

4B3) 대기질 통계예보 모델에 관한 연구

The Study of Air Quality Forecasting Model Based on Statistical Methods

윤원철 · 구운서 · 윤희영¹⁾

안양대학교 환경공학과, ¹⁾(주)에니텍

1. 서론

대기오염 예보의 목적은 인간의 건강과 생활환경을 보전하기 위하여 고농도의 대기오염이 발생하기 전에 대책을 수립하기 위한 것이며, 대기오염 예보체제는 미국이나 영국, 일본과 같은 나라들에서는 이미 운영되고 있으나, 대기오염이 날로 심각하게 대두되고 있는 우리나라에서도 대기오염 예보제의 적극적인 도입이 요구된다.

통계모델은 측정된 오염물질을 근거로 오염물질의 농도에 영향을 주는 기상인자(풍향, 풍속, 일사량, 기온, 강수량 등)와의 통계적 상관관계를 유추해내는 경험적 모델의 일종이다. 즉 관측 기상자료 및 측정 대기오염도 자료를 독립 변수로 하고, 특정 오염물질 농도 값을 종속변수로 하여 독립변수와 종속변수간의 상관관계를 통계적 기법을 이용하여 구하는 것이다.

본 연구에서는 고농도 미세먼지 예보모델을 개발하기 위해 대기오염물질자료와 기상자료를 이용하여 미세먼지의 중회귀모델과 의사결정모델을 개발하였다.

2. 연구 방법

본 연구의 대상지역은 수도권지역으로 서울, 경기, 인천지역이다. 대상기간은 2001년부터 2004년까지이며 대기오염물질자료와 기상자료를 이용하였다. 미세먼지 예보시에 사용하게 될 자료를 살펴보면, 대기오염측정소 자료의 5가지 오염물질(SO₂, PM₁₀, NO₂, O₃, CO)과 기상청 지상관측자료 중 기상인자(풍속, 습도, 기온, 강수량), 오산지역의 고층기상자료 중 기상인자(압력, 고도, 기온, 이슬점온도, 풍향, 풍속)이다.

미세먼지의 변화에 미치는 대기오염물질과 기상요소들의 영향을 분석하기 위하여 대상기간의 시계열 분석과 상관분석을 실시하고, 대기오염물질과 기상요소들의 독립성을 분석하여, 상관성이 높은 인자들에 대해 예보모델 입력자료로 선정하였다. 이러한 입력자료를 이용하여 중회귀 및 신경망 모형을 통하여 각 영향인자에 의한 미세먼지의 예측식을 각각 구하고, 또한 의사결정 모형을 통하여 최종적인 예보값을 결정하여 미세먼지 농도를 예측하였다.

3. 결과 및 고찰

식 1은 미세먼지의 회귀모델식을 나타낸 것이다. 2001년부터 2003년도의 자료를 이용하여 학습을 시켰으며, 2004년 자료를 적용하여 모형의 예측성을 평가하였다.

$$\text{식 1 : PM}_{10}\text{-target} = -18.264 + 0.502 \cdot \text{PM}_{10} + 3.342 \cdot \text{SO}_2 + 0.213 \cdot \text{HM} - 0.01 \cdot \text{RN}_{24} + 0.191 \cdot \text{NO}_2$$

그림 1은 회귀모델의 예측값과 관측값의 상관도를 나타낸 것으로 상관도값이 0.88로 높게 나타났으며, 표 1은 회귀모델의 예측값과 관측값의 지수별 일치도를 나타낸 것으로 지수별 일치

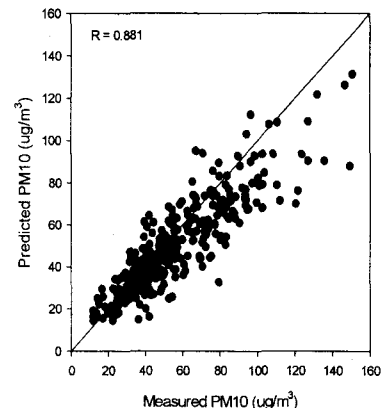


Fig. 1. Correlation between predictions and measurements.

Table 1. Comparison of statistical model predictions with measurements.

관측값 예측값	~50	50 ~100	100 ~150	150 ~200	총합계	
~50	179	42			221	일치 290(80.8%) 상향 11(3.1%) 하향 58(16.1%)
50~100	10	106	15		131	
100~150		1	5	1	7	
150~200						
총합계	189	149	20	1	359	

80.8%, 상향예측은 3.1%, 하향예측은 16.1%로 나타났다.

그림 2는 의사결정모델의 모형을 나타낸 것이고, 표 2는 의사결정 모델의 일치도를 나타낸 것이다. 의사결정모델로 적중시킨 $100\mu\text{g}/\text{m}^3$ 이하 범위에 들어올 수 있는 경우는 317개중 310개로 97.8%의 적중률을 나타내었고, $100\mu\text{g}/\text{m}^3$ 이상 범위에 들어올 수 있는 경우는 6개 중 5개로 83.3%의 적중률을 나타내었다. 그 외 모델로 결정할 수 있는 범위는 총 36개 자료 중 $100\mu\text{g}/\text{m}^3$ 이하가 28개, $100\mu\text{g}/\text{m}^3$ 이상이 8개의 자료로 나타났다.

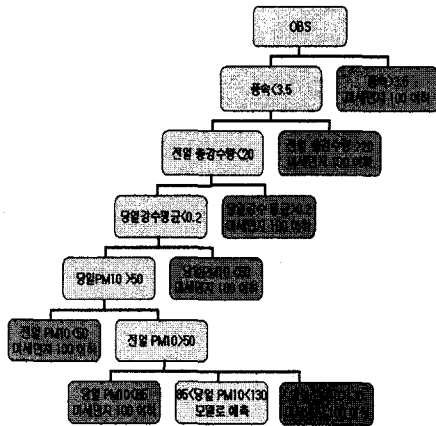


Fig. 2. Decision tree model.

Table 2. The application results of decision tree model.

의사결정 구분	100 이하	100 이상	모델결정 범위	총 계
일평균농도				
100 이하	310 (97.8%)	1 (16.7%)	28 (77.8%)	339
100 이상	7 (2.2%)	5 (83.3%)	8 (22.2%)	20
총 계	317 (100%)	6 (100%)	36 (100%)	359

사 사

본 연구는 환경부의 "차세대핵심환경기술개발사업(과제명 : 대기질 예경보 시스템 상용화 Package 개발)"으로 지원으로 수행되었으며, 이에 관계자 여러분께 감사드립니다.

참 고 문 헌

- 구윤서, 권희용, 윤희영(2003). 통계모델을 이용한 실시간 오염도 예보 시스템 개발, 대기환경학회 추계 학술대회 논문집
- Patricio Perez, Jorge Reyes(2002). Prediction of maximum of 24-h average of PM10 concentrations 30h in advance in Santiago, Chile, Atmospheric Environment 36, 4555-4561
- Robeson, S.M. and D.G. Steyn(1990). Evaluation and comparison of statistical forecast models for daily maximum ozone concentration, Atmospheric Environment, 24B(2), 303-312.