

## 실시간 데이터를 이용한 64M DRAM s-Poly 식각공정에서의 웨이퍼 상태 예측

### Wafer State Prediction in 64M DRAM s-Poly Etching Process using Real-Time Data

°이석주, 차상엽, 최순혁, 우광방

연세대학교 전기공학과 (Tel : +82-2-361-2767; Fax : +82-2-362-7611; E-mail : sjyi@bubble.yonsei.ac.kr)

**Abstracts** For higher component density per chip, it is necessary to identify and control the semiconductor manufacturing process more stringently. Recently, neural networks have been identified as one of the most promising techniques for modeling and control of complicated processes such as plasma etching process. Since wafer states after each run using identical recipe may differ from each other, conventional neural network models utilizing input factors only cannot represent the actual state of process and equipment. In this paper, in addition to the input factors of the recipe, real-time tool data are utilized for modeling of 64M DRAM s-poly plasma etching process to reflect the actual state of process and equipment. For real-time tool data, we collect optical emission spectroscopy (OES) data. Through principal component analysis (PCA), we extract principal components from entire OES data. And then these principal components are included to input parameters of neural network model. Finally neural network model is trained using feed forward error back propagation (FFEBP) algorithm. As a results, simulation results exhibit good wafer state prediction capability after plasma etching process.

**Keywords** 64M DRAM, Etching Process, Wafer State, Prediction, Real-Time Data

#### 1. 서론

반도체 소자의 집적도가 증가하면서 반도체 생산공정 및 장비에 대한 보다 효율적이고 정확한 모델링 및 제어가 필요해지고 있다. 특히 16M DRAM에서 64M DRAM 나아가 1G DRAM으로의 급속한 고집적화는 보다 정밀한 제어를 필요로 한다. 공정의 모델링과 제어는 반도체 제조공정의 발전에 중요한 위치를 차지하며, 대부분의 공정들은 복수개의 변수들에 의해 영향을 받는 난해한 비선형적 특성을 보인다. 그중에서도 최근 반도체 제조공정에의 적용이 급증하고 있는 플라즈마 식각공정은 건식식각을 이용한 공정으로서 매우 복잡한 물리적, 전기적 특성을 지니고 있다. 이로 인해 플라즈마 식각공정에 대한 정밀하고 효율적인 모델링 및 제어가 어려운 실정이다. 지금까지는 플라즈마 식각공정에 대한 모델링 및 제어입력 (recipe) 탐색은 통계적 공정제어 (statistical process control) 기법 및 전문가의 경험에 의존하여 왔고 [1], 최근에는 신경회로망 기법을 이용한 플라즈마 식각공정의 모델링이 활발하게 연구되고 있다. Rietman 등은 신경회로망을 이용한 플라즈마 식각공정의 모델링을 수행하였으나 고정된 입력변수만에 의한 모델링에 의존하여 실제 식각이 진행되는 동안 또는 식각 후의 장비 및 공정 상태를 적절하게 포함하지 못하고 있다 [2]. 또한 Lee와 Spanos는 식각이 진행되는 동안의 실시간 장비 데이터를 이용하여 식각공정의 모델링을 수행하였으나 단순히 실시간 데이터의 평균값만을 입력변수로 포함하였다 [3].

본 논문에서는 식각이 진행되는 동안의 공정상태를 보다 정확

하게 반영하기 위해 플라즈마 식각장비의 대표적인 in-situ sensing 기술인 OES (optical emission spectroscopy)를 활용하여 측정된 autoscan 데이터를 이용하여 64M DRAM s-poly 플라즈마 식각공정의 모델링을 수행한다. 본 논문에서는 플라즈마 식각공정과 필요한 autoscan 데이터 수집 방법에 대해 설명하고, 주 요소 분석 (principal component analysis) 기법을 이용한 신경회로망 모델링에 대해 설명한다. 그리고 구성된 플라즈마 식각공정의 모델을 통해 식각 후의 웨이퍼 상태 예측결과를 모의실험을 통해 설명한다.

#### 2. 대상 공정 및 데이터 수집

플라즈마 식각공정은 여러 가지의 전기적이고 물리적인 효과와 상호 작용하는 화학적인 반응을 가진 복잡한 비선형적인 요소들을 포함하고 있다. 공정과 제어변수상의 관계는 일반적으로 공정 제어입력을 만들기 위한 많은 시행착오와 실험적 설계의 조합으로 발전되어 왔다. 이러한 recipe들은 반복되는 사용과 함께 변화하는 반응실 안의 환경들에 유동적으로 적응하지 못한다. 플라즈마 식각공정에서 중요한 문제점은 식각이 진행되는 동안 공정과 웨이퍼 상태에 대한 정보를 획득하는 것이다. in-situ 측정 중 가장 일반적인 기술은 optical emission spectroscopy (OES)이다. OES는 일반적으로 식각시 소멸되거나 발생하는 반응물과 부산물에 대한 광학적인 방사측정으로 이루어진다. 실제로 반도체 생산라인에서는 반응물 또는 부산물종류에 관계되는

단일 파장에 대한 강도 (intensity)를 모니터링함으로써 식각의 종료시점을 결정하는데 이러한 OES 신호들을 사용하고 있다.

