

펄스드 이온에너지 패턴의 신경망 시계열 모델링과 플라즈마 감시에의 응용

김수연, 김병환*
 세종대학교, 전자공학과

Time-Series Neural Network Modeling of Pulsed Ion Energy Pattern and Applications to Plasma Monitoring

Suyeon Kim, Byungwhan Kim*
 Sejong University, Electronic Engineering

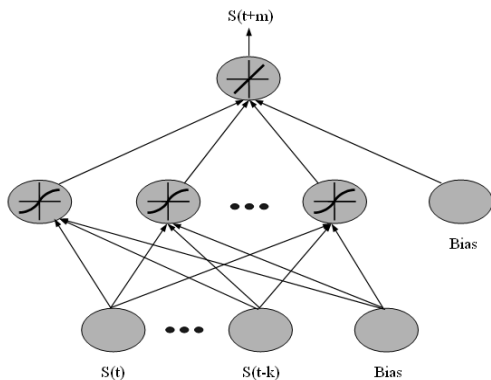
Abstract - 본 연구에서는 공정 중에 민감하게 반응하는 플라즈마로부터 수집되는 이온에너지 분포(IED : Ion Energy Distribution)와 시계열 신경망 모델링을 결합한 플라즈마 감시 기술을 개발하였다. NIEA(Non-invasive ion analyzer)를 이용하여 IED를 측정하였으며, 모델링에 사용된 신경망은 자기 상관 시계열 신경망(A-NTS : Auto-Correlated Neural Time-Series)이다. 모델 개발을 위한 학습과 테스트 데이터로는 Duty ratio 100%에서 수집한 IED를 이용하였으며, 개발된 모델의 감시 성능은 60%에서 수집된 IED로 평가하였다. 학습인자 k와 m의 범위는 각각 1-3 으로 총 9종류의 (k, m) 조합에 대해서 모델 성능을 평가하였다. 신경망 은닉층 뉴런수는 2-9의 범위에서 최적화하였다. 최적화된 모델은 (2, 3)과 뉴런수 2에서 구해졌으며, 0.335의 예측 에러를 보였다. 60% IED 데이터로 평가한 결과 플라즈마 고장예의 민감도는 62% 이상이었다. 이는 IED의 A-NTS 모델이 플라즈마 고장의 감시에 효과적으로 적용될 수 있음을 의미한다.

1. 서 론

플라즈마는 공정 조건에 따라서 민감하게 반응한다. 반도체 집적소자 제조에 있어서 공정의 질 (Quality)과 수율을 향상시키기 위해선 플라즈마 상태를 엄격히 감시하는 것이 요구된다. 플라즈마의 변화를 감시하기 위해 많은 in-situ 진단 센서들이 적용되었다. Optical emission spectroscopy [1-2], RF 임피던스 센서 [3], RF matching network monitoring 시스템 [4]등이 이에 해당한다. 이외에 ion energy analyzer를 이용하여 측정된 이온 에너지 분포(IED)를 플라즈마 상태를 감시하는데 이용할 수 있다. IED는 증착률이나 식각률과 같은 박막의 특성에 큰 연관성이 있다. 특히 박막 증착과 식각특성 메커니즘을 연구하는데 중요한 플라즈마 특성, 즉 이온밀도와 이온충격에너지 변화에 대한 자료를 제공한다. IED는 RF 전력이나 압력과 같은 공정 조건에 민감하게 변화를 보일 것으로 기대된다. 실시간 감시를 위해 시계열 신경망 모델링[1-2]이 보고된 바 있지만, IED의 신경망 모델링에 대한 연구는 보고된 바가 없다.

본 연구는 IED의 시계열 신경망 모델에 대한 것이다. IED의 데이터는 NIEA(Non-invasive ion analyzer)[5]을 이용하여 수집하였다. NIEA는 비침투식이기 때문에 실시간 공정 감시에 적합하다는 것은 이미 실험적으로 보고된 바 있다 [6].

Duty ratio라는 공정 변수를 변화시킴에 따른 플라즈마의 변화들을 알아보고자 하였다. IED 모델은 은닉층의 뉴런수와 시계열 신경망의 시간 변수들의 함수로 최적화 과정을 거쳤다.



<그림 1> A-NTS의 구조도

2. 본 론

2.1 실험데이터 수집

SiN 박막 증착 공정은 Pulsed-PECVD를 이용하였다. 원통형 챔버는 직경 0.27m, 높이 0.2m 크기이다. 챔버의 진공은 Turbo molecular pump, rotary pump 그리고 vacuum gauge에 의해 제어된다. 가스 유량과 공정 압력은 각각 mass flow controller들과 throttle valve에 의해 정밀하게 제어된다. 증착은 SiN₄-NH₃를 이용하였으며 각각 8sccm, 22sccm의 유량으로 흘러주었다. 증착시간은 5분이다. 13.56MHz의 소스 전력 500W를, 12.56MHz의 바이어스 전력 50W를 인가하였다. NIEA를 이용하여 Duty ratio 100%와 60% 각각의 IED 데이터를 수집하였다. NIEA는 High IED나 Low IED와 같은 다양한 이온 에너지의 분포를 제공하며, 이 중 본 연구에서는 Low IED를 모델링에 사용하였다.

