

현업운영 가능한 서울지역의 일 최고 대기오염도 예보모델 개발 연구

A Study on the Development of Operable Models Predicting Tomorrow's Maximum Hourly Concentrations of Air Pollutants in Seoul

김 용 준

기상청 응용기상국

(1996년 10월 7일 접수, 1997년 1월 6일 채택)

Yong-Joon Kim

Applied Meteorology Bureau, Korea Meteorological Administration

(Received 7 October 1996; accepted 6 January 1997)

Abstract

In order to reduce the outbreaks of short-term high concentrations and its impacts, we developed the models which predicted tomorrow's maximum hourly concentrations of O₃, TSP, SO₂, NO₂ and CO. Statistical methods like multi regressions were used because it must be operated easily under the present conditions. 47 independent variables were used, which included observed concentrations of air pollutants, observed and forecasted meteorological data in 1994 at Seoul and its surrounding areas. We subdivided Seoul into 4 areas coinciding with the present ozone warning areas. 4 kinds of seasonal models were developed due to the seasonal variations of observed concentrations, and 2 kinds of data models for the unavailable case of forecasted meteorological data.

By comparing the R² and root mean square error(hereafter 'RMSE') of each model, we confirmed that the models including forecasted data showed higher accuracy than ones using observed only. It was also shown that the higher the seasonal mean concentrations, the larger the RMSE. There was no distinct difference between the results of 4 areal models. In case of test run using 1995's data, the models predicted well the trends of daily variation of concentrations and the days when the possibility of outbreak of high concentration was high. This study showed that it was reasonable to use those models as operational ones, because the R² and RMSE of models were smaller than those of operational/research models such as in South Coast Air Basin, CA, USA.

Key words : operable prediction models, multi-regression, maximum concentrations of air pollutants

1. 서 론

1.1 연구 배경

1960년대 이후 지속된 산업화와 도시화로 많은 양의

대기오염물질이 배출되므로써 전국의 대도시 및 공단지
역의 대기질이 악화된 상태이다. 1980년에 환경청이 설
립되면서 저황연료 공급, 청정연료 사용 의무화, 자동차
배기가스 저감대책, 먼지 저감대책 등 여러가지 대기질
개선대책을 지속적으로 실시한 결과, 일부 대도시 지역

Table 1. Total numbers of exceeding the short-term environmental standards of each pollutants in each area (ref. Table 2) from 1994 to 1995. The numbers in paranthesis are the short-term environmental standards.

	O ₃ (100 ppb /hr)	TSP (300 µg/m ³ /day)	SO ₂ (250 ppb /hr)	NO ₂ (150 ppb /hr)	CO (25 ppm /hr)
North-West	28	4	0	2	0
North-East	43	8	0	1	0
South-West	39	4	0	21	0
South-East	48	1	0	1	0

의 연평균 이산화황과 총먼지 오염도가 장기환경기준치 이내로 개선되었다. 그러나 많은 지역에서 이산화질소와 오존의 장기 오염도는 개선되지 않고 있으며, 미세먼지 오염도는 환경기준치를 초과하고 있다. 더구나, 오존, 먼지, 이산화황과 이산화질소 등은 계절 및 지역에 따라 단기환경기준치를 자주 초과하고 있는 실정이다(환경부, 1996b; 환경부의 매월 대기오염도 발표자료, 표 1 참고).

따라서, 대기오염물질에 의한 인체, 생태 및 구조물 피해를 줄이기 위해서는 장기 오염도 저감대책을 지속적으로 추진하는 것은 물론 단기 고농도 발생과 인체 노출을 줄이기 위한 대책도 실시해야 한다. 다행히도, 단기 고농도 발생빈도가 높은 오존 피해를 줄이기 위해 오존 경보제를 서울지역에서는 1995년부터, 인천지역에서는 1996년부터 실시하고 있으며 1997년 이후에는 전국 6대 도시까지 확대 실시할 예정이다. 그러나 현재는 오존 농도가 경보치(120 ppb/시간)를 초과한 후에 경보를 내리므로, 오염도 저감과 시민의 노출 방지 대책을 실시할 만한 시간이 부족하여 경보발령 효과가 기대 이상으로 크지 않을 것으로 판단된다. 만일 오존을 비롯한 오염물질들의 고농도 발생 가능성을 하루 전에 예보한다면 관련 기관들은 대책을 미리 계획하여 실시하고 시민 스스로도 노출을 억제할 수 있으므로 그 피해를 더 줄일 수 있을 것이다.

1.2 대기오염 예보 현업 및 국내 연구 사례

일 최고 오염도는 통계기법 또는 확산모델을 이용하여 예측할 수 있다. 현재 우리는 일간 및 상세 배출원자료, 화학변환자료, 침적자료 등 기초자료가 부족하여 확산모델을 현업운영하기는 어려운 실정이다. 매시간 평균 대기오염도 자료, 매시간 또는 3시간 단위의 기상 관측자료와 일간 예보자료가 산출되므로 통계기법을 이용한 예

보모델의 현업운영은 가능하다. 대기오염도와 기상자료, 기타 오염도에 영향을 주는 인자들의 과거자료로 적절한 예측모델을 만들어 예보하는 통계기법은 과거와 현재의 행태에 큰 변화가 없을 때 비교적 정확한 예측이 가능하면서도 모델 개발 및 운영이 단순한 장점이 있다. 그래서 통계 예보모델을 중심으로 연구를 수행하였다.

현재, 통계기법을 사용한 단기 대기오염 예보모델을 현업운영하고 있는 대표적인 곳은 미국 California주의 South Coast Air Quality Management District이다(South Coast Air Quality Management District, 1994; Cassmassi, 1987, 1994, 이후에 'SoCAB (South Coast Air Basin)' 이라 함). Los Angeles County의 15개 지점, Orange County의 7개 지점, Riverside County의 10개 지점 및 San Bernardino County의 10개 지점을 대상으로 1시간 평균 오존, 8시간 평균 일산화탄소, 24시간 평균 미세먼지와 이산화질소의 일 최고 오염도를 예보한다. 또한, 예보 오염도를 사용하여 PSI 지수도 산출한다. SoCAB 외에도 미국의 워싱턴과 필라델피아, 일본 동경도와 오사카, 프랑스 파리를 비롯한 세계 여러 지역에서도 오염도 예보체계를 구축하기 위한 연구 및 현업화가 추진중이다(환경부, 1996c)

