

레이디얼 베이스스 함수망을 이용한 플라즈마 식각공정 모델링

Modeling of Plasma Etch Process using a Radial Basis Function Network

박경영¹, 김병환^{1,a}

(Kyoungyoung Park¹ and Byungwhan Kim^{1,a})

Abstract

A new model of plasma etch process was constructed by using a radial basis function network (RBFN). This technique was applied to an etching of silicon carbide films in a NF_3 inductively coupled plasma. Experimental data to train RBFN were systematically collected by means of a 2^4 full factorial experiment. Appropriateness of prediction models was tested with test data consisted of 16 experiments not pertaining to the training data. Prediction performance was optimized with variations in three training factors, the number of pattern units, width of radial basis function, and initial weight distribution between the pattern and output layers. The etch responses to model were an etch rate and a surface roughness measured by atomic force microscopy. Optimized models had the root mean-squared errors of 26.1 nm/min and 0.103 nm for the etch rate and surface roughness, respectively. Compared to statistical regression models, RBFN models demonstrated an improvement of more than 20 % and 50 % for the etch rate and surface roughness, respectively. It is therefore expected that RBFN can be effectively used to construct prediction models of plasma processes.

Key Words : Plasma etching, Radial basis function network, Modeling

1. 서론

플라즈마 식각공정은 전자소자 제조를 위한 미세패턴 형성에 이용되고 있다[1-3]. 플라즈마 공정은 공정변수에 따른 플라즈마 특성의 고도의 비선형성으로 인해 그 예측모델을 개발하기가 매우 어렵다. 물리적 모델은 많은 가정(assumptions)을 포함하고 있어 모델 예측치의 신뢰도가 낮으며, 동시에 모델개발과 시뮬레이션에 소요되는 시간이 상당하여 실제 공정개발에 응용하기에는 적합하지 않다. 이를 극복하기 위한 다른 방식으로 신경망과 통계적 실험계획법을 이용한 예측모델이 개발되어

왔다. 신경망은 주어진 입력력간의 비선형성을 가정 없이 정확히 학습할 수 있는 능력을 보였으며, 식각과 증착공정의 모델링에 많이 응용되었다[4-8]. 신경망 중에서, 역전파 신경망(error back propagation neural network-BPNN)[9]이 가장 많이 응용되었으나, BPNN 모델개발에는 다수의 학습인자가 관여하고 있어 모델성능을 최적화하기가 매우 어렵다[8]. 특히 뉴런간의 연결하중치의 불규칙성(randomness)은 성능최적화를 더욱 어렵게 하고 있다. BPNN에 비해 그 최적화가 용이한 다른 신경망에는 Radial basis function network (RBFN) [10]이 있으나, 반도체 플라즈마 공정에서의 그 응용 사례는 매우 드물다. 최근 RBFN은 플라즈마 방전 특성, 즉 Langmuir probe를 이용하여 수집한 플라즈마 특성에 응용된 바가 있었다[11]. 당시에 RBFN 모델은 통계적인 회귀 모델과 비교하였으나, 두 모델간의 성능차이는 미미하였다. 이는 RBFN 모델의 학습에 영향을 주는 학습인자, 즉

1. 세종대학교 전자공학과
(서울시 광진구 군자동 98)
a. Corresponding Author : kbwhan@sejong.ac.kr
접수일자 : 2004. 9. 2
1차 심사 : 2004. 11. 4
심사완료 : 2004. 12. 13

패턴 units, width, 그리고 초기웨이트 (initial weight distribution-IWD)의 크기에 대한 최적화가 이루어지지 않는다는 것에 큰 이유가 있다.