2.2 A-NTS 구조와 동작

A-NTS의 기반인 BPNN(Back-Propagation Neural Network)[7]의 구조가 그림 1에 도시되어 있다. 은닉층의 출력은 식 (1)의 bipolar sigmoid 함수에 의해 결정된다.

$$out_{i,k} = \frac{1 - \exp\left(-\frac{in_{i,k}}{g_b}\right)}{1 + \exp\left(-\frac{in_{i,k}}{g_b}\right)} \quad (1)$$

$in_{i,k}$ 와 $out_{i,k}$ 는 k번째 층에서 i번째 뉴런으로의 가중된 입력과 출력을 나타낸다. g_b 는 bipolar sigmoid 함수의 기울기를 나타낸다. 또한 A-NTS의 출력은 식 (2)의 linear 함수에 의해 결정되며, g_l 은 linear 함수의 기울기이다.

$$out_{i,k} = in_{i,k}g_l \quad (2)$$

Weight의 변화는 식 (3)으로 표현되는 일반화된 델타 규칙[7]에 따라 조정된다.

$$W_{i,j,k}(m+1) = W_{i,j,k}(m) + \eta \Delta W_{i,j,k}(m) \quad (3)$$

여기서 $W_{i,j,k}$ 는 k번째 층의 i번째 뉴런과 (k-1)번째 층의 j번째 뉴런의 연결강도이다. $\Delta W_{i,j,k}$ 는 모든 입, 출력 쌍의 에러(E)를 최소화시키는 방향으로 계산된다. η 는 학습률을 의미한다.

그림 1에서와 같이 IED는 과거와 현재의 데이터로 미래의 값을 예측한다는 구조이다. 그래서 A-NTS 모델의 수행은 (k, m)의 조합에 따라 달라진다. k는 학습에 이용되는 과거의 데이터 수이며, m은 예측하게

<표 1> Conditions of Modeling Factors

Parameter	Ranges
Error Tolerance	0.1
Learning rate	0.01
Weight Deviation	±1
Gradient of Slope	1
k	1, 2, 3
m	1, 2, 3
NHN	2 - 9

미래의 시간을 가리킨다. 본 연구에서는 k와 m을 각각 1부터 3까지 변화를 준, 즉 총 9가지의 조합으로 모델링을 하였다. 또한 은닉층의 뉴런 수의 함수로도 A-NTS 모델링을 수행하였으며, 2에서 9까지 1씩 증가시켰다. 학습에 이용된 다른 모델인자들은 표 1에 정리되어 있다.

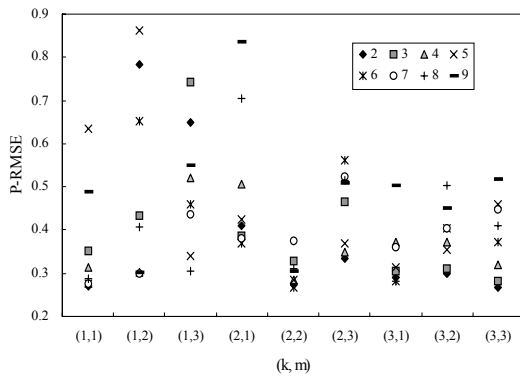
2.3 A-NTS 모델링

Duty ratio 100%에서 수집된 IED의 데이터로 A-NTS 모델을 만들었다. 모델 성능은 Duty ratio 60%에서 수집된 IED 데이터를 통해 그 감시 성능을 평가하였다. Duty ratio 100% IED 데이터는 학습과 테스트 데이터로 나누어졌으며 각각 66개의 패턴이 형성되었다.

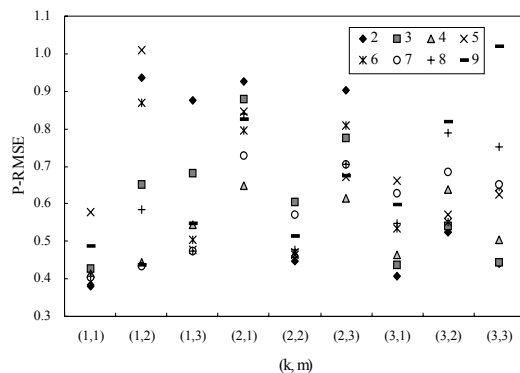
A-NTS 모델은 식 (4)와 같은 RMSE(Root Mean Square Error)가 계산된다.

$$RMSE = \sqrt{\frac{\sum_{j=1}^q (d_j - out_j)^2}{q}} \quad (4)$$

여기서 q는 총 테스트 데이터의 수이고, d_j와 out_j는 각각 목표의 출력값과 j번째 뉴런에서 계산되어진 출력값이다. RMSE는 학습데이터와 테스트 데이터에 따라 구해지며, 모델의 100% duty와 60% duty 테스트 데이터에 대한 RMSE (P-RMSE로 표현)들은 그림 3과 4에 도시되어 있다. 그림 3을 보면, RMSE가 대략 0.2-0.9의 범위를 갖는다. 조합 (1,1)-(2,1)에서 은닉층 뉴런수에 따라 RMSE가 크게 변화하는 것을 볼 수 있다. A-NTS의 성능은 (k, m)의 조합과 은닉층 뉴런수의 두 함수에 대해 동시에 최적화 과정을 거친다. 100%의 duty에서 개발된 모델의 최소 P-RMSE는 (3, 3)의 (k, m)과 뉴런수 2에서 얻을 수 있었다. 이때의 최소 P-RMSE 0.2673이다.



