

## 분산주성분 분석을 이용한 실내환경 중 PM-10 오염의 패턴분류

### Pattern Classification of PM-10 in the Indoor Environment Using Disjoint Principal Component Analysis

남 보 현\* · 황 인 조 · 김 동 술  
경희대학교 환경·응용화학부 대기오염연구실 및 환경연구센터  
(2001년 11월 27일 접수, 2002년 2월 21일 채택)

Bo-Hyun Nam\*, In-Jo Hwang and Dong-Sool Kim  
*Department of Environmental Science and Engineering School of Environment and  
Applied Chemistry and Institute of Environmental Studies, Kyung Hee University*

(Received 27 November 2001; accepted 21 February 2002)

#### Abstract

The purpose of the study was to survey the distribution patterns of inorganic elements of PM-10 in the various indoor environments and analyze the pollution patterns of aerosol in various places of indoor environment using a pattern recognition method based on cluster analysis and disjoint principal component analysis. A total of 40 samples in the indoor had been collected using mini-vol portable samplers. These samples were analyzed for their 19 bulk inorganic compounds such as B, Na, Mg, Al, K, Ca, Ti, V, Cr, Fe, Ni, Cu, Zn, As, Se, Cd, Ba, Ce, and Pb by using an ICP-MS. By applying a disjoint principal component analysis, four patterns of the indoor air pollutions were distinguished. The first pattern was identified as a group with high concentrations of PM-10, Na, Mg, and Ca. The second pattern was identified as a group with high concentrations B, Mg, Al, Ca, Fe, Cu, and Ba. The third pattern was a group of sites with high concentrations of K, Zn, Cd. The fourth pattern was a group with low concentrations PM-10 and all inorganic elements. This methodology was found to be helpful enough to set the criteria standard of indoor air quality, corresponding pollutants, and classification of indoor environment categories when making an indoor air quality law.

**Key words** : indoor air quality, cluster analysis, disjoint principal component analysis

#### 1. 서 론

현대인의 대부분이 하루 중 80~95% 이상의 시간을 다양한 실내공간(가정, 사무실, 학교, 지하시설

물, 상가, 음식점, 지하철 등)에서 생활하고 있어 (Wiley *et al.*, 1991; Szalazi, 1972) 실내공기오염이 인체에 미치는 영향은 크다고 할 수 있다(Chapin, 1974). 이러한 실내공기의 상태는 일차적으로 외부공기의 영향을 받게 되고 이차적으로는 담배연기, 시멘트, 건축자재, 페인트 및 벽면의 입자상 물질, 주방 및 생활용품 등과 같은 실내오염원으로부터 영

\* Corresponding author  
Tel : (062) 952-7833, E-mail : nambol@hanmail.net

향을 받아 오염의 정도가 심해진다.

최근 실내공기오염은 한정된 공간에서 오염물질의 지속적 순환으로 인한 공기질의 악화, 신규 건축 자재 등에서 방출되는 미확인 오염물질의 증가, 건물의 밀폐화로 인한 건강 악화 등이 거론되며 중요성이 커지고 있다. 이러한 이유로 외국의 경우는 1970년대부터 실내공기오염에 대한 연구가 활발히 진행되고 있는 반면에, 우리 나라의 경우는 실내공기오염에 대한 인식부족으로 관련 연구가 매우 열악한 실정이다(대기환경연구회, 2000). 특히, 실내환경에서의 입자상 오염물질에 대한 연구는 총부유분진(total suspended particulate; TSP)에 대한 연구만이 일부 수행되었을 뿐, 호흡성 분진(respirable suspended particulates) 및 PM-10에 대한 연구는 미흡한 실정이다.

