

PS9(SM32) PMF를 이용한 PM-10의 오염원 추정 PM-10 Source Estimation Using Positive Matrix Factorization

황인조, 김동술

경희대학교 환경학과 및 산학협력기술연구원

1. 서론

대기 연구자들은 대기오염의 일반적인 현황과 대기오염 유발의 근본 원인 파악, 저감 대책 등에 대한 연구를 활발히 수행하고 있다. 하지만 이러한 연구들 중에서 대기오염의 근본 원인을 파악하기 위한 오염원 (source) 추정 연구는 국내외적으로 매우 미진하다. 대기질의 평가와 예측은 분산모델과 수용방법론을 이용하는데, 분산모델에 내재되어 있는 한계성과 제약점 때문에 수용체에서 오염물질의 특성을 분석한 후, 오염원의 기여도를 평가하는 수용방법론이 지속적으로 개발되고 있다.

수용방법론은 크게 이화학적 분석분야와 여기서 도출된 자료에 입각하여 결과를 해석하는 응용통계학적 분석분야로 나눌 수 있다. 이화학적 분석은 전량분석 (bulk analysis)과 개개의 입자에 대하여 분석하는 입자별 분석 (particle by particle analysis)으로 나눌 수 있다. 이러한 이화학적 분석을 통하여 가스상·입자상 오염물질의 물리·화학적 특성을 파악한 후, 응용통계학적 분석을 통하여 오염원의 기여도를 정성·정량적으로 추정할 수 있다. 수용방법론 초기에는 농축계수법이나 시계열분석법 등이 많이 사용되었으나 최근에는 화학질량수지법과 인자분석법 등의 다변량 통계분석이 널리 사용되고 있다. 최근 가장 보편적으로 사용되는 다변량 통계분석은 인자분석으로서, 복잡하고 방대한 환경자료에 대한 해석에 주로 사용되며, 오염원의 기여도를 정량적으로 추정할 수 없다는 단점에도 불구하고 주로 오염원 분류표의 부재시 오염원의 추정에 이용되는 통계 기법이다 (황인조와 김동술, 1998; Poissant *et al.*, 1996; Huang *et al.*, 1999). 그러나 인자분석시 발생하는 음수의 인자부하량 발생, 불충분한 정보 제공 등의 문제점으로 인하여 Paatero와 Tapper (1994)에 의해서 인자분석보다 수학적으로 개선된 알고리즘인 PMF (positive matrix factorization) 모델이 개발되었다.

본 연구의 목적은 high-volume air sampler로 채집된 PM-10 중 무기원소와 이온성분의 농도분포 현황을 파악하고 변수들 간의 공분산 혹은 상관중심의 인자분석보다 향상된 PMF 모델을 이용하여 PM-10 오염원을 정량적으로 추정하고자 하였다.

2. 연구내용 및 방법

시료의 채집기간은 1997년 3월 1일부터 1998년 2월 28일까지 채집하여 총 254개의 시료를 채집하였다. 시료의 채집장소는 경기도 용인시에 위치한 경희대학교 자연과학대학 옥상 (5층)에서 시료를 채집하였다. PM-10의 채집은 high-volume air sampler (U.S.A., General Metal Works, Model IP10)를 사용하여 평균 1.13 m³/min의 유량으로 평균 24시간동안 매일 채집하였다. 여지는 수정점유여지 (quartz microfibre filter, QM-A, 8"×10", Whatman)를 사용하였다. 시료 채집 후 전자 테스게이터 (Sanplatec Corp., Model Oyin 09678BN) 내에 보관하여 항량시킨 후 0.01 mg의 감도를 갖는 전자저울 (A&D Co., Model HM-202)로 칭량하여 분진농도를 계산하였다. 채집된 PM-10 입자의 무기원소 분석을 위한 전처리법은 미국 EPA가 1992년 10월 13일에 고시한 CWA (Clean Water Act)의 microwave 전처리법을 준용하여 Questron (Questron사, Model Q-15 MicroPrep)을 이용한 질산 전처리방법을 사용하였으며, 수용성 이온성분을 분석하기 위한 전처리법은 여지를 초순수에 침적시켜 초음파 추출기로 추출하는 방법을 이용하였다. 전처리가 끝난 시료는 원자흡광광도계 (AAS, Atomic Absorption Spectrophotometer, Hitachi, Model Z-8100)를 사용하여 Zn, Cu, Fe, Pb, Al, Mn 등 6개의 무기원소를 분석하였으며, 수용성 이온성분은 이온 크로마토그래피 (Dionex사, Model DX-400)를 사용하여 Na⁺, NH₄⁺, K⁺, Mg²⁺, Ca²⁺ 등 5개 항목을, 음이온은 SO₄²⁻, NO₃⁻, Cl⁻ 등 3개 항목을 분석하였다.

