

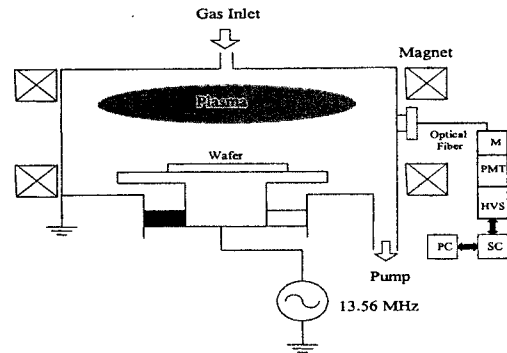
## 플라즈마 공정 감시를 위한 Actinometric 광방사분광기 정보의 신경망 모델링

권상희, 황보광, 이규상, 우형수, 김병환\*  
 세종대학교

### Neural Network Modeling of Actinometric Optical Emission Spectroscopy Information for Monitoring Plasma Process

Sang Hee Kwon, Bo Kwang, Lee Kyu Sang, Hyung Soo Uh, Byungwhan Kim\*  
 Sejong University

**Abstract** -플라즈마 공정은 집적회로 제작을 위한 미세 박막의 증착과 패턴닝에 핵심적으로 이용되고 있다. 본 연구에서는 플라즈마 공정감시와 제어에 응용될 수 있는 모델을 제안한다. 본 모델은 광방사분광기 (Optical emission spectroscopy-OES)정보와 역전파 신경망을 이용해서 개발하였다. 제안된 기법은 Oxide 식각공정에서 수집한 데이터에 적용하였으며, 체계적인 모델링을 위해 공정데이터는 통계적 실험계획법을 적용하여 수집되었다. Raw OES 정보대신, Actinometric OES 정보를 이용하였으며, 신경망의 예측성능은 유전자 알고리즘을 이용해서 증진시켰다. OES의 차수를 줄이기 위해 주인자 분석 (Principal Component Analysis-PCA)을 세 종류의 분산 (100, 99, 98%)에 대해서 적용하였다. 최적화한 모델의 예측에러는 323 k/min이었다. 이전에 PCA를 적용하고 은닉층 뉴런의 함수로 최적화한 모델의 예측에러는 570 k/min이었으며, 개발된 모델은 이에 비해 43% 증진된 예측 성능을 보이고 있다.



#### 1. 서 론

반도체 플라즈마 공정은 미세 박막의 증착과 패턴닝에 핵심적으로 응용이 되고 있다. 플라즈마 공정의 최적화와 실시간 감시, 그리고 제어를 위해서는 공정 예측모델이 요구된다. 플라즈마의 복잡한 물리화학적 특성을 모델링하기 위해 신경망, 특히 Backpropagation Neural Network(BPNN) [1]이 많이 응용되고 있다. 플라즈마 공정 감시를 위해 가장 많이 응용되고 있는 In-situ 센서로는 Optical Emission Spectroscopy(OES)이 있다. OES 정보와 신경망을 결합한 플라즈마 식각공정모델이 보고된 바가 있다 [2-3]. 부분적 OES 정보보다, 전체 OES 정보에 기반하여 개발된 모델이 증진된 예측성능을 보였다 [3]. 한편, BPNN의 학습에는 다수의 학습인자가 관여하고 있으며 [4], 따라서 이들의 영향을 최적화하기 위해 최근 유전자알고리즘 (Genetic Algorithm-GA) [5]를 이용한 식각공정모델이 개발된 바 있다 [6]. 공정의 화학적 분석의 신뢰도를 증진하기 위해 종종 Raw OES 대신에 Actinometric OES 정보를 이용하는 경우가 있으며, 후자의 정보를 이용한 모델이 개발된 사례가 없다.

본 연구에서는 Actinometric OES 정보와 BPNN, 그리고 주인자 분석 (Principal Component Analysis-PCA)를 이용하여 플라즈마 식각공정 모델을 개발한다. 다수 학습인자의 영향은 유전자 알고리즘 GA를 이용하여 최적화하였다. OES 정보는 Oxide 박막의 플라즈마 식각공정 중에 통계적 실험계획법을 이용하여 수집되었다. 개발된 식각률 모델은 이미 보도된 모델 [3]과 그 성능을 비교평가하였다.