플라즈마 식각공정은 많은 입력변수들을 갖고 있으며 이들중 중요한 변수들은 D.O.E (Design Of Experiments)에 의해 선택된다. 표 1은 본 연구에서 사용한 주요 입력변수와 식각장비가 갖는 특성에 의한 식각시의 입력범위이다.

표 1 입력변수와 범위  
Table 1. Input variables and their ranges

입력변수	범위	
	최소	최대
Pressure [mTorr]	20	26
Source Power [W]	1900	2500
Bottom Power [W]	70	90
Cl <sub>2</sub> Flow Rate [sccm]	150	200
N <sub>2</sub> Flow Rate [sccm]	5	15

그리고 "fractional factorial experiments"에 의해 17회의 실험을 수행하여 17개의 체계적인 입출력 데이터집합을 얻는다. autoscan을 위한 식각웨이퍼는 "dummy wafer"라고 불리우며 실제의 제품생산에 사용되는 "profile wafer"와 달리 웨이퍼 표면에 포토공정을 통한 마스크가 없는 것이 특징이다. 그림 1과 같이 autoscan은 식각중에서 가장 활발히 생성, 소멸되는 광파장을 탐색하는 기능이다. 그림 1은 2600~5500Å의 파장대역의 autoscan 모니터링 화면이다. 이 때 데이터 추출은 매 20Å마다 이루어 졌다. 세로축은 광의 세기를 0~10[V]의 전압대역을 나타낸다.

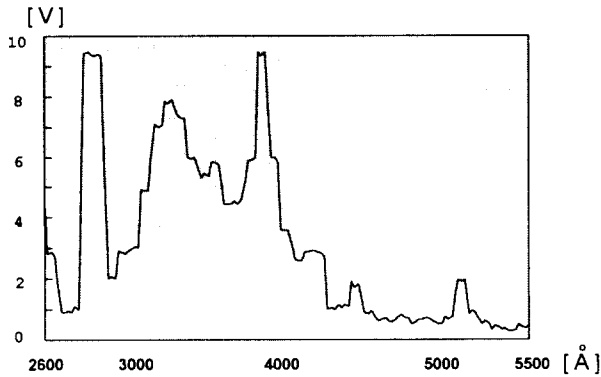


그림 1. Autoscan 데이터의 예  
Fig. 1 Example of autoscan data

### 3. 신경회로망 모델링

#### 3.1 주 요소 분석 (Principal Component Analysis)

일반적으로 optical emission spectral (OES)은 수백개의 파장에 대한 강도의 측정치로 구성되어 있다. 이러한 많은 수의 데이

터들을 분석하는 것은 매우 어렵다. 주 요소 분석 (principal component analysis: PCA)은 데이터 감소와 분석을 위한 매우 효과적인 방법으로 알려져 있다.

PCA의 기본적인 개념은 그림 2와 같다. 좌표축  $x_1$ 과  $x_2$ 는 점들 점들로 표시된 샘플 데이터를 가진 두 개의 zero-centered measurement 변수들을 가리키고 있다. 벡터  $v_1$ 과  $v_2$  위로의 표준편차  $\sigma_1$ 과  $\sigma_2$ 는 샘플 데이터와 자연좌표계의 관련성에 대한 정보를 제공하고 있다. 이때 자연좌표계  $v_1$ 과  $v_2$  위로 원래의 축  $x_1$ 과  $x_2$ 을 회전시킴으로써 직교공간은 1차축이 최대의 분산의 방향으로 정렬되는 곳에서 발견된다. 데이터의 최대 분산의 방향을 따라 좌표축을 회전함으로써, 두개의 변수  $x_1$ 과  $x_2$ 는 하나의 변수  $v_1$ 으로 잘 표현될 수 있다. PCA는 데이터의 최대 분산의 방향을 제공한다. 이러한 경우에 처음의 주 요소 방향으로서의  $x_1$ 과  $x_2$ 의 투영은 데이터의 차원을 2에서 1로 압축시킨다.

