

## 3C5) 신경망모델을 이용한 대기 중 오존농도 예측에 관한 연구

### Study on the Prediction of Atmospheric Ozone Concentration Using Neural Network

이종범 · 천태훈 · 구운서<sup>1)</sup>

강원대학교 환경과학과, <sup>1)</sup>안양대학교 환경공학부

#### 1. 서 론

통계모델은 지역적인 평균농도를 산출하는데 어려움이 많으며, 과거의 측정 자료에 의존하므로 발생빈도가 적은 고농도 오존발생 등 특이한 상황에 대한 예측에는 오차가 큰 단점이 있다. 따라서 오존농도에 영향을 미치는 중요한 기상요소 및 대기질 자료를 심도 있게 선정하고 적절한 활용법을 개발하는 것이 오존농도 예측 정확도 향상에 중요하다.

신경망 분석은 인간의 두뇌를 모델로 하여 등장하였다. 학습이란 생명체가 가지는 중요한 특성으로 다양한 분야에서 살아있는 것을 연구하는 과정에서 필연적으로 제기되어 학습 행동연구에 적절한 모형을 제공한다. 인공 신경망은 노드의 집합체로 구성되며, 입력노드는 입력신호를 받고 출력 노드는 신력신호를 보내며 임의의 개수의 중간 계층에는 중간노드들이 포함된다.

본 연구에서는 수도권지역을 대상으로 정확도 높은 오존예보모델을 개발하고 현재 시행중인 오존예보체에 적용할 수 있는 정확도 높은 예보시스템을 개발하고자 한다. 또한 수도권 지역의 예보시스템을 기초로 하여 우리나라 광역시에 대한 오존예보 시스템을 개발하고자 한다.

#### 2. 연구 방법

본 연구에 신경망모델의 실행을 위해서 사용된 변수는 종속변수 1개와 독립변수 25개로 총 26개이다. 독립변수로 사용한 자료는 기상측정자료, 대기질 측정자료, 기상모델 예보자료의 3가지로 대별된다. 종속변수는 예보 목표일(D-day)의 일최고 오존농도를 예보함을 목적으로 하므로 종속변수로 취하였다. 시간대는 01시부터 24시까지의 최고농도( $O_3max-d$ )이다. 독립변수는  $O_3maxy$ ,  $NO_2y$ ,  $NO_2maxy$ ,  $COy$ ,  $SO_2y$ ,  $PM_{10}$ ,  $RHy$ ,  $Tmaxy$ ,  $Tminy$ ,  $SRy$ ,  $WSd$ ,  $VISBy$ ,  $Prechry$ ,  $RHd$ ,  $Td$ ,  $Tmaxd$ ,  $SRd$ ,  $SRmaxd$ ,  $WSd$ ,  $MMHd$ ,  $U29d$ ,  $V29d$ ,  $U31d$ ,  $V31d$ , 편차변수를 사용하였다.

대상 지역은 서울시 4개 지역, 경기도 5개 지역과 광역시를 대상으로 하여 전국적으로 15개 지역으로 선정하였다. 각 지역별로 해당 기상대의 자료를 사용하고, 대기질은 그 지역 내의 모든 농도 평균치를 사용하였다.

대상기간은 오존농도가 발생하는 하계를 대상으로 2004년부터 2006년까지의 3년 동안 각각 4월 1일부터 9월 31일까지 학습기간으로 선정하였으며 2007년 5월 1일부터 8월 31일까지의 123일간을 평가기간으로 선정하였다.

신경망 모델 실행 시 표준화 과정을 통해 전 처리를 하여 입력 데이터로 사용하였다. Multilayer Perceptron(MLP)을 사용하였으며, 활성화함수와 결합함수를 사용하였다. 학습 알고리즘으로는 Quasi-Newton 알고리즘을 사용하였다. Quasi-Newton 알고리즘은 중간 크기의 신경망에 적합한 방법으로 적은 메모리를 요구하는 알고리즘이다.

