

반도체 공정 최적화를 위한 일반화된 회귀 신경망 플라즈마 모델

박성진, 김병환  
전남대학교

A Generalized Regression Neural Network Plasma Model for Semiconductor Process Optimization

Sung-jin Park, Byungwhan Kim  
Chonnam National University

초록 - 일반화된 회귀 신경망을 이용하여 반도체 공정 최적화를 위한 플라즈마를 모델링한다. 플라즈마는 Box-Wilson 실험계획표에 의해 특성화되었으며, 여기에서 변화시킨 인자로는 소스전력, 압력, 척지대의 위치, 그리고 염소의 유량이다. 총 24회의 실험이 수행이 되었으며, 플라즈마 변수는 Langmuir Probe를 이용하여 측정하였다. 측정된 주요 플라즈마 변수로는 전자밀도, 전자온도, 그리고 플라즈마 전위이다. 폭변수를 점진적으로 증가시켜 회귀신경망을 최적화하였으며, 최적화된 모델은 통계적인 반응표면모델과 비교하였다. 비교결과, 회귀신경망은 반응표면모델에 상응하는 예측능력을 보이고 있음을 알 수 있었다.

그림 1에 나타나 있다. 장비 외측은 코일로 감겨져 있으며, 권선 수는 10회이다. 코일에 rf 전력이 공급되면서 장비 안에 플라즈마가 발생하며, 이때 Langmuir probe에 의해 기본적인 플라즈마 변수가 측정된다. 측정된 변수로는 기본적인 전자 밀도, 전자 온도, 그리고 플라즈마 전위 (potential) 등이다. 플라즈마는 공정 (또는 장비) 인자의 변화에 매우 민감하게 반응을 하며, 따라서 이들 인자와 플라즈마 변수와의 체계적인 모델링을 위해 통계적인 기법 중의 하나인 Box-Wilson 실험 계획 [11]이 이루어졌으며, 이는 다시 2<sup>4</sup> full factorial 계획과 8개의 면 중심 어레이(array)로 구성된다. 변화시킨 공정 변수로는 소스 전력, 압력, 척 (chuck) 위치, 그리고 염소 (chlorine) 유량이며, 이들의 범위는 표 1에 나타나 있다.

중심어: 플라즈마, 일반화된 회귀신경망, 표면반응모델, 실험계획표

1. 서론

반도체 칩 (chip) 제조시, 플라즈마 (plasma) 공정은 다양한 박막의 증착과 미세 패터닝의 식각 (etching)에 이용되고 있다. 공정 개발에 소요되는 시간과 노력을 줄이기 위해서, 이들 공정의 시뮬레이션 (simulation)을 통한 최적화가 요구되고 있으며, 이는 다시 공정의 핵심인 플라즈마에 대한 정확한 예측 모델 개발을 요구하고 있다. 플라즈마 모델은 주로 복잡한 물리적인 식에 의해 이루어지고 있으나 [1-2], 이들 모델의 개발에는 고도의 물리, 화학적인 지식이 요구될 뿐 아니라, 모델을 형성했다 하더라도 임의의 공정 내지 장비 변수에 대한 플라즈마 예측치를 얻는 데에 과다한 계산속도가 요구됨으로 실제 칩 제조에 그리 기여를 못하고 있는 실정이다. 특히, 물리적인 모델 형성 과정에 수반되는 가정 (assumption) 들의 존재로 그 응답들이 실제 측정치에서 크게 벗어나, 이 모델들의 사용은 극히 미미한 형편이다.

앞서의 어려움을 극복하고자, 최근에 에러 역전파 (error back-propagation) 신경망을 이용한 장비 플라즈마 모델이 발표된 바가 있었다 [3]. 이에 앞서 신경망은 다수의 반도체 공정 모델링에 응용된 바가 있었다. [4-10], 본 연구에서는 신경망의 다른 구조인 일반화된 회귀신경망 (generalized regression neural network)을 이용하여, 공정 플라즈마를 모델링 한다. 구체적인 플라즈마 유형은 반구형 (hemispherical) 식각 장비에서 생성되는 자기유도 결합형 (inductively coupled) 플라즈마이며, 변화시킨 공정 변수로는 소스 전력, 압력, 척 (chuck) 위치, 그리고 염소 (chlorine)의 유량이다. 체계적인 모델링을 위해 이들 4개의 변수에 대한 Box-Wilson의 면 중심 (face-centered) 실험 계획이 이용되었으며, 총 24번의 실험이 행해졌다. 매 실험마다, Langmuir probe를 이용하여 모델링 하는 플라즈마 변수를 수집하였다.