본 연구에서는 식각공정의 모델링에 RBFN을 적용하며, 학습에 관여하는 세 학습인자의 영향을 실험적으로 최적화한다. 실험데이터는 유도 결합형 플라즈마(inductively coupled plasma-ICP)를 이용한 SiC 식각공정으로부터 수집하였으며, 통계적인 실험계획법에 의거하여 실험이 진행이 되었다. 모델에 이용된 식각특성에는 식각률과 표면거칠기가 있다. 식각률은 scanning electron microscope (SEM)을 이용하여 측정하였으며, 표면거칠기 측정에는 atomic force microscopy(AFM)이 이용되었다. 개발된 RBFN 모델들은 통계적인 회귀모델과 비교하여, 그 성능을 비교하였다.

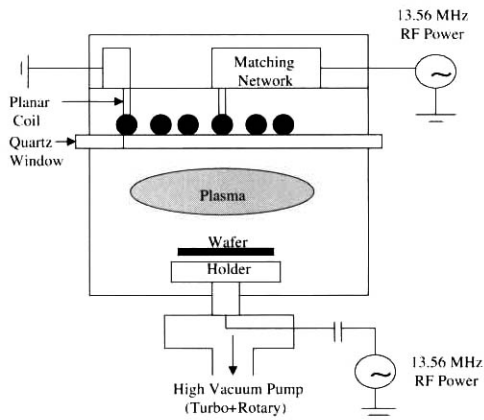


그림 1. ICP 식각 시스템의 개요도.

Fig. 1. A schematic diagram of ICP Etching system.

2. 실험

2.1 실험 데이터

ICP 식각 시스템은 그림 1에 도시되어 있다. 챔버 안에서 발생한 플라즈마는 석영유리에 의해 planar-coupled된 코일과 분리된다. 원통모양의 챔버는 각각 80 mm의 반지름과 40 mm의 높이를 가진다. 식각 가스를 공급하기에 앞서 챔버안은 두 펌프 (TUROVAC 3430MC와 Edward High Vacuum E2M40)를 이용하여 배기되며, 약 10^{-6} Torr의 기분압력을 유지한다. 가스유입은 질량유입제어기 (mass flow controller)로 정확히 조절되며, 바

라트론 gauge로 측정되는 생성압력은 throttle 밸브로 제어된다. 냉각수는 식각 공정 중 온도의 급격한 변화로부터 장비의 손상을 최소화하기 위해 적 혼더로 공급된다.

실험 패턴은 <0001> 방향성의 n-type, 2인치, 4H-SiC 웨이퍼로 제조 하였다. 웨이퍼의 열전도성은 $4.9 \text{ W/cm} \cdot \text{K}$ 이다. Cr negative mask를 이용하여 photoresist (PR)를 형성하였다. Magnetron sputtering을 이용하여 패턴된 PR 위에 Ni박막을 $0.18 \mu\text{m}$ 만큼 증착시킨 후, acetone으로 PR을 제거하여, Ni mask 층을 형성하였다. 그런 다음에, SiC 박막은 NF/CH_4 플라즈마로 식각 되었다. 식각 공정시간은 5분으로 설정하였다. 실험 변수로는 rf 소스파워, bias 파워, 압력, 그리고 NF_3 비를 이용하였다. NF_3 비는 NF_3 의 유량을 NF_3 와 CH_4 유량의 합으로 이루어진 전체 유량으로 나누는 것으로 정의된다. 총 유량은 30 sccm 로 고정되었다.

측정한 식각 특성에는 식각률과 표면 거칠기가 있었으며, 각 기 SEM과 AFM으로 측정하였다. 모델개발을 위해 통계적 2^4 전인자 실험계획법[12]을 이용하여 실험을 수행하였다. 실험계획표에 이용한 실험변수와 범위는 표 1에 나타나 있다. 결과적으로 수집된 16개의 데이터와 중앙점에 해당하는 하나의 실험데이터를 포함한 17개의 데이터를 이용하여 RBFN을 학습하였다. 학습된 모델의 적합도를 테스트하기 위해 추가적으로 16회의 실험을 수행하였으며, 따라서 모델개발을 위해 총 33회의 실험을 수행하였다. 16개의 테스트 데이터는 17개의 학습데이터에 포함되지 않은 데이터이다. 각 학습과 테스트 데이터의 입력 벡터 (vector)는 전술한 4개의 공정변수로 구성이 된다.