3. PMF (positive matrix factorization)

PMF는 측정 자료의 표준편차를 고려하여 인자부하량이 항상 양의 값으로 계산되게 하는 인자분석의 한 종류이다 (Paterson *et al.*, 1999). 인자분석과 PMF 분석의 주된 차이점으로 PMF의 경우는 인자부하량이 항상 양의 값으로 계산되며, 상관행렬의 정보에 의존하는 것이 아니라 개개 자료의 최소 자승값이 최소가 되게 하는 (least-squares minimization) 알고리즘 즉, 각 자료의 오차 추정분석 (error estimate)에 기초한다.

$$X = G F + E \quad (1)$$

여기서 행렬 X는 차원이 $n \times m$ 인 데이터 행렬이며, G 행렬은 $n \times p$, F 행렬은 $p \times m$ 차원으로 표시되며 여기서 p는 추출된 인자의 수를 나타낸다. F 행렬의 행 (row)은 특정한 하나의 오염원에 대한 오염원 분류표 (source profile)를 나타내며 이에 대응하는 G 행렬의 열 (column)은 특정한 하나의 오염원에 대하여 시료들에 존재하는 배출량 즉, 각각의 측정시간에서 오염원의 강도 (strength)를 의미한다. 행렬 G와 F의 값들은 항상 양의 값을 가지게 되며, 또한 행렬 E는 잔차 행렬을 의미한다. PMF 분석의 목적은 Q값을 최소화하는 것인데, 여기서 σ_{ij} 는 X_{ij} 의 표준편차를 나타낸다.

$$Q = \sum_{i=1}^n \sum_{j=1}^m (E_{ij}^2 / \sigma_{ij}^2)$$

PMF 분석 결과에서 중요한 변수는 EV값 (explained variation value)으로서, 이 값은 각 오염원에서 각각의 화학원소에 대한 기여도를 측정할 값을 의미한다.

$$EV_{kj} = \frac{\sum_{i=1}^n |G_{ik} F_{kj}| / S_{ij}}{\sum_{i=1}^n (\sum_{k=1}^p |G_{ik} F_{kj}| + |E_{ij}|) / S_{ij}}, \quad k=1, \dots, p$$

4. 결과 및 고찰

1997년 3월 1일부터 1998년 2월 28일까지 PM-10 시료를 채집하여 분진농도, 무기원소, 그리고 이온성분의 농도를 구하였다. 시료 채집기간 중 PM-10의 24시간 기준치를 초과한 횟수는 총 8회로, 연구지역의 대기질은 비교적 양호한 것으로 조사되었으며 초과 건수는 봄철인 3, 4, 5월과 겨울철인 12, 1, 2월에 집중되어 있는 경향을 나타내었다. 분진농도의 월평균은 3월에 $86.3 \mu\text{g}/\text{m}^3$ 으로 최고농도를 나타내었으며 8월에 $28.5 \mu\text{g}/\text{m}^3$ 으로 최저농도를 나타내었다 (그림 1). 무기원소의 계절별 평균농도에서 Zn, Cu, Al의 경우는 봄철이 가장 높은 농도를 보였으며, Fe와 Mn의 경우는 겨울철이 가장 높은 농도를 나타내는 것으로 조사되었다. Pb의 경우는 여름철을 제외한 봄, 가을, 겨울철에 거의 유사한 농도 수준을 나타내는 것으로 조사되었다. 양이온의 경우, 겨울철과 봄철에 고농도를 나타내고 여름철에 저농도를 나타내는 전형적인 농도분포를 나타내었지만, Ca^{2+} 의 경우는 여름, 가을, 겨울철 모두 유사한 수준의 농도분포를 보이고 봄철의 경우 다른 계절에 비해 약 2배정도 높은 농도를 나타내는 경향을 보였다. 음이온의 경우는 각 항목마다 특이한 경향을 나타내었는데, SO_4^{2-} 는 여름철에 가장 높은 농도를 나타내었으며, NO_3^- 는 봄철이 가장 높은 농도로 조사되었으며 나머지 계절은 거의 유사한 농도수준으로 조사되었다. Cl^- 는 봄철과 겨울철에 높은 농도를 나타내는 것으로 조사되었다.