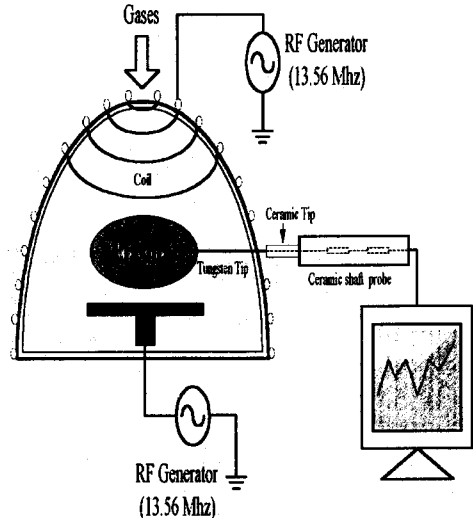


그림 1. 반구형 플라즈마 식각 장치.

Fig 1. Hemispherical plasma etch system.

공정변수	범위
소스전력	700~900 watts
척 위치	-30~90 mm
염소	60~120 seem
압력	5~10 mTorr

표 1: 실험 변수와 범위.

Table 1. Process factors and experimental range.

2. 실험 장비와 통계적인 실험계획

실험에 이용된 반구형 식각 장비의 개략적인 모습이

### 3. 일반화된 회귀 신경망

일반화된회귀신경망 (generalized regression neural network) [12]의 구조가 응답이 하나인 경우에 대해서 그림2에 도시되어 있다. GRNN의 예측치 ( $\hat{Y}$ )를 수학적으로 표현하면

$$\hat{Y} = \frac{\sum_{j=1}^n y_j \exp[-D(x, x_j)]}{\sum_{j=1}^n \exp[-D(x, x_j)]} \quad (1)$$

여기서 함수 D는

$$D_x(x, x_j) = \sum_{i=1}^p \left( \frac{x_i - x_{ji}}{\sigma_j} \right)^2 \quad (2)$$

로 표현되고, 변수 n과 p는 각각 독립변수와 학습 패턴의 수를 지칭한다. 그리고 변수  $\sigma$ 는 폭 (spread) 변수 (또는 smooth factor)라 칭하며, 이 폭이 작은 경우, 좋은 패턴 인식을 이룰 수 있으나, 모델의 일반화 능력을 저하시키며, 폭이 넓을 경우 반대의 현상이 발생한다. 따라서 최적의 일반화 능력을 얻기 위해 이 값의 조정이 필요하다.

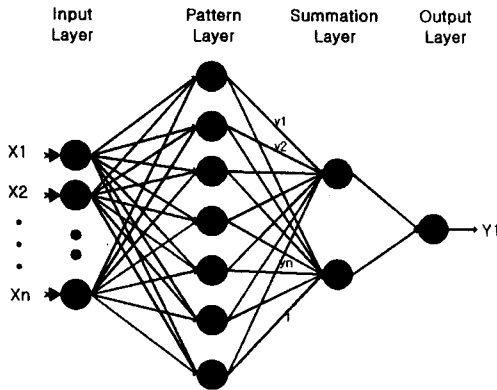


그림 2. 일반화된 회귀 신경망 구조.

Fig 2. Architecture of generalized regression neural networks.

그림 2에서와 같이, GRNN은 총 4개의 층, 즉 입력층, 패턴층, 합층, 그리고 출력층으로 구성된다. 입력층의 뉴런수는 독립 변수 (예 공정 변수)의 수와 일치하며, 패턴층의 뉴런 수는 학습패턴의 수와 일치한다. 합층은 두 개의 뉴런으로 구성되며, 처음의 뉴런은 패턴층의 모든 출력을 모두 합해, 식 (1)의 분자 항을 계산하며, 다른 하나의 뉴런은 분모 항을 계산한다. 입력층과 패턴층간의 하중치 ( $W_p$ )는 입력패턴 (X)에 의해 결정되며, 이를 표현하면,

$$W_p = X^t \quad (3)$$

패턴층의 하나의 뉴런은 합층의 두 개의 뉴런에 연결되며, 패턴층의 i번째의 뉴런과 합층 첫 뉴런간의 연결 하중치는  $y_i$ 가 된다. 이 i번째의 뉴런과 합층의 다른 하나의 뉴런과의 연결 하중치는 1이 된다. 합층과 출력층간의

하중치 ( $W_s$ )는  $y_j$ 와 1에 의해 다음과 같이 결정된다.

$$W_s = [Y \text{ ones}] \quad (4)$$

출력층에서는, 단순히 합층의 두 뉴런의 출력을 나누어 식(1)의 예측치를 출력한다.