표 1. 실험변수와 범위.

Table 1. Experiment parameters and ranges.

| 공정변수 | 범위 | 단위 |
|-----------------|---------|-------|
| RF 소스전력 | 700-900 | Watts |
| RF 바이어스 전력 | 50-150 | Watts |
| 압력 | 6-12 | mTorr |
| NF_3 비 | 20 100 | % |

2.2 RBFN 신경망

RBFN의 구조가 그림 2에 도시되어 있다. 그림 2에서와 같이, RBFN은 입력층, 패턴층, 그리고 출력층으로 나누어진다. 입력데이터는 k-means 알고리즘[13]을 통하여 분류된다.

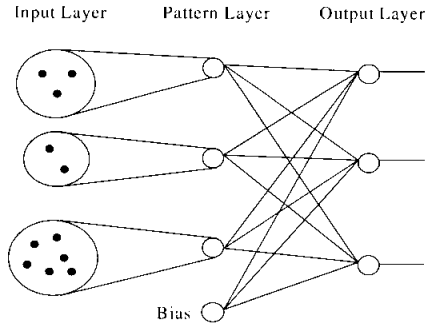


그림 2. RBFN에 대한 개요도.
Fig. 2. A schematic of RBFN.

우선 학습데이터를 구성하는 n개의 입력 벡터들로부터 주어진 k개의 패턴층 뉴런수에 해당하는 k개의 벡터들을 임의로 선택한다. 즉 n개의 학습 입력 벡터들은 k개의 클래스 (class)로 분류되게 된다. 선택된 k개의 벡터들은 초기 중앙 (center) 벡터로 일반적으로 지정된다. 나머지 (n-k)개의 입력 벡터들은 차례로 초기 중앙벡터들과의 유사성을 비교하게 되며, 이는 일반적으로 식 (1)로 정의되는 Euclidean distance에 의해 수행된다.

$$D = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (1)$$

비교결과 가장 유사성이 높은 패턴층 뉴런에 해당하는 클래스로 분류가 되며, 동시에 해당 뉴런의 초기 중앙값은 분류된 새로운 벡터와의 평균값으로 변화하게 된다. 이러한 분류과정이 완료된 후, k개의 패턴층 뉴런은 각기 서로 다른 안정화된 중앙 벡터를 가지게 되며, 이들 벡터들은 입력층과 패턴층 사이의 연결 웨이트를 제공한다.

패턴층 cluster units의 수는 주어진 입력데이터에 대해 자동적으로 결정될 수 있으나, 본 연구에서는 units의 수에 따른 성능을 평가하기 위해 일정범위내에서 그 수를 변화시켰다. 패턴층의 임의의 뉴런의 출력은 식 (2)로 구해진다.

$$g(x) = e^{-(x - \mu_j)^2 / \sigma_j^2} \quad (2)$$

여기서 x는 입력값, μ_j 은 k means 알고리즘으로 얻은 새롭게 분류된 영역 (receptive field라 불리는데) 지역의 center값이며, σ_j 는 receptive field의 width, g(x)는 j번째 뉴런의 output이다. 모델 성능은 σ_j 값에 상당한 영향을 받을 수 있으며, 따라서 RBFN 모델 최적화를 위한 중요한 학습인자이다. 그리고 패턴층과 출력층 weight은 일반적인 델타 규칙[6]을 이용하여 반복적으로 조정되어 이 식 (3)으로 표현이 된다.

$$\Delta w_i = \alpha (t_i - y_{ip}) x_i \quad (3)$$

여기서 t_i 는 target 값, y_{ip} 는 출력층에서 나오는 예측모델값, 그리고 x_{ip} 는 패턴층에서 나와 출력층으로 들어가는 입력값이다. 델타 규칙을 적용할 때, 초기웨이트는 불규칙하게 분포되어 있으며, 따라서 IWD에 따라서 모델성능이 영향을 받을 수 있다. 이상에서와 같이 RBFN의 성능에 영향을 미치는 학습인자에는 패턴 unit의 수, width, 그리고 IWD 등이 있으며, 그 영향을 고찰한다.