국내의 통계기법과 관련된 대기오염 연구는 1980년대 중반부터 시작되었다. 1985년에 국립환경연구원에서는 대기오염 경보체계 구성을 위한 조사연구를 하였으며, 1986년에는 대기오염도 예측기법 개발에 관한 연구를 수행하였다. 학계에서도 대기환경기준 오염물질의 오염도 특성을 분석하였고 기상인자와의 상관관계를 분석하는 등(전병일 등, 1994, 1995; 이기원 등, 1993; 신찬기 등, 1992; 김신도 등, 1992; 박일수 등, 1991; 송동웅과 김영만, 1991; 정용승과 정재섭, 1991; 김민영 등, 1989; 나진균, 1989; 김상진 등, 1988; 정용 등, 1986) 예측기법 개발을 위한 많은 기초 연구들이 이루어졌다. 동적확률모델을 이용한 SO₂ 농도 예측(김해경, 1991a, b), 회귀모형을 이용한 오존농도 예측(허정숙과 김동술, 1993; 이선기, 1991), 신경회로망을 이용한 오존농도 예측(김용국과 이종범, 1994) 등 통계기법을 이용한 SO₂와 오존의 단기예측 연구도 있었다. 그러나 기존의 국내 연구들이 오염물질의 농도 변화와 기상 상태와의 상관관계 분석에 집중되었던 반면 오염도 예측을 위한 연구는 그리 많지 않았다. 또한, 오존과 SO₂에 대한 오염도 예보 연구가 일부 있었으나 개발된 연구결과가 실용화되지 못하고 단순한 연구 수준으로 끝났으며, 다른 오염물질에 대한 예측연구 사례는 발견

할 수 없었다.

1.3 연구 내용

본 연구에서는 현실적으로 현업운영 가능성이 가장 높은 통계기법을 사용하여 대기오염물질의 일 최고 오염도를 하루 전에 예보하는 대기오염 예보모델을 개발하였고 외국 현업운영모델 결과 및 국내 연구결과와 비교, 검증하였다. 대상 오염물질은 오존(O₃), 총먼지(TSP), 이산화황(SO₂), 이산화질소(NO₂), 일산화탄소(CO) 등 5개 대기환경기준 오염물질이며, 대상 지역은 서울지역을 오존경보제 실시지역과 같이 4개 지역으로 세분한 지역이다(표 2). 본 논문의 2장에서는 1994년과 1995년 서울지역의 대기오염물질의 변동 특성을 분석하였고, 3장에서는 1994년도의 오염도 및 기상자료를 사용하여 오염 예보모델을 개발하였으며, 4장에서는 개발된 예보모델을 1995년 사례로 시험모의하여 검증하였다. 또한, 5장에서는 결론으로 모델의 개선방안과 현업운영 추진방안을 제시하였다.

2. 대기오염물질의 변동 특성

2.1 연변화 특성 분석

오염도의 연변화 경향, 단기환경기준치를 초과할 가능성이 높은 기간과 지역 등을 파악하여 통계 예보모델 개발의 기초로 활용하기 위해 5개 오염물질의 일 최고 오염도의 연변화를 분석하였다. 서울지역 20개 자동측정소에서 측정된 1994~1995년의 1시간 평균 자료 중 일 최고 오염도를 현행 오존경보제의 4개 지역(표 2)으로 세분하여 분석하였고, 자세한 분석 결과는 환경부(1996a)에 나타내었다.

오존은 겨울철에 낮고 여름철에 높은 뚜렷한 연변화 경향을 보였다. 6~9월에 고농도 오존이 자주 발생하고 매일매일의 변동이 크게 나타나 기상 조건에 따라 오존 오염도가 크게 변하고 있음을 나타내었다. 지역간에는 뚜렷한 차이가 없이 비슷한 형태를 보였다. 총먼지는 겨울철에는 높고 여름철에는 낮은 연변화 경향을 보였으나 뚜렷하지는 않았다. 일 변동이 크게 나타났고 지역간 차이도 보였다. 1년 내내 고농도 오염도가 나타날 가능성이 있으며, 특히 10~4월에 가능성이 높게 나타났다. 이산화황은 겨울철에 높고 여름철에 낮은 연변화 경향을 보였으며, 10~4월에 고농도 가능성이 높게 나타났다. 이 기간 동안 일 변동이 크며, 지역간에는 뚜렷한 차이가 없이 유사한 형태를 보였다. 이산화질소와 일산화탄소도 대체로 겨울철에 높고 여름철에 낮은 연변화

경향을 보였으나 변동폭은 크지 않았다. 두 오염물질 모두 지역간에는 뚜렷한 차이를 보이지 않았으며, 이산화질소는 일 변동이 크게 나타났다.

2.2 일변화 특성 분석

각 오염물질의 일 최고 오염도가 나타나는 시각을 파악하여 경보 발령시 고농도 발생 가능성이 높은 시각도 함께 알릴 수 있게 하기 위해 매시간 평균 오염도의 일변화를 분석하였다.

오존은 햇빛이 강해지면서 점점 증가하여 14~16시에 최고 오염도를 보이고 그 이 후 점차 줄어들다가 새벽 4~5시에 약간 증가하였다 아침 7~8시에 최저값을 보이는 bi-modal 분포를 나타냈다. 각 지역이 같은 변화 경향을 보이고 지역간 농도값 차이도 크지 않았다. 총먼지도 새벽 5시에 최저값이고 그 이후 점점 증가하여 9~10시에 최고값을 보인후 점점 감소하다 야간에 다시 증가하는 bi-modal 분포를 나타냈다. 각 지역이 비슷한 일변화 경향을 보이거나 최고값 발생 시각과 오염도가 다르게 나타났으며, 북서지역(NW)이 가장 높은 오염도를 보였다. 이산화황은 아침 7~9시 사이에 최고값을 보이고 그 이 후 점차 줄어들어 오후에 최저값을 보이고 야간에 다시 높아지는 일변화 경향을 보였다. 각 지역의 일변화 경향은 비슷하나 최고값 크기와 발생 시각, 일간 변동폭이 서로 다르게 나타났다. 이산화질소는 오전 8~10시, 오후 20~22시에 최고값이며, 두 극값의 크기가 비슷한 bi-modal 분포를 보여 두 번의 고농도 발생 가능성을 나타냈다. 각 지역의 오염도 크기와 변화 경향이 비슷하게 나타났다. 일산화탄소는 오전 8~10시에 최고값을 보이고, 점점 줄어들어 오후에 최저값을 보이다가 다시 증가하여 야간에 높은 값을 보였고, 지역간에 오염도 크기에는 차이가 있지만 일변화 경향은 비슷하게 나타났다.