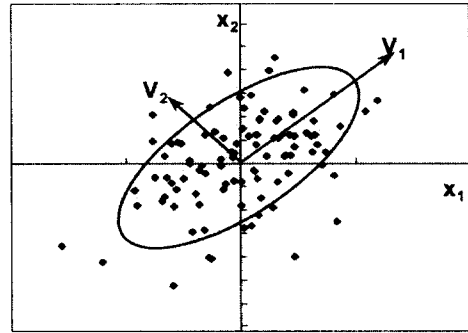


그림 2. 주 요소 분석  
Fig. 2 Principal Component Analysis

#### 3.2 실험 데이터에 대한 주 요소 분석의 적용

다수의 autoscan 데이터들을 감소시키기 위해서, PCA가 적용되었다. 원래의 OES 데이터는  $r \times n$  행렬로 형식화된다. 여기서  $r$ 는 실험을 한 횟수이며,  $n$ 은 각각의 실험에서 측정된 샘플 데이터의 수이다. 그러므로 결과적인  $17 \times 146$  데이터 행렬  $L_{ij}$ 에서 각각의 행  $i$ 는 17번의 실험 횟수중 하나의 데이터에 대응되고, 각각의 열  $j$ 는 260에서 550nm 범위의 146개의 파장중 하나에서의 강도를 표현하고 있다. 이때 raw 데이터 행렬은 각각의 파장의 관점에서 중심값에 대한 편차를 의미하도록 식 (1)을 통해 행렬  $M$ 으로 변환된다.

$$M_{ij} = (L_{ij} - \bar{L}_j) \quad \text{for } 1 \leq i \leq r \text{ and } 1 \leq j \leq n \quad (1)$$

여기서  $\bar{L}_j$ 는 전체실험횟수( $r$ )에대한 열  $j$ 의 샘플 데이터의 평균이다.  $n \times n$  (이 경우에는  $146 \times 146$ ) 공분산행렬  $C$ 는 행렬  $M$ 의 을 이용하여 다음과 같이 계산된다.

$$C = \left( \frac{1}{r-1} \right) \cdot M^T M \quad (2)$$

공분산행렬  $C$ 의 고유벡터들은 다음과 같이 정의된다.

$$CV = VA \quad (3)$$

여기서  $V=[v_1 \cdots v_{n=146}]$ 은  $n$ 개의 고유벡터를 가지는  $n \times n$  행렬이다. 그리고  $\Lambda$ 는 고유치의  $n \times n$  대각행렬이다. (3)에서 각각의  $n$  차원 고유벡터  $v_j$ 는  $j$ 번째의 고유치  $\lambda_j$ 와 대응한다. 고유치를  $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ 과 같은 내림차순으로 정렬하면 공분산행렬  $C$ 의 가장 큰 고유치  $\lambda_1$ 과 결합한 고유벡터  $v_1$ 은 데이터 내에서 가장 큰 분산의 방향을 나타낸다. 어떤 방향  $v_j$ 로의 분산은 이 벡터에 연관된 고유치  $\lambda_j$ 를  $n$ 개의 고유치의 합이나 공분산행렬  $C$ 의 trace로 나눔으로써 계산될 수 있다.

$$\frac{\lambda_j}{\sum \lambda_j} = \frac{\lambda_j}{\text{trace}(C)} \quad (4)$$

만일  $M$ 을 공분산행렬  $C$ 의 고유벡터  $v_j$ 에 투영한다면  $r$ 차원의 벡터  $t_j$ 를 얻을 수 있다.

$$Mv_j = t_j \quad \text{for } 1 \leq j \leq n \quad (5)$$

$M$ 을 고유벡터 행렬  $V$ 로 투영하면 대응하는  $r \times n$ 행렬  $T$ 는  $n$ 개의 열 벡터로 이루어진다. 이 변환은 데이터의 새로운 표현을 만든다.  $V$ 안의 고유벡터로  $M$ 을 투영함으로써 새로운 행렬  $T$ 는 직교하는 열 벡터를 갖고 모든 분산  $\Lambda$ 의 나머지를 보유한다.