3. 결과

3.1 RBFN 모델 성능

식각률 모델: RBFN 모델의 성능을 학습인자의 함수로 고찰한다. 우선 식각률 데이터를 모델링하는 과정을 설명한다. Pattern units의 수를 2에서 8까지 증가시켰으며, 다른 width와 IWD를 각각 중간값인 2.5와 1.0으로 고정시켰다. 주어진 pattern units에서 IWD의 불규칙성을 고려하여 100개의 모델을 연속적으로 발생하였으며, 그 중 가장 작은 예측에러를 가진 모델을 선정하였다. 모델의 예측 성능은 테스트 데이터에 대해 root mean squared error (RMSE)를 적용하여 계산을 하였다.

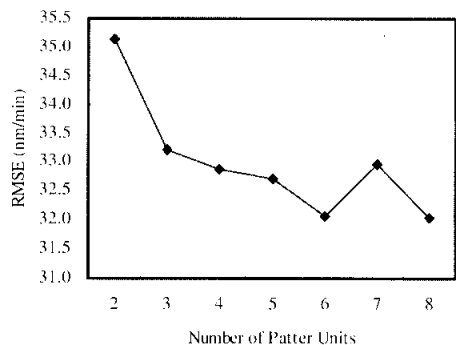


그림 3. Pattern unit 수에 따른 RMSE 변화.
Fig. 3. RMSE variations as a function of the pattern unit.

Pattern unit 수에 따른 RMSE의 특성이 위의 그림 3에 도시되어 있다. 그림 3에서와 같이, RMSE는 pattern unit 수의 증가에 따라 일반적으로 감소하고 있으며, pattern unit 수가 8일 때 32.0 nm/min으로 최적화 되었다.

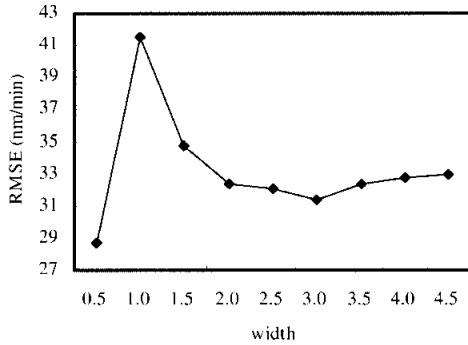


그림 4. Width 범위에 따른 RMSE 변화.

Fig. 4. RMSE variations as a function of the width.

다음은 width 변화에 따른 RMSE 변화를 알아보기 위하여 pattern unit 수를 이전에 최적화한 8에, 그리고 IWD를 중간값 2.5에 고정하고 width에 변화를 주며 RMSE의 변화를 살펴보았다. Width의 범위는 0.5-4.5이며, 0.5 간격으로 증가시켰다. 모델링 결과가 그림 4에 도시되어 있다. 그림 4에서와 같이, width가 0.5에서 1.0으로 증가되었을 때, RMSE는 상당히 증가하였으며, 1.5이상일 경우, RMSE는 width의 증가에 따라 감소되는 경향을 보인다. 그리고 2.0이상의 width에 대해서는 width의 증가에도 불구하고 RMSE의 변화가 작음을 알 수 있다. 결국 1보다 작은 width의 변화에 대해서 식각률 모델의 예측능력이 크게 변화를 보였으며, width가 0.5일 때 최적의 RMSE 28.6 nm/min를 가질 수 있었다. 이는 pattern unit 수만 최적화한 모델과 비교할 때 10 % 향상된 것이며, 이는 width의 변화가 모델의 학습능력에 상당히 영향을 미치고 있음을 의미한다.