3. 예보모델 개발

3.1 예보 대상지역과 오염물질

예보 대상지역은 서울지역을 현재 실시되고 있는 오존경보제의 북서(NW), 북동(NE), 남서(SW), 남동(SE) 지역으로 세분한(표 2) 지역이다.

예보 오염물질은 5개 환경기준 오염물질이며(환경부, 1996b), 다음날의 1시간 평균의 일 최고 오염도이다. 본 연구의 목적이 현업운영이기 때문에 실제 예보 시각이 합리적이면서도 현실적인 시각으로 결정되어야 하며, 이에 따라 일간 기준 시간도 결정되어야 한다. 오염 경

Table 2. Predicting areas.

Area	Included Districts	Included Monitoring Stations
North -West	Eunpyoung, Seodaemun, Mapo, Yongsan, Choong, Chongno	City Hall, Bulgkwang, Hanam, Mapo, Namgajwa
North -East	Tobong, Kangbuk, Sungbuk, Tongdaemun, Sungdong, Kwangjin, Choongryang, Nowon	Myunmog, Shinseol, Kilum, Ssangmun, Kuyui, Sungsoo
South -West	Kangseo, Yangchun, Youngdeungpo, Tongjak, Kwanak, Kuro, Keumchun	Munrae, Shinlim, Hwagok, Kuro, Oryu
South -East	Seocho, Kangnam, Songpa, Kangdong	Taechi, Chamsil, Banpo, Bangi

보발령에 따른 피해저감대책을 충분한 시간을 가지고 실시하기 위해서는 적어도 하루 전에 예보가 이루어져야 한다. 또한, 모델 입력변수의 이용가능 시각, 예보 담당자의 근무 시간, 신문 등 언론매체의 기사 마감 시각, 각 오염물질의 최고 오염도 발생가능 시각 등이 고려되어야 한다. 각 오염물질들의 최고값 발생 시각은 8~10시, 14~16시이므로 이 시간대에 예보한다면 예보된 최고농도가 오늘 또는 내일의 값인지 애매모호하므로 이 시간대 예보는 피하는 것이 좋다. 현재 대기오염도는 매시간, 관측 기상자료는 매시간 또는 3시간 간격으로 산출되며, 기상청의 정기 일기예보는 9시와 16시에 이루어진다. 24시간 근무체계가 아닐 경우 예보 담당자의 일상 근무시간은 9~18시이다. 그리고 언론매체의 기사 마감 시각 등을 고려할 때, 17:00를 기준으로 예보하는 것이 적합하다고 판단되었다. 따라서, 일 기준 시간을 오늘 17:00에서 부터 다음날 16:59까지로 하였다.

3.2 예보모델의 개발

현실적으로 현업운영 가능한 대기오염 예보모델은 오염도와 기상자료 만을 사용할 수 있는 통계모델이다. 사용가능한 통계기법이 여러가지이지만(김용국과 이종범, 1994), 모델이 단순하여 개발 및 운영이 비교적 쉽고 설명변수(독립변수)와 반응변수(종속변수)간 관계를 설명하기가 쉬우며 계산시간이 적은 중회귀방정식을 이용하였다. 회귀분석은 어떤 변수(반응변수 또는 종속변수)의 변화를 다른 변수들(설명변수 또는 독립변수)로 설명 또는 예측하려는 방법으로, 독립변수가 2개 이상인 모델을 중회귀모델이라 한다. 일반적인 중회귀 모델은

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \dots + \beta_p x_{pi} + \epsilon_i, \quad i=1, \dots, n \quad (1)$$

과 같으며, 여기에서

- β_0, \dots, β_p : 회귀계수
- x_{1i}, \dots, x_{pi} : 설명변수(독립변수)
- y_i, \dots, y_n : 반응변수(종속변수)
- $\epsilon_i, \dots, \epsilon_n$: 오차항

이다.

배출량과 지형효과가 비슷하면서도 2.1절에서와 같이 오염도 변화가 유사한 계절별로 모델 개발하였다. 각 계절의 기간은 겨울철은 1월 1일~2월 28일, 봄철은 3월 1일~5월 31일, 여름철은 6월 1일~8월 31일, 가을철은 9월 1일~11월 30일이다. 또한, 예보 기상자료의 사용이 불가능한 경우에 대비하여 오염도 자료와 관측 기상자료만을 사용한 경우(NOW)와 예보 기상자료까지 함께 사용한 경우(FORE)의 두가지 모델을 개발하였다. 따라서, 오염물질별(O₃, TSP, SO₂, NO₂, CO), 계절별(WI, SP, SU, AU), 지역별(NW, NE, SW, SE), 사용자료별(NOW, FORE)로 예보모델을 개발하였다.

통계예보모델 개발에 사용된 독립변수는 1994년도의 오염도와 기상자료이다. 오염도 자료는 이미 측정, 분석이 완료된 오늘의 일 최고 오염도로, 서울, 경기지역내 37개 관측지점에서 측정된 5개 기준오염물질의 오염도를 서울의 4개 지역(NW, NE, SW, SE), 서울주변의 서부(WE), 남부(SO), 동부(EA) 지역으로 세분하여 지역평균한 오염도이다. 각 지역으로부터의 수송효과를 고려하기 위해 사용된 변수들이므로 값의 크기가 다양하고 여러 개인 지역내 각 지점의 최고농도보다는 지역내 대표값이라 할 수 있는 지역평균 값이 더 의미가 있다. 또한, 각 지점 값을 그대로 사용할 경우 지역 평균값보다 자료의 오차 가능성이 높아 예보 정확도를 떨어뜨릴 가능성도 높아진다. 그래서 각 지점의 최고값을 그대로 사용하지 않고 지역평균한 값을 사용하였다. 기

Table 3. Lists of independent variables.

Data		Independent Variables
Pollution		O ₃ (NW, NE, SW, SE, WE, SO, EA)
		TSP (NW, NE, SW, SE, WE, SO, EA)
		SO ₂ (NW, NE, SW, SE, WE, SO, EA)
		NO ₂ (NW, NE, SW, SE, WE, SO, EA)
		CO (NW, NE, SW, SE, WE, SO, EA)
Meteorology	Observed	daily mean wind speed, U and V component winds daily mean temperature, cloud coverage, and mixing ratio maximum and minimum temperature total solar radiance and precipitation surface pressure difference (Seoul-Kangwha, -Inchun, -Yangpyung, -Suwon)
	Forecasted	wind speed, U and V component winds maximum and minimum temperature precipitation, cloud coverage

상자료는 서울·경기지역 기상청과 산하기관에서 정기 관측한 자료에서 산출한 관측자료와 서울·경기지역 예보자료이다. 각 지역의 통계 예보모델 개발에 사용된 독립변수 개수는 총 49개이며, 그 목록을 표 3에 나타내었다.

