

DR9) 퍼지전문가시스템과 신경망을 이용한 오존예보시스템 개발 Development of an Ozone Forecasting System using the Fuzzy Expert System and Neural Network

허정숙, 최형욱, 김동술

경희대학교 환경학과 대기오염연구실 및 환경연구소

1. 서론

오존은 극도로 활성적이며, 오존을 생성시키는 동일한 원인물질 (precursor)에 의해 소멸되는 2차오염물질로, 오존 오염문제는 복잡하고 문제 해결이 어렵다. 또한 전구물질의 오염원이 있는 풍하쪽 (down wind)에 위치한 지역에서 높은 오존농도가 관측되는 등, 지역에 따라 다양한 형태로 발생한다 (Jorquera *et al.*, 1998). 실제, 서울지역에서 오존주의보 발령횟수는 1998년 18회로 매년 증가하고 있는 실정이며, 인구 밀집지역인 대도시에서 고농도 오존 예보문제는 더욱 절실하다. 현재, 오존예측 모델 개발은 국내 · 외적으로 활발히 이루어지고 있으며, 국내에서는 김용국과 이종범 (1994)이 중회귀모델과 신경망모델을 이용하여 일 최고 오존농도를 예측한 바 있으며, 박진수 등 (1997)은 Wavelet 변환방법을 이용하여 단기오존농도를 예측하였다. 국외의 연구로는 신경망을 이용한 오존예측 방법이 여러 연구자에 의해 진행되고 있다 (Gardner and Dorling, 2000; Jorquera *et al.*, 1998; Yi and Prybutok, 1996).

허정숙과 김동술 (1999)은 이미 군집분석 (cluster analysis)과 분산주성분분석 (disjoint principal component analysis)을 응용한 형태인식법 (pattern recognition)을 이용하여 고농도 오존오염의 형태별 순수군집 (homogeneous class)을 분류한 바 있다. 본 연구의 목적은 서울지역의 고농도 오존을 진단, 예측하는 전일 오존예보시스템을 개발하고자 한다. 이를 위해, 퍼지전문가시스템 (fuzzy expert system)을 이용하여 고농도오존 발생 가능성을 진단하는 퍼지규칙 (fuzzy knowledge rule base)을 개발하고자 하였으며, 다음으로 오차역전파학습알고리즘 (back-propagation)을 이용한 다층퍼셉트론신경망 (multi perceptron neural network)에 의하여 고농도 오존예측모델을 개발하고자 하였다.

2. 연구 방법

본 연구는 환경부에서 입수한 1989년부터 1996년까지 8년간 대기오염자동측정망의 시간별 자료 (SO₂, CO, NO₂, O₃)와 기상청의 지상기상자료 (풍향, 풍속, 일사량, 온도, 습도) 및 오산 상층기상자료 (500hpa 고도면의 온도, 풍속, 풍향) 등 총 12개 변수를 이용하였다. 또한, 지상기상자료 중 온도와 습도, 상층기상자료는 내삽하여 이용하였고, 대상자료는 12개 변수 모두가 갖추어진 자료만을 선택하였다. 대상지역은 오존에피소드가 가장 빈번하며, 지역적 특성을 가장 잘 대표하는 4개 지역 즉, 광화문, 쌍문동, 구로동 및 방이동지역을 대상으로 하였다. 본 연구의 목적이 고농도 오존을 추정하는 것이므로, 오존농도 80 ppb (미국 EPA의 National Ambient Air Quality Standards 8시간 기준농도) 이상인 경우만을 대상자료로 하였다.

그림 1은 본 연구의 전체적인 연구 진행 방향을 제시한 것이다. 첫 번째 단계로 각 변수들의 특성을 조사하기 위하여 단변수분석을 실시하였다 (허정숙과 김동술, 1997). 다음단계로 대상자료를 풍향을 근거로 분류하였고, 분류된 각각의 대상자료에 대하여 변수별 자료변환을 통하여 군집분석과 주성분분석을 실시하였다. 허정숙과 김동술 (1999)이 패턴인식 분석과정을 자세히 기술하였다. 분류된 고농도 오존오염 패턴을 근거로 퍼지규칙을 작성하였다. 퍼지의 원리는 0 (완전한 참) 아니면 1 (완전한 거짓)이라는 단순하고 딱딱한 개념에서 벗어나 0과 1사이의 애매모호함의 정도를 값으로 표현하는 것이다. 퍼지논리 (logic)는 어떤 사실의 정도 (0~1)를 소속함수 (membership function)로 대응할 수 있다. 퍼지전문가시스템은 인공지능의 한 분야로 특정 전문영역과 전문가적 문제 해결 방식에 관한 전문가적 지식 (knowledge)을 분류시키는 것이다. 본 연구에서는 이미 전단계의 분석과정에서 분류된 각 대상지역별 순

수한 군집들 (homogeneous classes)의 통계값을 이용하여 소속함수를 결정하였고, 퍼지규칙을 작성하였다.