실내공기오염에 대한 합리적인 대책과 관리방안 등을 마련하기 위해서는 실내공기 오염원에 대한 연구가 선행되어야 하며, 이러한 연구는 수용방법론(receptor methods)을 이용하여 수행할 수 있다. 수용방법론 중에서 오염원에 대한 정보가 부족할 때 인자분석, 다중회귀분석, 군집분석 등의 다변량 분석법이 보편적으로 많이 사용되고 있다(황인조 등, 2001; 유상준과 김동술, 1997; Hopke, 1985). 본 연구에서는 군집분석과 인자분석의 한 부류인 분산주성분 분석법을 이용하였다. 이러한 응용통계 방법론들은 오염원에 대한 정보가 부족할 때 자료 집단에서 유사한 성질을 갖는 특정 군집을 분류하고, 분류 작업이 올바르게 수행되었는지의 여부를 정량적으로 규명하는 방법으로서, 국내에서는 김동술과 김형석(1990)이 실내환경에서 여러 오염물질의 농도를 새로운 각도에서 분석하고 오염의 패턴을 분석하기 위하여 이용한 바 있다.

본 연구는 다양한 실내환경에서의 PM-10 오염현황과 오염의 형태(pattern)를 파악하고자 하였으며, 이를 위해 다양한 실내공간에서 채취된 PM-10의 무기원소성분을 ICP-MS를 이용하여 분석한 후, 첫째, 실내공간에서의 PM-10 농도경향을 파악하고자 하였으며 둘째, 군집분석을 이용하여 분석된 PM-10 무기원소 자료를 유사한 특성을 지닌 군집으로 분류하였다. 그리고 마지막으로 분산주성분 분석(disjoint principal component analysis)을 이용하여 분류된 군집을 확률적으로 검증한 뒤, 최종적으로

산출된 순수군집(homogeneous cluster)의 특징과 실내환경 중 PM-10 오염의 패턴을 파악하고자 하였다. 본 연구방법론은 추후 실내 공기질법이 마련될 때 오염물질의 선정, 기준의 설정, 규제지역의 분류 시 도움이 될 것이다.

## 2. 연구 및 방법

### 2.1 실험방법

PM-10의 채취는 다양한 실내공간(가정, 사무실, 학교, 지하시설물, 상가, 음식점, 지하철 등)의 총 40개 측정지점에서 수행되었으며 이를 표 1에 나타내었다. 시료의 채취기간은 2000년 4월~10월까지이며, 시료는 mini-vol portable sampler (Model 4.1, Airmetrics Co., USA)를 사용하여 채취하였다. 채취유량은 5 L/min로 고정하여 12 시간 이상 채취하는 것을 기본으로 하였으며, 분진채취기가 과부하 등으로 정지하는 경우를 대비하여 채취기에 부착된 시간기록계의 수치를 기록하였다. 시료 채취에 사용된 여지는 미국 Corning Costas사의 직경 47 mm, pore size 0.2 μm의 Membra-Fil membrane filter를 사용하였다. 여지는 시료 채취 전·후에 전자 건조기에 보관하였으며, 시료채취 전·후의 무게차를 채취유량으로 나누어 분진농도를 계산하였다.

여지 위에 채취된 입자의 무기원소분석을 위해 우선 microwave 전처리법을 사용하였으며 구체적인 절차는 문헌을 참고할 수 있다(황인조와 김동술, 1998). 다양한 실내공간에서 채취된 PM-10 시료는 ICP-MS(ICP-MS 4500/Hewlett Packard)를 이용하여 B, Na, Mg, Al, K, Ca, Ti, V, Cr, Fe, Ni, Cu, Zn, As, Se, Cd, Ba, Ce, Pb 등 총 19항목에 대하여 분석을 수행하였으며 표 2에 ICP-MS에 대한 실험조건을 나타내었다.