각 모델의 중요 독립변수는 STEPWISE 기법으로 선택하였고(유의수준 0.15), 상용 통계 package인 「SAS」를 활용하였다. 각 모델의 최종 step 수, 결정계수(R²), C(p), 자유도, 오차 등과 선택된 독립변수와 그 계수값은 환경부(1996a)에 나타내었다.

3.3 예보모델의 적합도 분석

개발한 각 예보모델의 적합도를 파악하기 위해 결정계수(R²)와 RMSE(root mean square error)를 분석하였다. 표 4는 각 모델의 선택된 독립변수 개수(N), 결정계수, RMSE의 오염물질별, 사용자료별 연평균 값으로, 일반적으로 NOW 모델보다는 FORE 모델이 더 많은 개수의 독립변수를 사용하였다. 추가된 독립변수의 대부분이 예보 기상자료로 예보 기상자료가 모델 정합에 중요한 역할을 함을 보였다. 또한, 모든 FORE 모델이 NOW 모델보다 결정계수는 크고 RMSE는 작게 나타나 관측 기상자료만 사용할 때보다는 예보자료까지 함께 사용한 경우에 더 우수한 적합도를 보였다. 지역별로는 연평균 결정계수와 RMSE에 큰 차이가 없었으며, 계절별로는 결정계수는 유사한 값이었으나 평균 오염도가 큰 계절에 큰 RMSE를 보였다. 그러나 각 오염물질별, 지역별, 계절별로 다른 결과를 보이는 경우도 있었다(각 오염물질별, 지역별, 계절별 모델과 분석 결과는 환경부

Table 4. Annual means of numbers of selected independent variables (N), R² and RMSE (ppb, µg/m³) of models.

Pollutants	Models	N	R ²	RMSE
O ₃	NOW	6.1	.57	9.56
	FORE	7.6	.69	8.39
TSP	NOW	5.9	.53	39.25
	FORE	8.0	.65	33.91
SO ₂	NOW	6.2	.57	10.78
	FORE	8.1	.73	8.53
NO ₂	NOW	9.1	.65	9.33
	FORE	12.1	.76	7.92
CO	NOW	11.4	.63	5.87
	FORE	12.9	.76	4.71

(1996a)에 자세하게 나타내었음).

본 연구에서 개발한 통계예보모델들의 적합도가 현실적으로 사용가능한 수준인지를 판단하기 위해 현재 유사하지만 좀 더 복잡한 일 최고 오염도 예보모델을 현업운영하고 있는 SoCAB의 결과와 비교하였다. 직접 비교 가능한 오염물질은 오존으로, 본 연구의 FORE 모델과 SoCAB 모델의 연평균 결정계수는 각각 0.69, 0.59, RMSE는 각각 8.39, 34.96 ppb로(Cassmassi, 1994), FORE 모델이 더 나은 결과를 보였다. 비록 두 인자의 비교만으로는 어느 모델이 더 우수하다고 확언할 수는 없지만, SoCAB 모델 결과와 비교할 때, NOW 및 FORE 모델을 오존의 일 최고 오염도 예보모델로 사용할 수 있을 것으로 판단되었다. 기타 오염물질에 대해서는 비교할 수 있는 자료가 없었다.

4. 예보모델 검증

1994년 자료를 사용하여 개발한 각 오염물질의 통계 예보모델을 1995년 사례에 적용하여 각 모델의 예보 성능을 검증하였다. 사용자별, 지역별, 계절별 시험예보 결과의 최저, 최고, 평균 오염도를 측정값과 비교하였고, 각 사례에 대한 RMSE, NMSE (normalized mean square error), MRE (mean relative error) 등을 분석하였다. RMSE, NMSE와 MRE는 관측값과 예보값 사이의 분산 정도를 나타내는 척도로, 각각의 식은 다음과 같다.

$$RMSE = [\sum(P_i - O_i)^2/n]^{1/2} \quad (2)$$

$$NMSE = \frac{[\sum(P_i - O_i)^2/n]^{1/2}}{(\sum P_i/n) (\sum O_i/n)} \quad (3)$$

$$MRE = \sum \frac{(P_i - O_i)/(P_i + O_i)}{n} \quad (4)$$

여기에서, P_i 는 예보 오염도, O_i 는 측정 오염도, n 은 예보 날짜수이다.

표 5는 각 예보모델을 시험예보한 연평균 결과이다. 각 오염물질별 NOW와 FORE모델의 평균값(MEAN), 최소값(MIN), 최대값(MAX)의 연간 그리고 지역별 평균한 값이 관측된 값(OBS)들과 유사한 크기로 개발

된 모델들이 관측값들을 잘 예측하고 있음을 보였다. 표 4의 모델 정합도 분석에서와 같이, 오존을 제외한 기타 오염물질의 경우 예보 기상자료를 함께 사용한 FORE 모델이 관측 기상자료만을 사용한 NOW 모델보다 더 우수한 결과를 보였다. 그러나 오존은 모델 정합도 분석과는 달리 NOW 모델이 더 작은 RMSE를 보였지만 그 차이가 크게 나타나지 않아 어느 모델이 더 우수하다고 단언할 수 없었다. 각 오염물질의 RMSE는 평균 오염도에 비례하고(0.34~0.55배) 1994년 모델 정합도 분석(표 4)의 RMSE보다 1.2~2.6배 정도 크게 나타났다. 연평균 MSE는 모두 양의 값으로 본 모델들이 전반적으로 측정값보다 높게 예보하는 경향을 보였다. 지역별로는 비슷한 오차를 보였고, 계절별로는 평균 오염도가 낮은 계절에는 작은 오차를 보였다.

그림 1은 북서지역(NW)의 측정 및 예보된 일 최고 오존 농도로, NOW와 FORE 모델 모두 오염도의 상승과 하강 경향, 고농도 오존 발생일 등을 잘 예보하여 현업운영 모델로 사용하여도 될 것으로 나타났다. 그러나 두 모델 모두 MRE는 양의 값으로 전반적으로 높게 예보하는 경향이 있으며, 예보된 극값이 관측된 극값보다 낮게 나타나는 경향을 보였다. 그래서 고농도 발생 가능성이 높은 계절인 6~9월의 오염도의 변동폭이 관측값보다 작은 경향이 있었다.