본 연구에서는 4개 대상지역의 퍼지규칙을 상호 비교하여 서울지역의 공통적인 오존오염형태를 재분류하였다. 또한, 이 공통적인 군집에 해당되지 않은 군집들은 해당지역의 고유한 오존오염형태로 정의하였다. 지금까지의 분석은 오존농도 80 ppb 이상이 나타난 현재 시점만을 대상으로 오존오염형태를 분류하였다. 다음단계는 퍼지규칙에 의해 재분류된 서울시의 공통적인 군집과 각 지역별 군집을 대상 자료로 해당 군집의 현재시점부터 24시간 전 자료를 추적하여 데이터베이스화하였다. 이 자료는 전일 오존예측모델의 기본자료가 된다. 24시간 전부터의 1차오염물질의 거동은 다음일 고농도 오존발생에 거의 영향이 희박할 수 있으므로, 전일의 고농도 오존예측 가능성 진단에 관한 퍼지규칙의 작성은 온도, 습도, 풍속, 일사량, 상층풍속 및 상층온도 등 6가지 변수만을 대상으로 하였다. 또한, 6가지변수의 예측 시간대별 현재상황의 통계값 뿐만 아니라 현재 시점을 기준으로 1시간전 자료와의 변화폭 즉, 변화속도를 변수로 첨가하여 12개 변수를 가진 퍼지규칙이 만들어졌다. 오존농도 80 ppb 이상이 나타난 자료 (1,078 case) 중 고농도 오존이 가장 많이 발생하는 시간대는 오후 3시로 약 23 % (248 case)를 차지하였다. 따라서, 예측시스템 개발시 고농도 오존예측 목표시점은 내일 (tomorrow) 오후 3시로 정하였고, 오존예측이 시작되는 시점은 금일 (today) 4시부터 1시간별 고농도오존 가능성을 진단한다. 실제 오존예측값이 본격적으로 제시되는 시점은 아침시간부터 시작하는 것으로 결정하여 오존농도 80 ppb 이상이 나타난 현재 시점으로부터 7시간 전 시간을 신경망모델 가동시점으로 정하였다. 따라서, 금일 아침 8시부터 예측농도를 확인 할 수 있다. 아침 8시부터 퍼지규칙의 변수는 SO₂, CO, NO₂, O₃, 기상 자료를 포함한 12개 변수가 모두 이용되었고, 현재 예측시점에서 1시간전과 아침 8시 자료와의 변화속도를 변수로 추가하였다. 이 퍼지규칙은 단지 고농도 오존 발생 가능성만을 진단하는 것이다.

오존농도를 예측값으로 제시하는 모델은 신경망 기법을 이용하였다. 신경망의 기본개념은 수학적 모델에 의하여 인간 뇌의 구조와 기능을 모방한 것이다. 본 연구에서는 은닉층의 수가 하나 이상을 갖는 다층퍼셉트론을 이용하였고 학습은 오차역전파 알고리즘을 이용하였다. 또한, 학습회수는 30000번 이상으로 하였다. 신경망모델 개발시 이용된 변수는 퍼지규칙의 변수와 동일하며, 다만 1시간전 오존농도와 아침 8시의 오존농도가 변수로 추가되었다.

3. 결과 및 고찰

퍼지전문가시스템을 이용하여 고농도 오존오염의 형태를 재분류시킨 결과, 공통규칙은 8개 군집으로 결정되었다. 이 공통규칙은 서울시의 일반적인 오존오염 현상으로 정의될 수 있다. 또한, 광화문지역의 고유한 오존오염현상은 10개 군집 즉, 10개 퍼지규칙으로 결정되었으며, 쌍문동지역은 11개 퍼지규칙, 구로동은 8개 규칙, 방이동 지역은 20개 퍼지규칙을 확보하였다. 특히, 방이동지역은 서울지역에서 가장 고농도 오존이 빈번하게 발생된 지역으로 그 만큼 오존오염현상도 다양한 것으로 조사되었다.