### 2.2 군집분석과 주성분분석

군집분석(cluster analysis)은 자연과학 분야에서 널리 사용되고 있는 응용통계 분석법으로서, 자료집단 내에서 유사한 성질을 갖는 특정 그룹을 유도·분류하는데 이용하고 있다. 군집분석은 군집의 수, 내용, 구조 등이 완전히 알려지지 않은 상태에서 개체 사이의 거리(distance) 또는 유사도(similarity)에

**Table 1. Sampling sites in this study.**

ID	Type of indoor	Sampling site	Ground / Underground	Smoking / Non smoking
1	Entertainment Place	Billiard room A	G*	Smoking
2		Billiard room B	U**	Smoking
3		Billiard room C	U	Smoking
4		Cartoon room A	U	Smoking
5		Cartoon room B	U	Smoking
6		Cartoon room C	U	Smoking
7		Singing establishment A	U	Smoking
8		Singing establishment B	G	Smoking
9		Game room	G	Smoking
10	Restaurant	Eating house A	G	Smoking
11		Eating house B	G	Smoking
12		Eating house C	G	Smoking
13		Beer hall A	U	Smoking
14		Beer hall B	U	Smoking
15		Beer hall C	G	Smoking
16		Beer hall D	G	Smoking
17		Coffee shop A	G	Smoking
18		Coffee shop B	G	Smoking
19	Store	Complex store A	G	Smoking
20		Complex store B	G	Smoking
21		Complex store C	G	Smoking
22	School	Library A	G	Smoking
23		Library B	G	Smoking
24		Library C	U	Smoking
25		University A	G	Smoking
26		University B	G	Smoking
27		University C	G	Smoking
28		University D	G	Smoking
29		University E	G	Smoking
30		University F	G	Smoking
31	Office	Office A	G	Smoking
32		Office B	G	Smoking
33	Residential area	House A	G	Smoking
34		House B	G	Smoking
35		Apartment	G	Smoking
36	Underground structures	Subway A	U	Non smoking
37		Subway B	U	Non smoking
38		Subway C	U	Non smoking
39		Subway D	U	Non smoking
40		Subway E	U	Non smoking

\* : Ground, \*\* : Underground

**Table 2. Analytical conditions of ICP-MS.**

ICP-MS	HP 4500
Reflected power	1,300 W
Carrier gas (Ar)	1.15 L/min
Sampling depth	8.0 mm
Integration time	0.1 sec
Sampling period	0.3~0.9 sec
Ar Plasma temp.	8,000~10,000°C
	Extraction lens
Ion lense	Einzel lens
	Omega lens
Neblizer	Babington type

근거하여 군집을 분류하고, 분류된 군집의 특성을 파악하며, 군집들간의 관계를 분석하는 것을 주요 목적으로 한다. 또한 군집간의 분산(variance)은 최대한화하며 군집 내부의 분산은 최소화시키는데 그 목적이 있다(김기영과 전명식, 1997; Gomez and Martin, 1987). 이 분석법의 기본 원리는 두 object 사이의 거리를 기준으로 공간에서 비유사도를 측정 함으로서 동일한 패턴을 갖는 object를 파악하여 거리 또는 비유사도가 작은 두 object를 같은 군집으

로 소속화 시키는 방법이다(Hopke, 1991). 군집분석법은 연구과정에서 자료의 형태 윤곽은 쉽게 파악할 수 있으나, 유사도 측정방법의 선정, 군집수를 결정하는 유사도 준위의 결정 등 분석자의 주관적인 판단이 개입될 수 있으므로 연구의 최종 결과를 얻고자 할 때 세심한 주의가 필요하다.

군집분석을 이용하여 유사한 object들이 각각의 군집을 형성하고 미완성이지만 일단 object들의 분류작업이 끝났을 때, 각각의 군집으로 분류된 object들이 과연 그 군집(또는 패턴)에 소속되어 있는지의 여부를 정량적으로 규명하고, 분류된 군집을 확률적으로 검증한 뒤, 최종적으로 순수 군집을 도출할 필요가 있다. 본 연구에서는 분산주성분 분석을 응용한 형태인식법(pattern recognition)을 적용하였고, 이를 위해 SIMCA(Soft Independent Modeling of Class Analogy) package(SIMCA-3B manual, 1984)를 이용하여 연구과정을 수행하였다. 분산주성분 분석법은 이미 알고 있거나 군집분석 등을 통해 생성된 군집 각각에 대해 독립적으로 주성분분석을 수행하는 통계적 방법론이다. 한편, SIMCA는 Wold(1976)에 의해 개발된 형태인식기법으로, 분산주성분 분석 모델을 기초로 한 통계적 모델이다. 분산주성분 분석을 통해 각 군집은 선형구조로 모델화되고, 한개의 object가 고정된 확률 값에서 이미 모델화 된 특정 군집속에 소속할 수 있는지를 임계거리(critical distance)를 이용하여 결정할 수 있다.