극값의 예보 특성을 분석하기 위해 측정값과 예보값을 비교하여 그림 2에 나타내었다. 70 ppb 이하의 농도 발생시에는 예보값이 측정값과 잘 일치하는 경향을 보였다. 그러나, 이 지역에서는 90 ppb 이상의 고농도 발생

Table 5. Annual means of minimum, maximum and average concentrations, RMSE (ppb, $\mu\text{g}/\text{m}^3$), NMSE and MRE of test runs.

Pollutants	Models	MIN	MAX	MEAN	RMSE	NMSE	MRE
O ₃	NOW	9.0	53.2	31.8	12.14	.0125	.000
	FORE	13.8	55.9	35.3	13.31	.0120	.050
	OBS	8.4	59.8	32.4			
TSP	NOW	75.8	277.4	158.7	68.36	.0030	.046
	FORE	58.3	263.3	148.8	63.00	.0031	.015
	OBS	52.3	345.7	147.9			
SO ₂	NOW	9.3	67.6	36.8	17.59	.0157	.049
	FORE	6.5	67.3	37.2	15.49	.0151	.039
	OBS	8.4	76.0	33.9			
NO ₂	NOW	28.4	91.9	57.8	24.55	.0084	.067
	FORE	17.6	85.6	51.7	18.38	.0073	.013
	OBS	21.8	95.0	50.2			
CO	NOW	8.4	57.4	28.3	15.38	.0263	.114
	FORE	6.6	47.6	26.4	12.17	.0226	.081
	OBS	7.6	46.7	22.0			

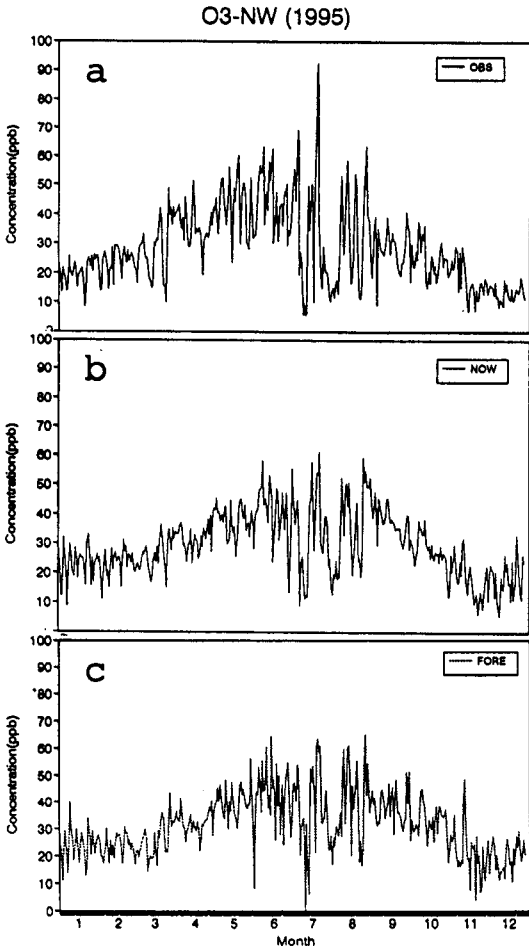


Fig. 1. O₃ concentrations (ppb) of observed (a), predicted by NOW model (b) and FORE model (c) at North-West area in 1995.

이 2회에 그쳤지만 두 모델 모두 예보값이 측정값보다 작게 나타났으며, 관측 기상자료만을 사용하는 NOW 모델보다는 예보 기상자료를 함께 사용하는 FORE 모델이 고농도 오존을 더 잘 예보하는 것으로 나타났다. 본 연구에서는 농도 수준에 따라 모델을 분리하여 개발하지 않았기 때문에 발생 횟수가 상대적으로 많은 저농도 경우에 더 많은 가중치가 주어진 모델이 만들어졌다. 따라서 상대적으로 발생 횟수가 적은 고농도 발생시 극값이 낮게 예보되는 경향이 있었다. 이 문제를 해결하기 위해서는 모델을 고농도와 저농도 경우로 추가 세분하여 개발할 필요가 있으며, 본 모델이 고농도 발생 가능성을 높게 예보할 때 고농도 모델로 다시 예보하는 것도 고농

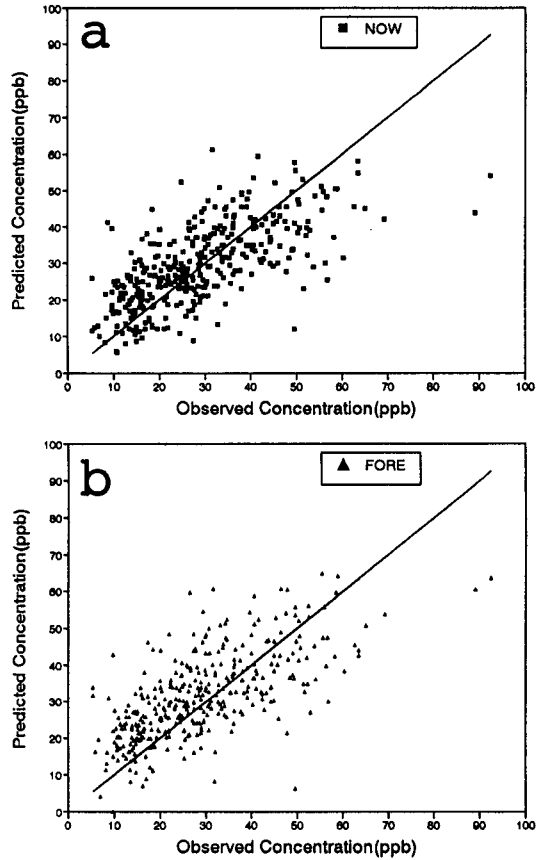


Fig. 2. Comparisons of O₃ concentrations of observed with those of predicted by NOW model (a) and FORE model (b) at North-West area in 1995.

도 예보 정확도를 높이는 한 방법이 될 것이다.