신경망의 원리가 과거자료의 학습을 통해 미래를 예측하는 것이므로, 입력자료의 특성에 따라 다양한 결과를 얻을 수 있다. 그러므로, 퍼지전문가시스템을 통해 고농도 오존 가능성이 있는 것으로 진단된 사례를 학습하여 예측 모델을 개발하는 것은 매우 좋은 예측력을 얻을 수 있다. 또한, 본 연구에서 확보된 사례는 약 1,368개 (군집당 24시간전부터 고농도오존 발생시점까지의 자료 포함)로 각 시간대별 사례에 해당될 경우, 고농도 오존 예측력을 충분히 상승시킬 수 있다. 신경망 모델은 금일 아침 8시부터 가동되기 시작하여, 시간별로 예측이 이루어진다. 개발된 시간별 예측모델의 정확도는 공통퍼지규칙인 경우, 아침 8시의 오후 3시 오존 예측오차는 8개 군집 평균 6.8 %였으며, 아침 9시의 예측오차는 7.2 %, 아침 10시의 예측오차는 5.1 %, 아침 11시의 경우 5.2 %, 아침 12시의 경우 5.5 %, 오후 1시의 예측오차는 4.9 %, 1시간전인 오후 2시의 예측오차는 4.1 %였다. 오후 3시 예측시점이 가까울 수록 예측력은 증가하였다. 또한, 광화문지역의 경우도 마찬가지로, 시간대별 예측오차가 각각 7.3 %, 3.2 %, 2.0 %, 2.2 %, 2.1 %, 2.1 %, 1.7 %로 예측율은 92 % 이상이었다. 쌍문동, 구로동 및 방이동지역에서의 오존예측율은 각각 98 %, 98 %, 92 % 이상으로 분석되었다.

본 연구에서 개발된 오존예보모델의 장점은 확보된 퍼지규칙에 해당되지않은 고농도 오존 사례가 발생 할 경우, 새로운 데이터베이스가 확보되고, 퍼지규칙이 추가되어 모델을 보강할 수 있으며, 통계분

석의 약점인 자기수정 (self-correction) 문제를 보완 할 수 있다.

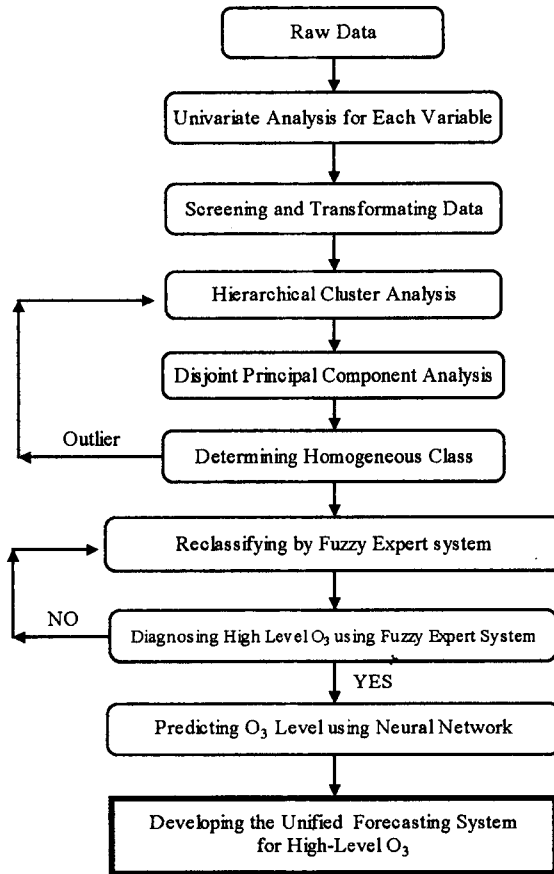


Figure 1. Flow chart for the scope of study.

참 고 문 헌

- 김용국과 이종범 (1994) 하계의 일최고 오존농도 예측을 위한 신경망 모델의 개발, 한국대기보전학회지 10(4), 224-232.
- 박진수, 김신도, 조용준 (1997) Wavelet transform을 이용한 오존 특성분석 및 다원 ARIMA 모형을 통한 예측, 1997년도 한국대기보전학회 추계학술대회 요지집, 89-90.
- 허정숙과 김동술 (1999) 고농도오존의 추정을 위한 오존오염 형태인식, 1999년도 한국대기환경학회 춘계 학술대회, 81-83.
- 허정숙과 김동술 (1997) 고농도오존의 추정을 위한 예비연구, 1997년도 한국대기보전학회 추계학술대회, 91-92.
- Yi J.S. and Prybutok V.R. (1996) A neural network model forecasting for prediction of daily maximum ozone concentration in an industrialized urban area, Environmental Pollution, 92(3), 349-357.
- Gardner M.W. and Dorling S.R. (2000) statistical surface ozone models: An improved methodology to account for non-linear behaviour, Atmospheric Environment, 34, 21-34.
- Jorquera H., Perez R., Cipriano A., Espejo A., Letelier M.V., and Acuna G. (1998) Forecasting ozone daily maximum levels at Santiago, Chile, Atmospheric Environment, 32(20), 3415-3424.