#### 3. 결과 및 고찰

신경망 예보 대상일은 2007년 5월 1일부터 8월 31일까지로 하였다. 2004년부터 2007년 예보 대상일의 하루 전까지의 자료로 학습시키고, 2007년의 예보 대상일의 일 최고농도를 예측 후 예보대상일의 측정치로 모델의 오차를 평가하였다.

대상지역은 15개 지역이며, 그 중 경기북동지역과 서울 서부지역의 결과를 예시 하였다. 그림 1은 평

가 대상기간인 2007년 5월 1일부터 8월 31일까지의 실측자료와 예측자료의 시계열이다. 경기북동지역은 6월 18일의 고농도를 모사하지 못하였으며 그 외 고농도는 비교적 잘 모사하였다. 서울 서부는 대체적으로 고농도를 잘 모사하였다.

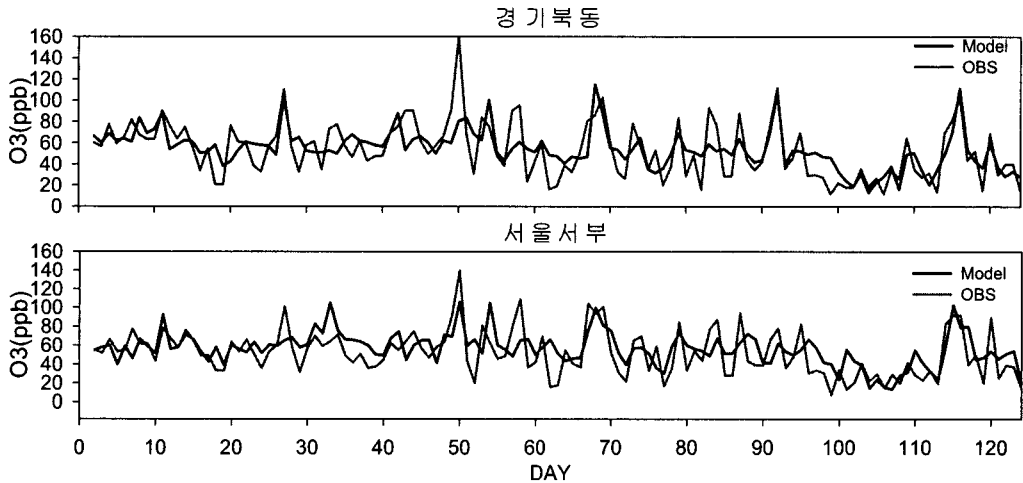


Fig. 1. Comparison of observed(thin line) and estimate(thick line) O<sub>3</sub> concentration on May 1-30, 2007.

오존측정 농도자료에 대한 신경망모델 결과를 평가한 결과는 표 1에 제시하였다. r 값의 경우 경기북동지역은 0.69이며, 서울서부는 0.59이다. RMSE 경우 경기북동지역은 18.96ppb, 서울 서부는 20.18ppb로 다소 큰 오차범위를 보인다. 서울 서부 지역에서 40ppb 이하의 오존농도를 예측함에 있어 40ppb 이상으로 예측하여 나타나는 현상으로 사료된다. 이것은 일최고 농도가 40ppb 이상인 기간만을 선정하여 학습하였기 때문에 40ppb 이하를 잘 예측하지 못한 것으로 사료된다.

Table 1. Model evaluation statistics for ozone concentrations for the period of May 1-30, 2007. (unit : ppb)

	NO. of data	avr. obs.	avr. model	r	RMSE	MB
경기북동	123	53.58	54.82	0.69	18.96	1.24
서울서부	123	52.34	56.77	0.59	20.18	4.43

## 사 사

본 연구는 환경부의 차세대 핵심환경기술개발 사업(오존예보시스템 개발, 과제번호 012-062-033)의 일환으로 수행 되었으며, 이에 감사드립니다.

## 참 고 문 헌

- 국립환경과학원 (2006) 대도시 대기질 관리방안 조사연구V-광화학 대기오염의 생성과정규명 및 저감대책 수립.
- 배화수 등 (2005) SAS Enterprise Miner를 이용한 데이터마이닝.
- 서울특별시 보건환경연구원 (2003) 서울시에 맞는 오존예보시스템 개발.