그림 3은 북서지역(NW)의 측정 및 예보된 일 최고 총먼지 농도로, FORE 모델은 농도의 상승과 하강 경향, 고농도 발생일 등은 잘 예보하였으나, NOW 모델은 특히 1, 2월과 9월에 높게 예보하는 등 계절적 편이가 나타나 겨울과 가을철 모델은 개선이 필요함을 보였다. 연평균 오염도는 측정, FORE, NOW 모델이 각각 141.9, 141.1, 175.1 $\mu\text{g}/\text{m}^3$ 으로 FORE 모델은 잘 예보하나 NOW 모델은 높은 값을 보였다. FORE 모델과 NOW 모델의 연평균 RMSE는 각각 64.07, 70.58 $\mu\text{g}/\text{m}^3$ 으로 나타났다. 오존의 경우와 같이 고농도 발생시 관측값보다 낮게 예보하는 경향이 있는데, 고농도 발생 경우에 대한 모델을 추가 개발하여 고농도 발생 가능성이 높을 때에는 고농도 예보 모델로 다시 예보

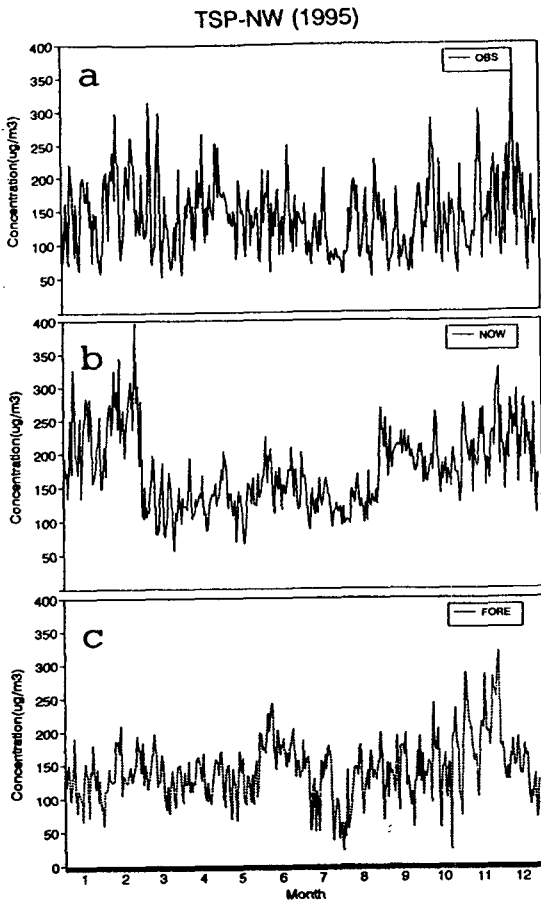


Fig. 3. Same as in Fig. 1 except for TSP concentrations ($\mu\text{g}/\text{m}^3$).

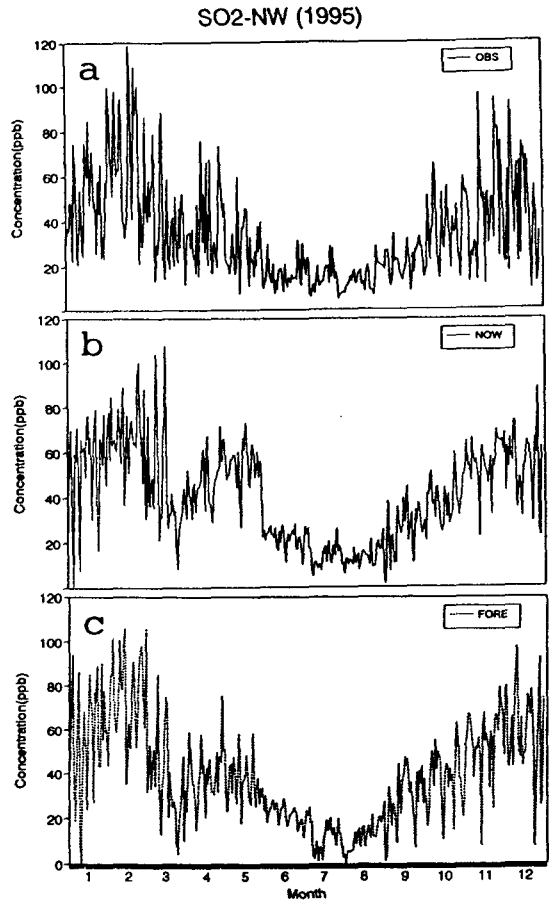


Fig. 4. Same as in Fig. 1 except for SO_2 concentrations (ppb).

한다면 예보 정확도가 높아질 것이다.

그림 4는 북서지역(NW)의 측정 및 예보된 일 최고 이산화황 농도로, NOW와 FORE 모델 모두 오염도의 상승과 하강 경향, 고농도 발생일 등은 잘 예보하였으며, NOW 모델이 측정값보다, 특히 봄철에, 높게 예보하는 계절적 편이가 나타나 봄철 예보모델의 개선이 필요함을 보였다. FORE와 NOW 모델의 연평균 RMSE가 각각 16.75, 18.16 ppb로 FORE 모델이 더 우수한 결과를 보였다. 이산화황도 오존의 경우와 같이 고농도 발생시 관측값보다 낮게 예보하는 경향이 있는데, 고농도 발생 경우에 대한 모델을 추가 개발하여 고농도 발생 가능성이 높을 때에는 고농도 예보 모델로 다시 예보한다면 예보 정확도가 높아질 것이다.

그림 5는 북서지역(NW)의 측정 및 예보된 일 최고

이산화질소 농도로, NOW와 FORE 모델 모두 오염도의 상승과 하강 경향, 고농도 발생일 등은 잘 예보하였다. 그러나 고농도 발생시 극값을 잘 예보하지 못했고, NOW 모델은 겨울철에는 높게, 봄철에는 낮게 예보하고 FORE 모델은 봄철에 높게 예보하는 등 계절적 편이를 보이므로 해당 계절의 예보 모델을 개선할 필요가 있음을 보였다. 두 모델의 연평균 RMSE가 각각 23.64, 18.64 ppb로 FORE 모델이 더 나은 결과를 보였다. 이산화질소도 오존의 경우와 같이 고농도 발생시 관측값보다 낮게 예보하는 경향이 있는데, 고농도 발생 경우에 대한 모델을 추가 개발하여 고농도 발생 가능성이 높을 때에는 고농도 예보 모델로 다시 예보한다면 예보 정확도가 높아질 것이다.

