Semiconductor Manufacturing,* Vol.19, No.4, pp.404-410, 2006.

[9] B. E. Goodlin, D. S. Boning, H. H. Sawin, and B. M. Wise, “Simultaneous Fault Detection and Classification for Semiconductor Manufacturing Tools,” *Journal of the Electrochemical Society,* Vol.150, No.12, pp.778-784, 2003.

[10] M. D. Ma, D. H. Wong, S. S. Jang, and S. T. Tseng, “Fault detection based on statistical multivariate analysis and microarray visualization,” Industrial Informatics, *IEEE Transactions on,* Vol.6, No.1, pp.18-24, 2010.

[11] D. Montgomery, E. A. Peck, and G. Vining, “Introduction to Linear Regression Analysis,” 4th Edition, Wiley, 2007.

[12] R. D. Cook, “Detection of influential observations in linear regression,” *Technometrics,* Vol.19, pp.15-18, 1977.

[13] R. D. Berger, “Comparison of the Gompertz and Logistic Equations to Describe Plant Disease Progress,” *Phytopathology,* Vol.71, No.7, pp.716-719, 1981.

[14] George Y. Wong, William M. Mason, “The hierarchical Logistic Regression Model for Multilevel Analysis,” *Journal of the American Statistical Association,* Vol.80, No.391, pp.513-524, 1985.

[15] W. H. David, S. Lemeshow, “Applied Logistic Regression,” John Wiley, New York, 1989.

[16] A. Albert, J. A. Anderson, “On the existence of maximum likelihood estimates in logistic regression models,” *Biometrika,* Vol.71, No.1, pp.1-10, 1984.

[17] S. Galit, R. P. Nitin, and C. B. Peter, “Data mining for Business Intelligence,” Wiley, 2009.

[18] E. W. Ronald, H. M. Raymond, L. M. Sharon, and Y. Keying, “Probability & statistics for engineering and scientists,” 9th Edition, Pearson Press, 2011.

심 현 식

e-mail : hssim@yonsei.ac.kr

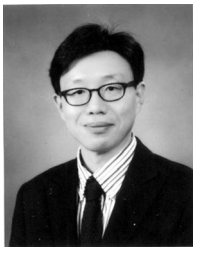
(전) 삼성전자반도체 그룹장

(전) 삼성전기 그룹장

2014년 연세대학교 정보산업공학과(공학박사)

현 재 연세대학교 정경대학 외래교수

관심분야 : 생산공정관리, 데이터 마이닝(제조), 정보시스템(생산정보, 경영정보)



김 창 욱

e-mail : kimco@yonsei.ac.kr

1996년 미국 Purdue대학교 산업공학(박사)

현 재 연세대학교 정보산업공학과 교수

현 재 연세대학교 공과대학 연구부학장

관심분야 : 제조업 데이터 마이닝, 에이전트 기반 시뮬레이션