SIMCA를 수행할 때 입력자료는 군집분석과 같은 방법으로 이미 알고 있는 군집인 training set와 군집으로 분류되지 않는 test set으로 구성된다. 또한 SIMCA 수행 후 어떠한 군집에도 포함되지 않는 object를 특이점(outlier)이라 부른다.

SIMCA의 원리는 다중 Taylor 급수에서부터 출발하며, 자료  $y_{ik}$ 는 식(1)과 같이 주성분 수  $k$ 개를 가진 주성분모델에 의해 근사할 수 있다.

$$y_{ij}^{(q)} = y_i^{(q)} \sum_{k=1}^q b_{ik}^{(q)} t_{kj}^{(q)} e_{ij}^{(q)} \quad (1)$$

여기서,  $y_{ik}$ 는  $q$ 번째 군집에서  $k$ 개의 주성분을 가진  $i$ 번째 변수와  $j$ 번째 object의 측정치를 의미한다. 즉, object는 각 시료(시료채취 장소에서 채취한 시료)

를 의미한다. 또한 우변의  $y$ 는 변수  $i$ 의 평균치,  $b$ 는 변수 관련항,  $t$ 는 object의 값을 나타내며, 잔차  $e_{ij}$ 는  $y_{ik}$ 의 불규칙한 부분(random part)을 설명하며, 측정 오차(measurement error)와 모델링 오차를 포함하고 있다. 모델링 오차보다 측정 오차가 크다면 모델은 자료를 잘 해석한다는 것을 의미하는 것이다(Wold, 1976). 잔차  $e_{ij}$ 의 제곱 합은 모델화 된 군집과, 그 군집에 속해 있는 한 object 사이의 거리를 나타낸다. 또한, 잔차로부터 각 모델에 대한 잔여표준편차(residual standard deviation)는 식(2)에 의해 계산되며, 이것은 training set에서 특이점을 감지하는데 사용된다(유상준과 김동술, 1997).

$$S_0^{(q)} = \left[ \sum_i \sum_j e_{ij}^{2(q)} / (n_q - a_q - 1) (m - a_q) \right]^{1/2} \quad (2)$$

여기서  $n_q$ ,  $a_q$ ,  $m$ 은 각각 군집  $q$ 내의 object 수, 주인자의 수, 그리고 변수의 수를 의미한다.

식(1)에서 주성분모델은 주성분의 수가 증가할수록 잔차  $e_{ij}$ 는 작아질 것이고, 따라서 모델은 자료에 더욱 근접하여 fitting이 가능하게 되는 반면에 결과에 대한 모델의 유효성은 제약을 받게 될 것이다(Wold et al., 1984). 따라서 주성분의 수를 결정할 때는 이러한 두 가지 측면을 모두 고려해야 하며, SIMCA에서는 횡유효도 검사(cross validation)를 통해 주어진 군집에 대해 통계적으로 유효한 주성분의 수를 결정한다. 자료행렬  $y$ 에서  $A$ 개의 주성분에 대한  $y_i$ ,  $b_{ik}$ ,  $t_{kj}$ 가 계산되었다면, 주성분의 수를 추가할 경우 좀더 양호한 fitting이 가능한지의 여부를 결정하기 위하여 횡유효도 검사를 수행하고(Wold and Sjöström, 1977; Wold, 1976) 주성분이 통계적으로 유효한지의 여부를 결정한다.