그림 6은 북서지역(NW)의 측정 및 예보된 일 최고

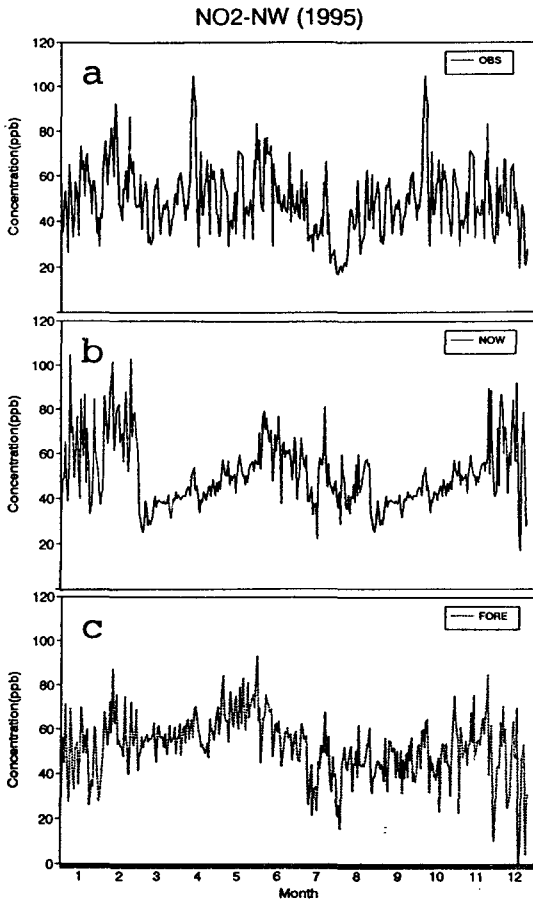


Fig. 5. Same as in Fig. 1 except for NO₂ concentrations (ppb).

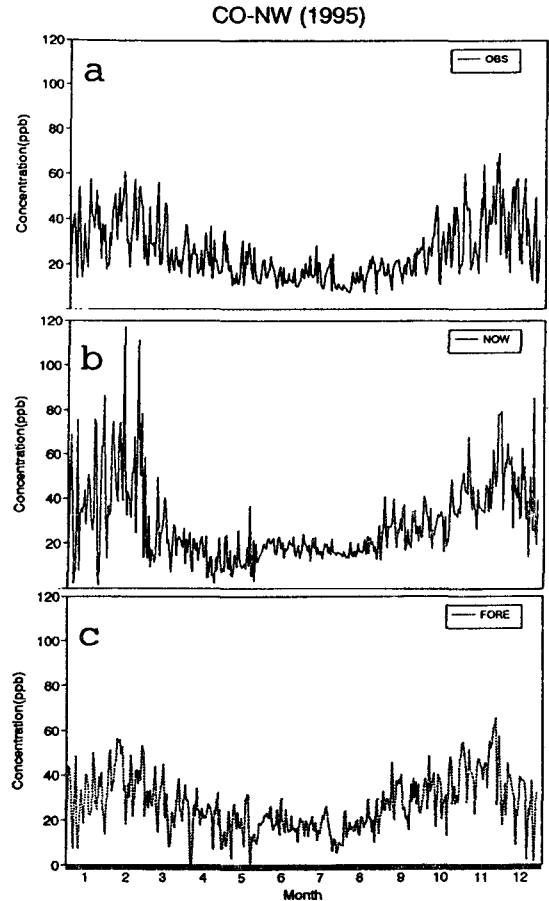


Fig. 6. Same as in Fig. 1 except for CO concentrations (ppb).

일산화탄소 농도도, 두 모델 모두 농도의 상승과 하락 경향, 고농도 발생일 등은 잘 예보하였다. 그러나 NOW 모델은 겨울철에 고농도 발생시 측정값보다 높게 예보하는 경향이 있어 겨울철 모델에 대한 개선이 필요함을 보였다. FORE, NOW 모델의 연평균 RMSE가 각각 12.28, 16.31 ppb로 FORE 모델이 더 우수한 결과를 보였다. 일산화탄소의 경우 다른 오염물질과는 달리 뚜렷한 고농도 발생 사례가 없어서 극값은 잘 예보되었다.

본 예보모델의 결과를 김용국과 이종범(1994)이 강수가 없고 전운량이 60% 미만의 맑은 날의 서울시 전지역 평균 일 최고 오존농도를 1989년과 1990년 5~9월 자료를 사용하여 개발한 통계모델을 1991년 5~9월 사례에 대해 시험예보한 결과와 비교하였다. 본 연구의

RMSE는 김용국과 이종범(1994)의 신경망모델의 13.2 ppb와 유사한 오차범위를 보였고 중회귀모델보다는 작은 오차를 보였다. 다른 오염물질에 대해서는 비교자료가 없었다.

5. 결 론

대기오염물질의 고농도 발생에 의한 피해를 줄이기 위한 방안의 하나로, 현실적으로 현업운영 가능한 대기오염 예보모델을 개발하여 5개 환경기준 오염물질의 일 최고 오염도를 하루 전에 예보할 수 있게 하였다. 서울지역을 현행 오존경보제 구역으로 세분한 4개 지역의 1994~1995년의 일 최고 오염도를 분석한 결과, 오염물질에 따라 여름철 또는 겨울철에 오염도가 높은 뚜렷

한 연변화를 보여 예보모델을 계절별로 개발하였다. 또한, 예보 기상자료의 사용 불가능한 경우에 대비하여 관측 기상자료만을 사용하는 NOW 모델과 예보 기상자료까지 함께 사용하는 FORE 모델을 개발하였다. 개발한 예보모델의 정확도는 FORE 모델이 NOW 모델보다 더 우수한 결과를 보였고, 지역별로는 큰 차이를 나타내지 않았다. 1995년 사례로 시험예보한 결과, 각 오염물질의 상승과 하강 경향, 고농도 발생일 등은 잘 예보하였으나, 예보값과 측정값의 크기에 차이를 보이는 계절이 있어 그 모델의 개선이 필요함을 보였다. 또한, 고농도 발생 경우에 대한 모델을 추가 개발하고 고농도 발생 가능성이 높게 나타나는 경우에 고농도 예보모델로 다시 예보하므로써 고농도 예보 정확도를 높일 수 있을 것이다. 오존을 제외한 오염물질은 FORE 모델이 더 우수한 결과를 보였고 지역별로 큰 차이를 보이지 않았지만 평균 오염도가 높은 계절에 큰 RMSE를 나타내는 경향을 나타냈다. 그러나 오염물질, 지역, 계절에 따라 각 모델의 특성을 파악한 후 사용 모델을 결정해야 할 것이다. 본 오존 예보모델 결과를 현재 현업운영하고 있는 미국 SoCAB 모델 결과 및 김용국과 이종범(1994)의 국내 연구 결과와 비교하였을 때, 본 연구의 예보모델을 현업예보모델로 사용 가능한 것으로 판단되었다. 다른 오염물질에 대해서는 비교 자료가 없었다.