#### 2. 본 론

##### 2.1 실험적인 세부사항

실험에 이용된 식각 장비의 개요도가 그림 1에 도시되어 있다. 하부전극은 13.56 MHz에서 동작하는 RF 전력이 연결되어 있다. 하부전극과 평행한 자력선은 챔버의부에 설치되어 있는 Electromagnetic 코일에 DC 전류를 전달하여 발생한다. 테스트 패턴은 Boron이 도핑된, (100)방향의 Si 기판위에 형성하였으며, 보다 구체적인 내용은 참고문헌 [3]에 상세히 기술되어 있다. 식각공정은 전술한 MERIE 장비를 이용하여 CHF<sub>3</sub>/CF<sub>4</sub> 플라즈마에서 수행하였다. 체계적인 모델링을 위해 2<sup>4-1</sup> 부분인자 실험계획법 [7]이 적용되었으며, 실험계획법에 이용된 공정변수와 범위는 표 1에 있다. 수집된 데이터 공정변수 중심점에 해당하는 실험과 함께 BPNN의 학습에 이용되었다. 학습된 모델의 예측성능은 다른 8개의 실험 데이터를 이용하여 평가하였다. OES의 파장 범위는 200-800 nm 이었다. Actinometric OES 데이터는 Raw OES 데이터를 7504 Å에서 수집한 Ar 유량에 대한 Intensity로 나누어 구해진다. 각 OES Spectra는 총 1881개의 라디칼 변수 정보를 포함하고 있다. 이를 신경망 입력패턴으로 사용하기에는 너무 변수가 많으므로 변수를 줄이기 위해 PCA를 적용하였으며, 데이터 분산 (Variance)는 100, 99, 그리고 98%로 설정하였다. 그 결과, 100, 99, 그리고 98%에 대해 16, 7, 그리고 5개로 변수의 차수가 감소하였다.

<그림 1> 식각 장비의 개요도

<표 1> 실험변수와 범위

Parameters	Range	Unit
CHF <sub>3</sub> 유량	20-80	sccm
CF <sub>4</sub> 유량	10-40	sccm
RF 전력	300-800	W
압력	50-200	Torr

##### 2.2 신경망 모델링

BPNN의 은닉층은 하나이며, 학습 규칙으로는 일반화된 델타 규칙을 적용하였으며, 뉴런간 WEIGHT 조정은 (1)과 (2)에 의해 이루어졌다.

$$W_{i,j,k}(m+1) = W_{i,j,k}(m) + \eta \Delta W_{i,j,k}(m) \quad (1)$$

여기서 W<sub>i,j,k</sub>는 (k-1) 층의 j 번째 뉴런과 k층의 i 번째 뉴런사이의 연결강도 이고 ΔW<sub>i,j,k</sub>는 학습편차를 줄이기 위해 인가되는 연결강도의 변이이며 다음과 같이 정의된다.

$$\Delta W_{i,j,k} = \frac{\partial E}{\partial W_{i,j,k}} \quad (2)$$

BPNN의 학습에는 다수의 학습인자가 관여하고 있으며, 이들은 모델의 학습과 예측성능에 상당한 영향을 미친다. 학습인자에는 학습 허용도 (Training Tolerance-TT), 은닉층 뉴런수 (Number of Hidden Neurons-NHN), 초기웨이트의 크기 (Magnitude of Initial Weight Distribution-MIWD), 그리고 은닉층 뉴런의 단극성 시그모이드함수의 경사 (Gradient of Bipolar Sigmoid Function-GBSF), 출력층 뉴런의 선형함수의 경사 (Gradient of Linear Function-GLF) 등이 있다. 경우에 따라서는 주어진 TT에서 학습이 종료되지 않을 수도 있으며, 과도한 학습이 진행되는 것을 방지하기 위해 다른 학습종료 기준으로 Epoch를 각 기 10,000과 50,000으로 설정하여 평가하였다.

##### 2.3 GA-BPNN 모델

BPNN의 학습인자를 최적화하기 위해 GA를 응용하였다. 각 학습인자의 범위는 표 2에 나타나 있다. GA에 관여하는 초기 해의 규모

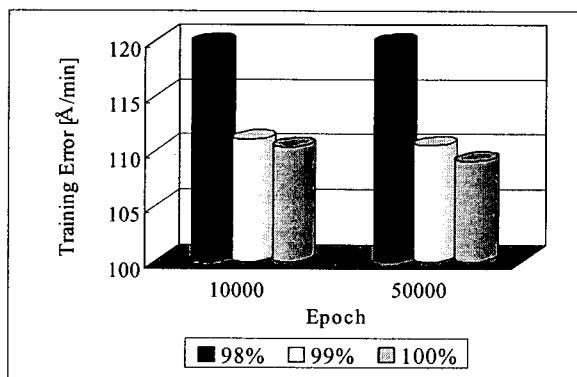
는 100으로 교배와 돌연변이 확률은 각각 0.9와 0.1로 설정하였다. 해의 적합도를 평가하기 위한 적합도 함수는 식 (3)로 정의된다.