$$MV = T \quad (6)$$

그리고

$$\Lambda = \left( \frac{1}{r-1} \right) \cdot T^T T \quad (7)$$

$\Lambda$ 은 (3)에서와 같은 고유치이다. 다시 말하면, 행렬  $T$ 는 그림 1에 나타난 것처럼  $M$ 의 방위를 새로운 직교 좌표  $V$ 로 나타낸다. 주 요소나  $C$ 의 고유벡터가 최대 또는 최소 분산의 방향을 따라 정렬되었기 때문에 행렬  $V$ 의 차원  $n$ 은 원하는 만큼의 분산을 포함하는 수까지 줄어 들 수 있다. 예를 들면,  $y$ 개의 주 요소가 데이터에서 분산의 99%를 포함한다면,  $n \times n$  고유벡터 행렬  $V$ 를  $n \times y$  행렬  $\hat{V}$ 로 줄일 수 있다. 그러면 데이터 행렬  $M$ 은  $r \times n$  행렬  $T$ 로부터  $r \times y$  행렬  $\hat{T}=[t_1, t_2, \dots, t_y]$ 까지 축소된 데이터를 줄이도록 이들 주 요소들로 투영될 수 있다.

$$M\hat{V} = \hat{T} \quad (8)$$

### 3.3 플라즈마 식각공정 모델링을 위한 신경회로망

신경회로망은 입력력 변수 사이의 비선형성을 학습할 수 있는 능력을 가지고 있다. 최근의 연구들은 신경회로망을 이용한 플라즈마 식각공정의 모델링은 기존의 통계적 모델링 기법에 비해 보다 정확한 모델링을 수행할 수 있다는 것을 보여준다 [2],[3]. 본 논문에서는 64M DRAM s-poly 식각공정에 대한 모델링을 위해 3개의 층을 갖는 신경회로망을 구성하였다. 입력층은 OES 데이터의 주 요소와 실험계획법에 의한 입력변수로 구성된다. 입

력층은 주 요소 분석에 의해 결정된 특정 허용오차 범위 내의 누적분산을 갖는 주 요소들이 모두 포함된다. 출력층은 식각률과 균일도로 구성된다. 신경회로망의 학습은 feed-forward error back propagation 알고리즘을 통해 이루어진다. 일정 범위의 랜덤한 값으로 초기화된 신경회로망의 모든 연결하중과 문턱치를 통해 신경회로망의 출력값이 계산되면 강하법 (gradient descent method)을 이용하여 목표값과 신경회로망 출력값과의 오차가 역전과되어 이 오차의 미분값에 비례하는 값 만큼 이 오차를 감소시키는 방향으로 신경회로망의 연결하중들을 갱신하게 된다. 사용된 신경회로망의 구조는 그림 3과 같다.

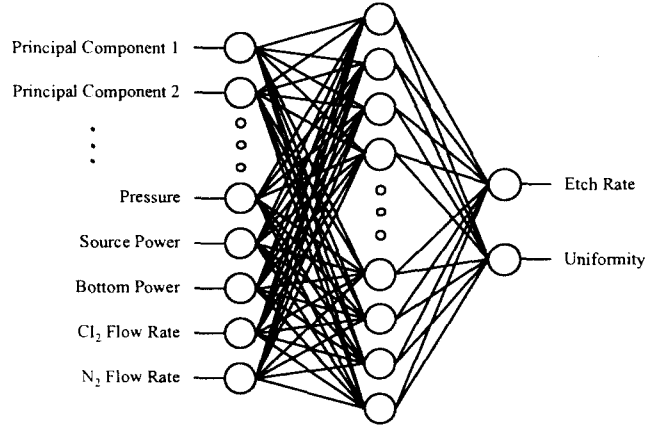


그림 3. 신경회로망의 구조  
Fig. 3 Structure of neural network

### 4. 웨이퍼 상태예측

각 데이터에 해당하는 고유벡터에 대한 분산을 계산함으로써 전체 데이터의 특성을 가장 잘 표현할 수 있는 주 요소들을 내림차순으로 정렬할 수 있다. 이 때 가장 큰 분산값을 갖는 고유벡터에 해당하는 데이터가 전체 데이터의 특징을 가장 잘 표현한다고 할 수 있다. 따라서 고유벡터의 누적분산의 비율을 구해보면 내림차순으로 정렬된 고유벡터에 해당하는 데이터가 전체 데이터의 분산에 얼마만큼의 영향을 미치는지 알 수 있다. 그림 4는 가장 큰 고유벡터로부터 하나씩 증가시키며 이러한 누적분산의 백분율을 산출한 것이다.

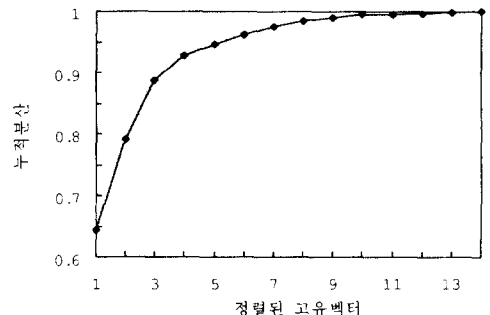


그림 4. 정렬된 고유벡터의 누적 분산  
Fig. 4 Cumulative variance of sorted eigenvectors

#### 4.1 신경회로망의 구조

그림 4를 살펴보면 첫 번째 주 요소가 전체 누적분산의 약 65%를 차지하고, 세 번째 주 요소 까지의 누적분산은 약 90%, 다섯 번째 주 요소까지의 누적분산은 전체 누적분산의 약 95%를 차지하는 것을 볼 수 있다.