### 4. 결과

그림 3은 폭 변수를 0.2에서 2.0까지 변화시킬 때, GRNN 예측치들을 실제치와 비교하여 도시하고 있다. 학습 패턴으로는 full factorial 실험계획에서 얻은 16개의 패턴이 이용되었으며, 테스트 패턴으로는 총 8개의 면중심 어레이가 이용되었다. 그림 3에서와 같이, 폭변수 값이 0.5 이상의 큰 값을 가질 경우, 폭변수의 증가에 따라 예측치들은 실제치에서 크게 벗어나고 있으며, 한편 폭변수의 값이 작을 경우, 그림에서 0.5 이하, 예측치들에서의 변화가 거의 없음을 알 수 있다. 이는, 폭변수 값이 작을 경우 (2)의 값이 커지게 되며, 그 절대값 (absolute value)이 2.5보다 클 경우 (2)의 exponential 값이 일정한 값을 가지게 되는 것에 연유한다. 그림 3에서 최적의 전자 밀도 모델은 폭변수 0.5에서 결정되었다. 그림 4에서 앞서의 최적의 전자밀도와 통계적인 반응표면 모델 (RSM)의 예측치들이 실제치와 비교하여 도시되고 있다. RSM 모델은 SAS Package를 이용하여 구하였으며, fit 에러와 예측 편차는 각기 0.239와 0.462이었다. 여기서 편차는 Root mean-squared error (RMSE) metric에 의해 계산이 되었으며, 이는 아래와 같이 표현된다.

$$RMSE = \sqrt{\frac{(y_a - y_b)^2}{n-1}} \quad (5)$$

여기서  $y_a$ 와  $y_b$ 는 각기 실제 측정치와 모델 예측치를 나타내며, n은 전체 테스트 벡터의 수이며, 여기에서는 8이다. 한편, 앞서 구한 최적의 전자밀도의 RMSE는 0.497이었으며, RSM 모델 (0.493)과 비교할 때, 서로 비슷한 예측 능력을 보이고 있다. 전자온도와 플라즈마 전위에 대해, 최적의 GRNN 모델은 모두 동일한 폭변수 0.5에서 결정이 되었다.

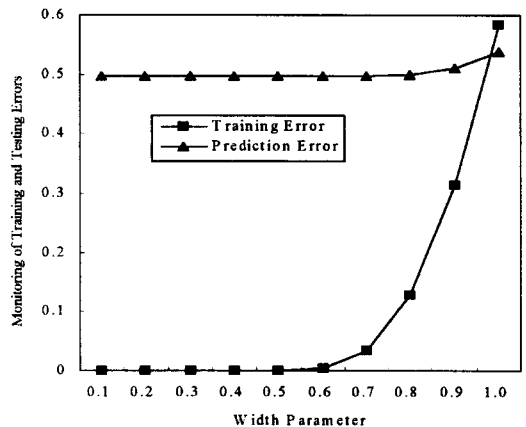


그림 3. 폭변수에 따른 GRNN 전자밀도 모델의 예측편차 비교.

Fig 3. Comparison of prediction error of GRNN electron density model with spread parameter.

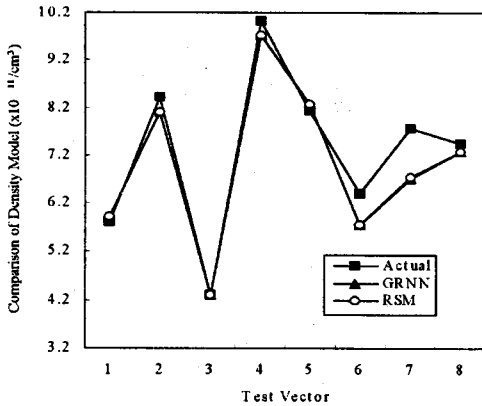


그림 4. 전자 밀도 비교.

Fig 4. Comparison of electron density model.

표 2. GRNN과 RSM 모델.

플라즈마 변수	GRNN		RSM	
	학습	테스트	Fit	테스트
전자밀도	0.152	0.497	0.239	0.493
전자온도	0.06	0.279	0.296	0.279
플라즈마전위	0.131	0.716	1.249	0.716

Table 2. Comparison of GRNN and RSM.