### 3. 결과 및 고찰

#### 3.1 연구대상 실내공간의 특성 및 PM-10 농도분포 현황

다양한 실내공간에서 측정된 PM-10의 농도값과 그 평균값을 표 3에 나타내었다. 또한 13곳 측정장소를 장소별로 분류하여 각 장소별 PM-10 평균 농도값을 24시간 기준치와 비교하여 그림 1에 나타내었다. 실내공기기준에 대한 국내기준은 최근까지 마

런되지 않았지만, 본 연구에서는 지하역사와 지하도상가 등에 대해 제정된 「지하생활공간 공기질 관리법」을 잠정적으로 활용하였다. 이 법에 의하면, PM-10의 실내기준은 1999년 12월 31일까지는 24시간 평균 250 µg/m<sup>3</sup> 이하를 적용하고 2001년 12월 31일까지는 24시간 평균 200 µg/m<sup>3</sup> 이하를 적용하며 그 이후로는 24시간 평균 150 µg/m<sup>3</sup> 이하를 적용한다고 명시되어 있는데 (환경부 1998), 본 연구에서는 24시간 평균 150 µg/m<sup>3</sup> 이하를 적용하여 평균값과 비교하였다.

유흥업소 중 당구장의 경우 PM-10 농도가 167.3 µg/m<sup>3</sup>으로서 기준치를 초과하는 높은 농도값을 나타내었다. 당구장의 경우는 창문이 거의 존재하지 않고 환기에 어려움이 있으며, 실내에서의 흡연, 이용자의 잦은 이동 등에 의해 실내 자체내에서 발생된 PM-10과 바닥에 침착된 PM-10이 다시 부유되어 높은 농도값을 나타낸 것으로 사료된다. 만화방의 경우는 평균 농도값이 120.9 µg/m<sup>3</sup>으로서 기준치를 초과하지는 않았지만 비교적 높은 농도값을 보이고 있다. 이는 당구장에 비해 이용자의 이동율이 작지만 흡연과 불충분한 환기 등에 의해 비교적 높은 농도값을 보이고 있었다. 노래방의 경우는 86.5 µg/m<sup>3</sup>의 평균 농도값을 나타내는 것으로 조사되었다. 이 장소의 경우 창문 없었으며 기계환기 방식에 의해 환기가 잘 이루어지고 있어 비교적 낮은 농도값을 보였다. 게임룸 실내공간에서의 PM-10 평균농도는 117.0 µg/m<sup>3</sup>으로서 만화방의 수준과 유사한 값을 나타내었다.

음식점 중 식당의 PM-10 농도는 148.3 µg/m<sup>3</sup>으로 조사되었다. 이 장소는 환기설비가 작동중임에도 불구하고, 실내에서 숯불을 연료로 고기를 구움으로서 발생하는 입자상 오염물질과 흡연의 다량배출로 높은 농도값을 보이고 있었다. Dyremark *et al.* (1995)에 의하면 고기를 숯불에 구울 때 다량의 입자상 PAH 등이 발생한다고 보고하고 있다. 술집 (beer hall)의 경우는 기준치의 거의 2배를 상회하는 270.0 µg/m<sup>3</sup>의 매우 높은 농도로서, 본 연구 수행 중 최대 PM-10 농도값을 나타내었다. 이는 측정 장소의 경우, 지하에 위치하며 환기가 제대로 이루어지지 않았고, 외부 유입 PM-10의 배기가 원활하지 않았던 것으로 사료되었다. 또한 많은 양의 흡연과 취사도구 등에서 발생하는 입자상 물질도 크게 영

향을 미친 것으로 사료되었다. Coffee shop의 PM-10 평균 농도는 110.2 µg/m<sup>3</sup>으로서, 실내에 특이한 오염원은 존재하지 않았지만 근처의 건물 신축공사의 영향을 받은 것으로 사료되었다. 한편, 김동술과 김형석 (1990)에 의하면, 실내공간에서 난방 및 취사를 위해 화석연료 액체연료를 열원으로 사용하는 곳은 예외 없이 높은 농도를 보였다고 보고한 바 있다.