마지막으로 IWD의 변화에 따른 모델의 RMSE 변화를 고찰한다. 다른 pattern unit 수와 width는 이전에 결정된 최적값, 즉 8과 0.5에 각기 고정하였다. IWD의 범위는 0.2-1.8이며, 0.2간격으로 증가시켰다. 모델링 결과가 그림 5에 도시되어 있다. 그림 5에서와 같이 IWD의 범위 0.6 1.2에서

RMSE는 매우 큰 변화를 보이고 있으며, 0.8 일 때 최적의 RMSE 26.0 nm/min을 얻을 수 있었다. 이는 이전의 모델에 비해 RMSE가 7 % 향상된 수치이며, pattern unit 수만을 최적화한 초기 모델에 비해서는 18.8 % 향상된 수치이다. 결국, 그림 3 5로부터, RBFN의 예측능력이 학습인자에 의해 상당히 영향을 받고 있음을 확인할 수 있었다.

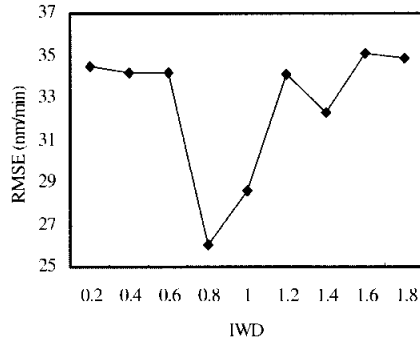


그림 5. IWD에 따른 RMSE 변화.

Fig. 5. RMSE variations as a function of IWD.

표면 거칠기 모델: 표면 거칠기 역시 식각률 데이터와 마찬가지로 세 학습인자의 함수로 모델링을 수행하였다. 세 인자의 실험범위도 식각률에서 적용했던 범위와 일치한다. 각 학습인자의 변화에서 결정된 모델의 RMSE와 최적 학습인자가 표 2에 정리되어 있다.

표 2. 표면거칠기 모델의 최적화된 성능.

Table 2. Optimized performance of surface roughness model.

| 학습인자 | 최적값 | RMSE(nm) |
|----------------|--------------|----------|
| Pattern Unit 수 | 2 | 0.104 |
| width | 4.0, 4.5 | 0.103 |
| IWD | 0.6, 1.0 1.8 | 0.103 |

표 2에서 같이 width와 IWD의 변화는 모델의 성능을 개선하는 데 영향을 미치지 못하고 있음을 알 수 있다. 결국, 식각률 모델성능은 각 학습인자에 상당히 의존 하였지만, 표면거칠기 모델의 경우, pattern unit 수에만 강한 영향을 받고 있음을 알 수 있었다.

3.2 통계적 회귀모델과의 비교

최적화된 RBFN 모델과 통계적 회귀 모델을 비교한다. 전형적인 회귀 모델은 다음과 같이 표현된다.

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{i^2} x_i^2 + \sum_{i=1}^k \sum_{j=1}^k \beta_{ij} x_i x_j \quad (4)$$

여기서 y 는 식각특성 예측치를 나타내며, β_i 와 β_{ij} 는 회귀변수, 그리고 x_i 는 실험변수를 표현한다. 지수 k 는 전체 실험변수의 수를 의미한다. RBFN 모델에 이용한 학습데이터를 통해 회귀변수를 추정하였다. 개발된 회귀모델의 테스트 데이터에 대한 예측 RMSE는 표 3에 정리되어 있다. 한편, 회귀모델에 대한 RBFN 모델의 성능개선도가 표 3에 함께 나타나 있다. 표 3에서와 같이, RBFN의 성능이 회귀모델에 비해, 두 식각 특성에 대해 20 % 이상 증진되었으며, 식각률의 경우 50 % 이상 증진되었다. 이는 RBFN의 예측성능이 기존의 통계적인 회귀모델에 비해 훨씬 우수함을 보여준다.