예를 들어 분산의 5% 손실을 가지고 모델링을 수행한다면 식 (8)에서 계산된 것처럼 행렬  $M$ 은  $17 \times 146$  행렬로부터 분산에서 5%의 손실을 가지고 있는  $17 \times 5$  행렬  $\hat{T} = [t_1, t_2, t_3, t_4, t_5]$ 까지 전체 데이터를 감소시킬 수 있다. 결과적으로 64M DRAM s-poly 플라즈마 식각공정의 모델링에 사용될 신경회로망은 입력층의 뉴런 수는 제어입력 (recipe)에 포함된 입력변수 5개와 주 요소 분석에 의해 결정된 5개의 autoscan 데이터를 포함하여 10개가 된다. 따라서 이 경우 신경회로망의 구조는  $10 \times 16 \times 2$ 로 하여 학습을 수행하였다.

#### 4.2 웨이퍼 상태 예측결과

위와 같은 방법을 통해 구성된 64M DRAM s-poly 플라즈마 식각공정의 모델을 이용하여 2개의 실험데이터를 이용하여 식각이 끝난 후의 웨이퍼 상태를 예측하였다. 이때 예측이 필요한 출력변수는 식각률과 균일도이다. 5%의 손실을 갖는 예측결과를 표 2에 나타내었다.

표 2를 살펴보면 학습에 사용되지 않은 실험 데이터에 대한 예측결과는 실제 식각률과 균일도를 거의 정확하게 예측하고 있음을 알 수 있다. 또한 신경회로망 학습에 소요된 시간은 recipe의 입력변수 5개만을 취한 신경회로망보다 학습시간이 더 짧은 경향을 보였다.

표 2. 식각 후의 웨이퍼 상태예측 결과 (5% 손실)

Table 2. Wafer state prediction after etching (5% Loss)

예측항목	실험 1		실험 2	
	목표값	예측값	목표값	예측값
식각률	4316	4309	3958	3959
균일도	10.7	10.9	13.4	13.6

## 5. 결론

본 논문에서는 실시간 신호 중 하나인 autoscan 데이터를 이용하여 64M DRAM s-poly 플라즈마 식각공정의 신경회로망 모델링을 수행하였다. 주 요소 분석을 통해 추출된 autoscan 데이터와 recipe의 입력변수를 통해 구축된 모델을 이용하여 식각 후의 웨이퍼 상태를 예측한 결과 우수한 예측성능을 가짐을 알 수 있었다. 또한 recipe의 입력변수만을 이용해 신경회로망을 학습시킨 경우와 비교해볼 때 입력층 뉴런의 개수와 은닉층 뉴런의 개수가 증가하여 더 많은 양의 연산이 필요함에도 불구하고 학습시간이 감소되는 경향을 보였다.

앞으로의 연구는 autoscan 데이터와 실시간 장비 데이터를 이용한 시계열 모델링 (time series modeling) 및 공정의 표동 (drift)이나 장비의 연속적인 사용으로 인한 영향 (예를 들면, 반응실 내벽에 생성되는 박막 등)에 대한 연구가 포함되어야 할 것이다.

## 참고문헌

- [1] S. May, J. Huand and C. Spanos, "Statistical Experimental Design in Plasma Etch Modeling", *IEEE Trans. Semicond. Manufact.*, vol.4, no.2, May 1991.
- [2] E. A. Rietman and E. R. Lory, "Use of Neural Networks in Modeling Semiconductor Manufacturing Process: An Example for Plasma Etch Modeling", *IEEE Trans. Semicond. Manufact.*, vol.6, no.4, Nov 1991
- [3] S. F. Lee and C. J. Spanos, "Prediction of Wafer State After Plasma Processing Using Real-Time Tool Data," *IEEE Trans. Semicond. Manufact.*, vol.8, no.3, pp.252-261, Aug. 1995.
- [4] Xinix 1014, *End Point Detection manual*, 1993.
- [5] P.Box, W.B.Hunter and J.S.Hunter, *Statistics for Experimenters*, New York : Wiley, 1978.
- [6] S. M. Sze, *VLSI Technology*. 2nd edition, pp.248, 1990.