상가는 대학교내에 위치한 복합상가를 대상으로 측정된 결과, 기준치를 초과하는 153.8 µg/m<sup>3</sup>으로 조사되었다. 학교 내의 복합상가는 식당가와 잡화점들이 모여 있어 학생들의 유출·입이 빈번하며, 이로 인해 신발에 묻혀 들어오는 토양성분들과 흡연, 그리고 식당가 취사도구 연소 등에 의해 높은 농도값을 나타낸 것으로 사료된다.

학교의 실내공간에 대한 PM-10 측정은 도서관과 여러 연구실, 기숙사 등에서 이루어졌다. 먼저 도서관의 경우, PM-10 평균 농도는 78.4 µg/m<sup>3</sup>으로서 낮은 농도값을 나타내었다. 도서관의 경우 PM-10 오염원이 거의 존재하지 않았으며, 도서관 내에서의 금연과 빈번한 물청소 등의 영향으로 생각되었고, 대부분 실외 대기에서 유입되었을 것으로 사료된다. 학교 내의 연구실은 55.5 µg/m<sup>3</sup>으로 조사되어 상대적으로 양호한 공기질을 유지하고 있었다.

Table 3. PM-10 concentrations in each sampling site.

Type of indoor	Sampling site	PM-10 Conc. (µg/m <sup>3</sup> )
Entertainment place	Billiard room	167.3
	Cartoon room	120.9
	Singing room	86.5
	Game room	117.0
	<b>AVG.</b>	<b>128.3</b>
Restaurant	Restaurant	148.3
	Beer hall	270.0
	Coffee shop	110.2
	<b>AVG.</b>	<b>193.9</b>
Store	Complex store	153.8
School	Library	78.4
	University	55.5
	<b>AVG.</b>	<b>63.1</b>
Office	Office	57.5
Residential area	Homes	107.9
Underground structures	Subway	146.4

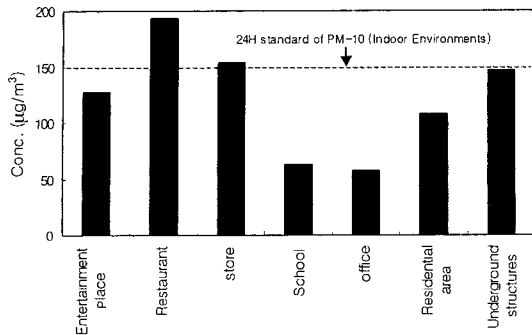


Fig. 1. Average PM-10 concentrations in the various indoor environments.

사무실의 경우도 PM-10의 농도값이 57.5 µg/m<sup>3</sup>를 나타내었다. 이 사무실의 경우 특이한 오염원이 존재하지 않았으며 환기가 잘 이루어지고 있었고, 외부 주변에도 특이한 오염원이 존재하지 않았다.

가정집의 PM-10 평균 농도는 107.9 µg/m<sup>3</sup>으로 조사되었다. 가정집의 경우는 실내에 특이한 오염원이 존재하지 않으며 측정장소는 바로 앞에 한남대교가 위치하고 있어 차량통행량이 매우 많은 곳으로서 자동차 통행에 의해 발생된 PM-10이 실내로 유입된 것으로 사료되었다.

지하 시설물의 경우는 서울시 지하철의 승강장과

대합실을 대상으로 PM-10 농도값을 조사한 결과 146.4 µg/m<sup>3</sup>으로서 기준치에 육박하는 농도를 보였다. 이는 실외 대기 중 부유분진이 지하역사로 유입되거나 이용 승객들의 신발, 옷 등에 묻은 분진이 지하철의 열차풍에 의한 비산, 지하철과 레일과의 마찰 등에 의해 비교적 높은 농도를 나타내는 것으로 사료되었다.