$$F = 1/(1+RMSE) \quad (3)$$

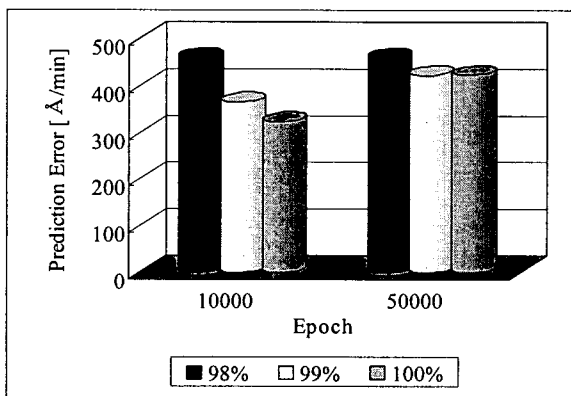
여기서 RMSE는 학습에러를 의미한다. 세대 수는 100으로 설정하였으며, 따라서 유전자 동작은 세대수 100에서 종료된다. 전체 세대에서 발생한 모델 중 가장 작은 예측 에러를 보이는 모델을 결정하였다.

<표 2> 유전자 알고리즘 최적화를 위한 인자와 범위

학습인자	범위
TT	0.04-0.12
NHN	5-10
MIWD	±0.4-±1.2
GBSF	0.4-2.0
GLF	0.4-2.0



<그림 2> Epoch에 따른 PCA-BPNN 학습에러



<그림 3> Epoch에 따른 PCA-BPNN 모델의 예측에러

GA-BPNN 모델의 학습오차는 그림 2에 도시되어 있다. 그림 2에서와 같이 Epoch에 상관없이, 분산의 증가에 따라 학습에러가 감소하고 있다. Epoch가 10,000일 경우 100%의 분산에서 가장 작은 학습에러 110 Å/min를 구했다. Epoch가 50,000인 경우에서도 동일한 분산에서 최소 학습에러 109 Å/min를 구했다. 두 모델의 에러가 비슷하며, 이는 Epoch가 BPNN 학습에 별 영향을 주지 않음을 의미한다. 한편, GA-BPNN 모델의 예측에러는 그림 3에 도시되어 있다. 그림 3에서와 같이, Epoch가 10,000일 경우 분산의 증가에 따라 예측에러가 감소하고 있으며, 100%의 분산에서 가장 작은 323 Å/min를 얻었다. 98% 모델의 예측에러 (459 Å/min)와 비교하여, 이는 29.6% 증진된 예측성능을 의미한다. 한편, Epoch가 50,000일 경우에는 99%의 분산에서 가장 작은 예측에러 422 Å/min를 얻었다. 이 모델과 비교하여, Epoch 10,000에서 구한 최적의 모델은 예측성능이 23% 증진되었다. 결국, 최적의 예측모델은 Epoch 10,000, 분산 100%에서 얻을 수 있었다.

한편, 개발된 모델을 최근 발표한 모델 [3]과 비교하였다. 이 모델은 100%의 분산에서 학습인자의 영향을 최소화하지 않고 개발되었으며, 모델의 예측에러는 570 Å/min 이었다. 본 연구를 통해 개발된 모델은 이 모델과 비교하여 43.3%의 증진된 예측치를 보이고 있다. 이는 제안하는 기법이 보다 우수한 예측성능을 가지는 모델개발에 효과적임을 보여준다.

## 감사의 글

이 논문은 2006년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임 (KRF-2006-311-D00047).

## [참고 문헌]

- [1] D. E. Rummelhart and J. L. McClelland, *Parallel Distributed Processing*, Cambridge, M.I.T. Press, 1986.
- [2] S. J. Hong, G. S. May, and D. C. Park, "Neural network modeling of reactive ion etching using optical emission spectroscopy data," *IEEE Trans. Semicond. Manufact.* 16, 598-608, 2003.
- [3] B. Kim, J. Bae, and W. S. Hong, "Plasma control using neural network and optical emission spectroscopy," *J. Vac. Sci. Technol. A* 23, p. 355-358, 2005.
- [4] B. Kim and S. Park, "An optimal neural network plasma model: a case study," *Chemometr. Intell. Lab. Syst.* 56, 39-50, 2001.
- [5] D. E. Goldberg, *Genetic Algorithms in Search, Optimization & Machine Learning*, Addison Wesley, Reading, MA, 1989.
- [6] 권민지, 김병환, 광방사분광기와 신경망을 이용한 플라즈마 식각공정 모델링, 2007년도 대한전기학회 하계학술대회논문집, p. 1807-1808, 2007. 7.
- [7] D. C. Montgomery, *Design and Analysis of Experiments*, John Wiley & Sons, Singapore, 1991.