참고 문헌

- 황인조, 김동술 (1998) Submicron 부유분진의 화학적 조성 및 분포에 관한 연구, 한국대기보전학회지, 14(1), 11-23.
- Huang, S., K.A. Rahn, and R. Arimoto (1999) Testing and optimizing two factor-analysis techniques on aerosol at Narragansett Rhode Island, Atmospheric Environment, 33(14), 2169-2185.
- Paatero, P. and U. Tapper (1994) Positive matrix factorization: A non-negative factor model with

optimal utilization of error estimates of data values, *Environmetrics*, 5, 111-126.

Paterson, K.G., J.L. Sagady, D.L. Hooper, S.T. Bertman, M.A. Carroll, and P.B. Shepson (1999) Analysis of air quality data using positive matrix factorization, *Environ. Sci. & Technol.*, 33(4), 635-641.

Poissant, L., J.W. Bottenheim, P. Roussel, N.W. Reid, and H. Niki (1996) Multivariate analysis of a 1992 SONTOS data subset, *Atmospheric Environment*, 30(12), 2133-2144.

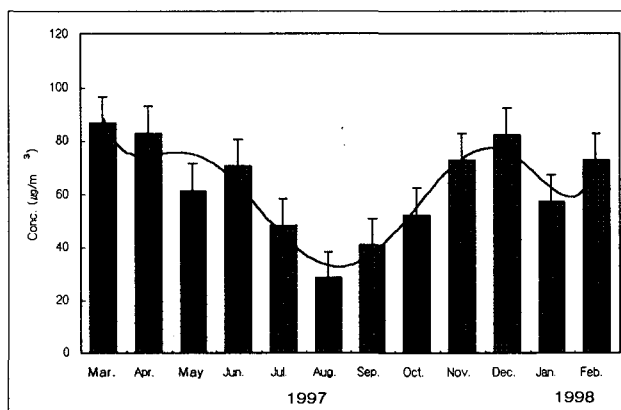


Fig. 1. Monthly average concentration of PM-10 particle in the study area during Mar. 1997 ~ Feb. 1998.

Table 1. Average seasonal concentration of inorganic and ionic species in the PM-10 during Mar. 1997 ~ Feb. 1998.

(unit : ng/m³)

	Concentration (ng/m ³)				Concentration (ng/m ³)				
	Spring	Summer	Fall	Winter	Spring	Summer	Fall	Winter	
Zn	552.9	364.9	269.4	284.2	NH ₄ ⁺	6939.1	6286.2	5228.6	7442.1
Cu	55.1	43.6	39.8	42.9	K ⁺	659.1	372.2	691.0	795.2
Fe	1061.7	812.9	1263.2	1438.2	Mg ²⁺	268.2	60.6	139.6	183.8
Pb	198.8	152.3	208.6	185.1	Ca ²⁺	2376.6	1038.8	1341.4	1096.2
Al	1710.4	1184.7	846.0	1154.6	Cl ⁻	3443.9	683.1	3232.4	4530.4
Mn	31.2	22.3	37.7	66.3	NO ₃ ⁻	10344.3	7067.5	7277.3	7212.3
Na ⁺	816.2	128.6	436.6	780.0	SO ₄ ²⁻	11714.1	13345.3	7098.5	8807.7