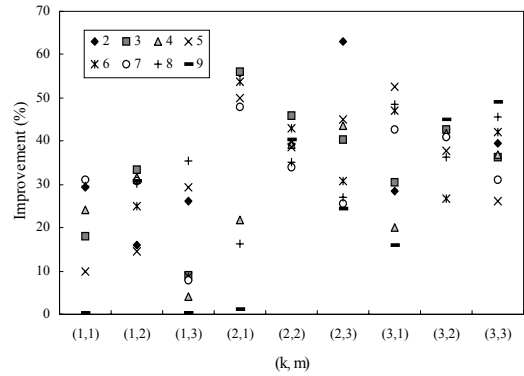
〈그림 3〉 조합과 은닉층 뉴런수에 따른 테스트 에러 (Duty 100%)



〈그림 4〉 조합과 은닉층 뉴런수에 따른 테스트 에러 (Duty 60%)

또한 조합과 은닉층 뉴런수에 대한 A-NTS 모델의 감시 성능을 평가하기 위해 최적화된 모델의 민감도를 계산 하였다. 그림 4의 P-RMSE에 대해 식 (5)를 이용하여 민감도를 계산하였으며 그 결과가 그림 5에 도시되어 있다.

$$Improvement = \left| \frac{P-RMSE_{100\%} - P-RMSE_{60\%}}{P-RMSE_{60\%}} \times 100 \right| \quad (5)$$



〈그림 5〉 Duty ratio 60%에 대한 duty 100% 모델의 민감도

그림 5를 보면, 전체 82개의 모델 중 8개를 제외한 모델들은 공정이상을 탐지하기에 충분한 10% 이상의 민감도를 보이고 있다. 10%이하에 해당하는 모델개발은 피해야하며, 이는 모델최적화의 중요성을 의미한다. 그림 5에서 최고의 민감도는 (2, 3)과 뉴런수 2에서 발생하였으며, 그 값은 62.91%이다. 이 모델의 P-RMSE는 그림 3에서 0.335이다. 그림 3에서 얻은 최적 모델에 대한 민감도는 39.36%이며, 따라서 그림 5에서 구한 모델에 비해 민감도가 낮다. 결국, 60% duty 데이터를 감시하는 데에는 (2, 3)과 2에서 구한 모델이 가장 적합한 것을 알 수 있다.

3. 결 론

본 연구에서는 BPNN을 이용하여 Low IED의 A-NTS 모델을 개발하였다. A-NTS 모델 성능은 은닉층의 뉴런수, (k, m) 조합, 그리고 모델 민감도의 함수로 최적화하였다. 소수의 A-NTS 모델을 제외한 다른 모델들은 플라즈마 고장을 감지하기에 충분한 탐지력을 보였으며, 개발된 모델 방식은 Wafer-to-Wafer 플라즈마 감시에 효과적으로 응용될 것으로 기대된다.

[감사의 글]

이 논문(저서)은 2007년도 정부재원(교육인적자원부 학술연구 조성사업비)으로 한국학술진흥재단의 지원을 받아 수행된 연구임 (KRF-2007-511-D00232).

[참 고 문 헌]

- [1] S. J. Hong and G.S. May, "Neural network based time series modeling of optical emission spectroscopy data for fault detection in reactive ion etching", Proc. SPIE., Vol. 5041, pp. 1-8, 2003.
- [2] B. Kim and M. Kwon, "Optimization of PCA-applied in-situ spectroscopy data using neural network and genetic algorithm," Appl. Spectros. Vol. 62, pp. 73-77, 2008.
- [3] S. Bushman, T. F. Edgar, I. Trachtenberg, "Radio Frequency Diagnostics for Plasma Etch Systems", J. Electrochem Soc., Vol. 144, No. 2, pp. 721-732, 1997.
- [4] B. Kim, C. Lee, "Monitoring plasma impedance match characteristics in a multipole inductively coupled plasma for process control", J. Vac. Sci. Technol. A, Vol. 18, No. 1, pp. 58-62, Jan/Feb 2000.
- [5] Non-invasive Ion Energy Analyzer Plasma Monitoring System, NIEA™ System, PLASMART. Inc.
- [6] M. A. Sobolewski, "Real-time, noninvasive monitoring of ion energy and ion current at a wafer surface during plasma etching", J. Vac. Sci. Technol. A, Vol. 24, No. 5, pp. 1892-1904, 2006.
- [7] D. E. Rummelhart and J. L. McClelland, "Parallel Distributed Processing", Cambridge, M.I.T. Press, 1986.