표 3. 통계적 회귀 모델과의 비교.

Table 3. Comparison with statistical regression models.

| | 통계적 모델 | RBFN 모델 | 향상률 |
|--------------------|--------|---------|------|
| Etch Rate (nm/min) | 55.2 | 26.1 | 52 % |
| Roughness (nm) | 0.136 | 0.103 | 24 % |

4. 결론

본 연구에서 RBFN을 이용하여 식각공정을 모델링하였다. 통계적인 실험계획법을 이용하여 SiC 식각공정데이터를 수집하였으며, RBFN의 성능을 세 학습인자 (pattern unit의 수, width, IWD)의 함수로 고찰하였다. 각 학습인자는 식각률 모델의 성능개선에 상당한 영향을 미쳤으나, 표면 거칠기 모델의 경우에는 그 영향이 미미 했다. 한편, 통계적인 회귀 모델에 비해 최적화된 RBFN 모델의 예측 성능은 상당히 증진되었으며, 따라서 비선형성이 매우 강한 플라즈마 공정의 모델링, 최적화, 그리고 제어에의 활발한 응용이 기대된다.

참고 문헌

[1] 백인규, 임완태, 이재원, 조관식, "BCl₃/Nc 혼합가스를 이용한 III-V 반도체의 고밀도 유도

결합 플라즈마 식각", 전기전자재료학회논문지, 16권, 12S호, p. 1187, 2003.

[2] 조남규, 구상모, 우용득, 이상권, "플라즈마 에칭으로 손상된 4H-실리콘 카바이드 기판위에 제작된 MOS 커패시터의 전기적 특성", 전기전자재료학회논문지, 17권, 4호, p. 373, 2004.

[3] 김장일, 권광호, 서경수, 남기수, "Al(Si, Cu) 합금막의 플라즈마 식각후 표면특성", 전기전자재료학회지, 9권, 3호, p. 291, 1996.

[4] B. Kim and S. Park, "Characterization of inductively coupled plasma using neural network", IEEE Trans. Plasma Sci., Vol. 30, No. 2, p. 698, 2002.

[5] B. Kim and K. Kwon, "Qualitative modeling of silica plasma etching using neural network", J. Appl. Phys., Vol. 93, No. 1, p. 76, 2003.

[6] B. Kim, S. S. Han, T. S. Kim, B. S. Kim, and I. J. Shim, "Modeling refraction characteristics of silicon nitride film deposited in a SiH₄/NH₃/N₂ plasma using neural network", IEEE Trans. Plasma Sci., Vol. 31, No. 3, p. 317, 2003.

[7] B. Kim, S. M. Kong, and B. T. Lee, "Modeling SiC etching in C₂F₆/O₂ inductively coupled plasma using neural networks", J. Vac. Sci. Technol. A, Vol. 20, No. 1, p. 146, 2002.

[8] B. Kim and S. Park, "An optimal neural network plasma model: A case study", Chemom. Intell. Lab. Syst., Vol. 56, No. 1, p. 39, 2001.

[9] D. E. Rumelhart, G. R. Hinton, and R. J. Williams, Parallel Distributed Processing, Cambridge, Vol. 1, MIT Press, 1986.

[10] J. Park and I. Sandberg, "Universal approximation using radial-basis-function networks", Neural Computation, Vol. 3, p. 246, 1991.

[11] B. Kim and S. Park, "Modeling of process plasma using a radial basis function network: A case study", Trans. Contr. Autom. Syst. Eng., Vol. 2, No. 4, p. 268, 2000.

[12] D. C. Montgomery, Design and Analysis of Experiments, John Wiley & Sons, Singapore, p. 270, 1991.

[13] C. M. Bishop, Neural Networks for Pattern Recognition, Oxford, England: Oxford University Press, 1995.