### 3.2 무기원소의 농도분포 현황

연구기간 중 mini-vol portable sampler를 이용하여 다양한 실내공간에서 채취한 PM-10을 질산 전처리법으로 전처리 한 후 ICP-MS로 무기원소의 농도를 분석하였다. 측정항목은 B, Na, Mg, Al, K, Ca, Ti, V, Cr, Fe, Ni, Cu, Zn, As, Se, Cd, Ba, Ce, Pb 등 총 19 항목으로, 다양한 실내공간에서의 19개 무기원소의 평균 농도를 표 4에 나타내었으며, 다양한 실내공간에서 각 무기원소의 농도값을 그림 2에 나타내었다. 그림 2의 (a)~(h)는 농도의 경향과 수준이 비슷한 무기원소를 분류하여 도식한 것이다. 분석된 19개 항목 중 B, Na, Mg, Al, Ti, V, Fe, Cu, As, Ba, Pb 등은 지하 시설물에서, Zn의 경우는 가정집에서, Cr, Ni은 사무실에서, K는 학교에서, Ca는 상가의 실내공간에서, Cd, Ce는 음식점에서, Se는 유흥업소에서 가장 높은 농도를 나타냈다.

Table 4. Summary of inorganic species concentration data in the various indoor environments. (unit : ng/m<sup>3</sup>)

	Entertainment place	Restaurant	Store	School	Office	Residential area	Underground structures
B	29.8	57.7	3.1	57.1	539.9	53.2	884.6
Na	811.8	1,011.5	880.2	1,081.6	N.D	497.8	1,267.6
Mg	183.8	98.8	581.5	38.6	8.9	161.5	647.5
Al	204.1	399.0	982.6	316.9	1,096.8	1,293.7	3,057.7
K	64.8	74.8	75.7	186.5	N.D	47.3	78.3
Ca	1,036.2	673.2	5,484.7	2,196.0	N.D	4,820.3	3,609.4
Ti	8.2	14.4	43.4	7.4	5.3	23.0	94.1
V	0.7	2.6	2.3	0.7	0.6	1.5	5.3
Cr	N.D	2.4	3.8	12.0	139.9	26.2	57.6
Fe	6.6	301.8	464.7	91.6	486.6	463.1	16,844.4
Ni	2.5	14.0	8.3	12.5	35.0	16.0	22.7
Cu	27.9	38.7	43.2	9.6	1.3	8.3	651.4
Zn	90.3	111.3	69.2	830.4	27.4	1,730.2	N.D
As	2.0	2.1	1.1	0.2	N.D	N.D	90.2
Se	3.0	2.3	2.7	1.4	0.5	2.6	2.7
Cd	2.4	5.0	2.5	1.6	0.6	0.6	2.2
Ba	19.4	328.5	10.9	36.9	446.9	56.0	744.1
Ce	18.4	30.3	13.0	7.6	1.9	1.4	3.5
Pb	22.4	82.8	39.1	26.2	27.0	21.9	116.2