표 2는 GRNN과 RSM 모델간의 학습과 예측 편차를 비교하고 있다. 예측치들에 대한 상대적이 퍼센티지 (percentage) 편차를 비교하면, 두 모델이 서로 비슷한 예측능력을 보이고 있음을 알 수 있다. 한편, 전자온도 모델의 경우 두 모델 예측치들이 실제 측정치에서 상당히 벗어나고 있었으며, 그 원인이 학습데이터의 분포와 밀접한 연관성이 있으리라 사료되어 각 학습 데이터에 대한 표준 편차 (standard deviation error)를 구하여 서로 비교하였다. 전자온도의 경우 그 편차는 0.294이었으며, 전자 밀도와 플라즈마 전위에 대해서는 각각 3.313과 1.701이었다. 여기에서 알 수 있듯이 전자온도는 상대적으로 매우 작은 표준 편차를 보이고 있으며, 따라서 표준편차가 작은 데이터에 대해서 GRNN은 최적의 모델을 만드는데 어려움을 겪고 있음을 알 수 있다

### 5. 결론

본 연구에서는 GRNN를 이용하여 반도체 제조 공정 시뮬레이션을 위한 플라즈마를 모델링 하였다. 공정 변수와 플라즈마 변수와의 체계적인 모델링을 위해 Box-Wilson 실험 계획표가 이용되었으며, GRNN의 예측능력은 통계적인 RSM 모델에 비교되었다. GRNN은 RSM 모델과 비슷한 예측 능력을 보이고 있어, RSM 모델을 대체할 수 있는 새로운 신경망 모델로서의 가능성을 충분히 입증했다. 한편, 두 모델 방식은 전자온도 모델의 경우에서 알 수 있었듯이 학습데이터의 표준편차가 너무 작을 경우 모델 예측능력이 크게 저하되는 특성을 보이고 있어, 이를 극복하기 위한 연구가 요구되고 있다.

### 감사의글

본 연구는 전남대 신진교수 연구비에 의해 지원되었으며, 이에 감사사를 드립니다.

### 참고 문헌

- [1] P. L. G. Ventzek, T. J. Sommerer, R. J. Hoekstra, and M.J.Kushner, Two dimensional hybrid model of inductively coupled plasma sources for etching, *Appl. Phys. Lett.*, 63 (5), 1994.
- [2] R. A. Stewart, P. Vitello, and D. B.Graves, Two-dimensional fluid model of high density inductively coupled plasma sources, *J. Vac. Sci. Technol.*, B 12(1), p. 478-485, 1994.
- [3] B. Kim, G. Park, and C. Lee, A Neural network plasma model of semiconductor manufacturing equipment, 98 international conference on fuzzy systems, vol. I, p. 289-291, Aug. 1999.
- [4] C. D. Himmel and G. S. May, Advantages of plasma etch modeling using neural networks over statistical techniques, *IEEE Trans. Semicond. Manufact.*, vol. 6, p. 103-111, 1993.
- [5] E. A. Rietman and E. R. Lory, Use of neural networks in modeling semiconductor manufacturing processes: An example of plasma etch modeling, *IEEE Trans. Semicond. Manufact.*, vol. 6, no. 4, p. 343-347, 1993.
- [6] B. Kim and G. S. May, An optimal neural network process model for plasma etching, *IEEE Trans. Semicond. Manufact.*, vol. 7, no. 1, p. 12-21, 1994.
- [7] B. Kim and G. S. May, Reactive ion etch modeling using neural networks and simulated annealing, *IEEE Trans. Comp., Pack. and Manufact. Technol., C*, vol. 19, no. 1, p. 3-8, 1996.
- [8] B. Kim and K. H. Kwon, Modeling a magnetically RIE of aluminum alloy films using neural networks, *IEEE Trans. Semicond. Manufact.*, vol. 11, no. 4, p. 692-95, 1998.
- [9] B. Kim, K. H. Kwon, and S. H. Park, Characterizing metal-masked silica etch process in a CHF<sub>3</sub>/CF<sub>4</sub> inductively coupled plasma, *J. of Vac. Sci. and Technol.*, A, vol. 17, no. 5, p. 2593-2597, 1999.
- [10] B. Kim, J. H. Sun, C. J. Choi, D. D. Lee, and S. Y. S. Use of neural networks to model low temperature tungsten etch characteristics in SF<sub>6</sub> plasma, *J. of Vac. Sci. and Technol.*, A, March/April, vol. 18, no. 2, p. 417-422, 2000.
- [11] G. Box and N. Draper. *Empirical Model Building and Response Surfaces*, New York: Wiley, 1987.
- [12] D. F. Specht, "A general regression neural network," *IEEE Trans. Neural Networks*, 2(6), p. 568, 1991.