본 예보모델의 예보 정확도를 더 높이기 위해서는 1) 종관 기상상태와 오염도 수준에 따라 예보모델을 세분하여 개발되 고농도 사례에 대한 추가 모델 개발이 필요하다. 2) 독립변수 목록에 혼합고, 상층 기상자료 등 오염도에 크게 영향을 주는 기상요소들을 추가 포함시키며 3) 각 독립변수와 종속변수의 관계가 선형 또는 비선형(예, log, exp, 다차항) 인지를 조사하는 후 중회귀방정식을 만들거나 신경회로망 방법과 같은 다른 통계기법을 사용한다. 4) 1년전의 자료에 근거한 모델보다는 어제까지의 자료를 모두 사용한 모델을 매일매일 개발하여 예보한다. 또한, 5) 새로운 자료로 계속 최신화하여 오전과 오후에 예보하는 등 예보 회수를 늘려 예보관이 최종 판단하게 한다. 6) 독립변수와 종속변수를 정규화하여 고농도 발생 가능성의 확률을 예보하므로써 오보 위험성을 줄이는 것도 한 방법이다. 장기적으로는, 7) 통계 모델과 확산모델을 동시에 운영하는 체계로 발전시켜야 하며, 확산모델의 개발 및 현업 운영을 위해서 기초 자료 조사와 모델 개발에 투자가 이루어져야 한다.

본 예보모델의 현업운영에 필요한 오염도 자료는 환경부와 서울시에서 측정하여 환경부 전산실에서 온라인 수집하고 있으며, 기상자료는 기상청에서 관측, 수집, 배

분하고 있다. 예보결과에 따른 대책은 환경부와 서울시가 실시하고, 언론기관에서 시민들에게 알려야 한다. 따라서, 예보모델의 현업운영은 기상청, 환경부, 서울시 및 언론기관 사이에 공동으로 추진되어야 한다. 대기오염 예보가 실제 실행되기 위해서는 우선 유관기관 간에 현업운영의 필요성을 공감하고 기초자료와 예보결과의 공동 이용, 기관간 역할분담 등에 대한 행정협약이 이루어져야 한다. 또한, 현업운영 기관은 자료의 수집과 배포망을 구축하고 전산기와 모델 운영 담당자도 구비해야 한다.

감사의 글

연구에 사용된 대기오염도 자료를 제공한 환경부와 기상자료를 제공한 기상청의 담당자들과 자료 정리에 도움을 준 한국환경기술개발원의 김현주 연구원에게 감사드립니다. 이 연구는 환경부의 용역과제 "대기오염 예보기법 개발연구"의 일부로 이루어졌다.

참고 문헌

- 김민영, 강희곤, 이완종, 이상철(1989) 오존농도의 동태 및 영향인자에 관한 연구, 한국대기보전학회지, 5(2), 55-71.
- 김상진, 이산훈, 정 용(1988) AIRMA model에 의한 서울시 일부지역 SO₂ 오염도의 월변화에 대한 시계열 분석, 한국대기보전학회지, 4(2), 72-81.
- 김신도, 김종호, 이정주(1992) 겨울철 서울 도심의 대기 중 부유분진의 특성, 한국대기보전학회지, 8(2), 105-111.
- 김용국, 이종범(1994) 하계의 일최고 오존농도 예측을 위한 신경망 모델의 개발, 한국대기보전학회지, 10(4), 224-232.
- 김혜경(1991a) 대기오염농도에 관한 확률모델, 한국대기보전학회지, 7(2), 127-136.
- 김혜경(1991b) 대기오염농도에 관한 동적확률모델, 한국대기보전학회지, 7(3), 156-168.
- 나진균(1989) 도심지역에서의 풍속에 따른 SO₂ 농도변화, 한국대기보전학회지, 5(2), 97-105.
- 박일수, 이덕길, 강인구(1991) 서울지역 겨울철 SO₂ 농도를 지배하는 기상인자, 한국대기보전학회지, 7(2), 96-194.
- 송동웅, 김영만(1991) 서울시 지역별 SO₂ 오염도 분석에 관한 연구, 한국대기보전학회지, 7(1), 23-30.

- 신찬기, 한진석, 김윤기(1992) 대기오염농도와 기상인자의 관련성 연구 - 서울 광화문 지점을 중심으로, 한국대기보전학회지, 8(4), 212-220.
- 이기원, 권숙표, 정 용(1993) 서울시 대기 중 오존 오염도의 연도별 변화와 그 영향인자 분석 - 광화문 지역을 중심으로, 한국대기보전학회지, 9(1), 107-115.
- 이선기(1991) 서울지역의 광화학 오염농도 추정을 위한 중회귀 모델 개발, 이학석사 학위논문, 강원대학교 환경학과, 54pp.
- 전병일, 김유근, 이화운(1994) 부산지역의 SO₂ 농도변화 특성에 관한 고찰, 한국대기보전학회지, 10(4), 245-251.
- 전병일, 김유근, 이화운(1995) 부산지역의 오존농도 특성과 시정인자에 관한 연구, 한국대기보전학회지, 11(1), 45-56.
- 정 용, 장재연, 권숙표(1986) 서울시 대기 중 오존의 오염도와 그 영향인자 분석, 한국대기보전학회지, 2(1), 73-79.
- 정용승, 정재섭(1991) 서울 수도권 지역의 광화학 오존에 관한 연구, 한국대기보전학회지, 7(3), 169-179.
- 허정숙, 김동술(1993) 다변량 통계분석을 이용한 서울시 고농도 오존의 예측에 관한 연구, 한국대기보전학회지, 9(3), 207-215.
- 환경부(1996a) 대기오염 예보기법 개발 연구, 한국환경기술개발원 연구보고서, 108pp.
- 환경부(1996b) 1996년 환경백서, 환경부, 597pp.
- 환경부(1996c) 대기분과위원회 회의자료, 환경부, 43 pp.
- Cassmassi, J.C.(1987) Development of an objective ozone forecast model for the South Coast Air Basin, Proceeding of the 80th Annual Meeting of APCA, June 21-26, 1987, New York, U. S. A.
- Cassmassi, J.C.(1994) Weather and smog forecasting in the SoCAB.
- South Coast Air Quality Management District (1994). *Brochure of South Coast Air Quality Management District.*