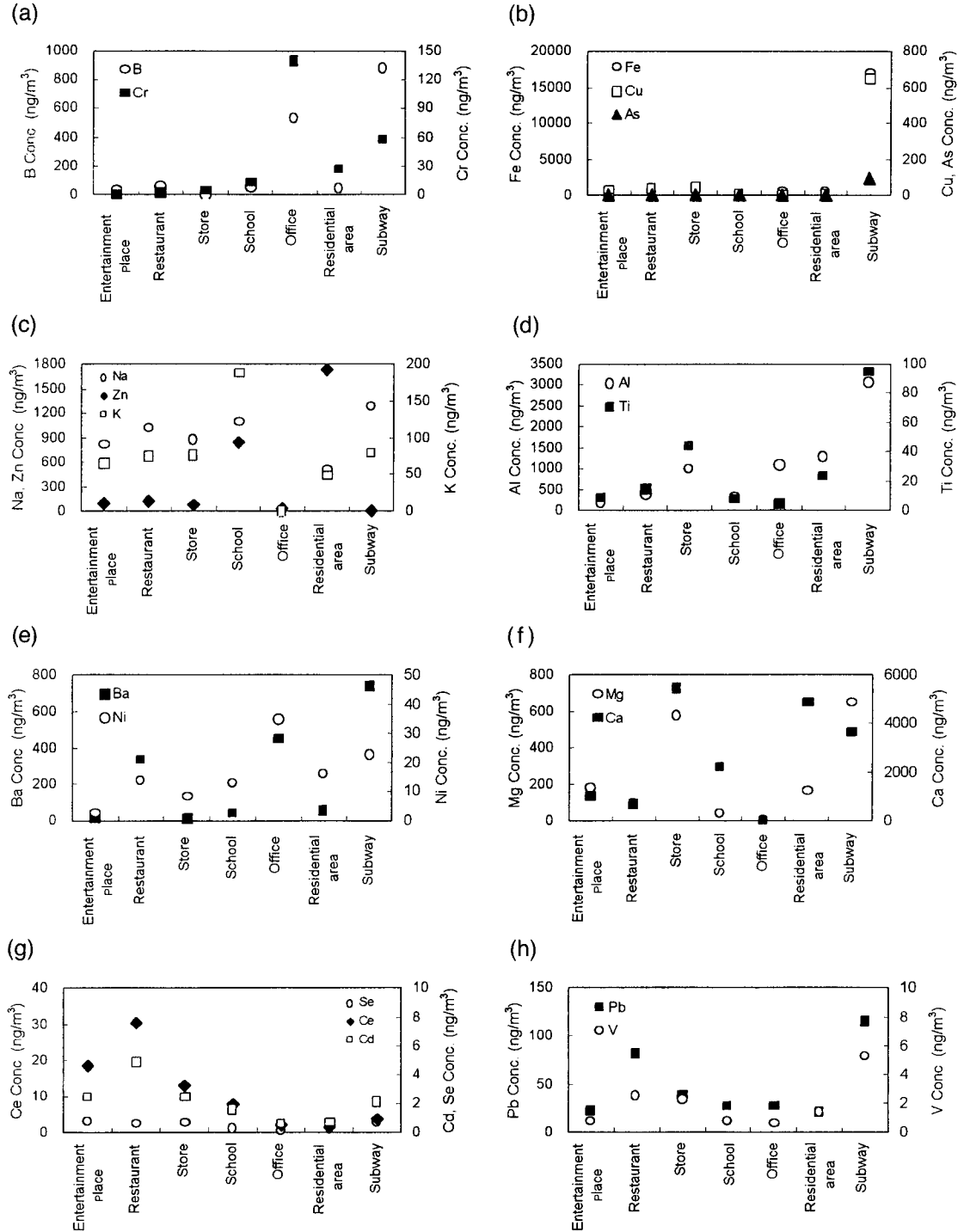


Fig. 2. Average concentration of inorganic elements in the various indoor environments (similar trends and levels of inorganic elements concentration).

유형업소의 실내공간에서 채취한 PM-10 중의 무기원소 농도를 분석한 결과, Ca이 1,036.2 ng/m<sup>3</sup>으로 가장 높은 농도를 나타내었으며 Na 811.8 ng/m<sup>3</sup>, Al 204.1 ng/m<sup>3</sup>, Mg 183.8 ng/m<sup>3</sup>의 순으로 조사되었다. 특히, 당구장의 경우는 Ca의 농도값이 다른 장소에 비해 아주 높은 농도로 조사되었는데, 이는 당구장에서의 초크(chalk) 사용에 의한 것으로 사료된다. 음식점의 경우는 Na가 1,011.5 ng/m<sup>3</sup>으로 가장 높은 농도로 조사되었으며, Ca, Al, Ba, Fe 등이 비교적 높은 농도값을 나타내었다. 상가의 경우는 Ca가 5,484.7 ng/m<sup>3</sup>으로 가장 높은 농도값을 나타내었으며, Al, Na, Mg, Fe 등도 비교적 높게 조사되었다. 이 장소는 시료채취 기간동안에 내부공사가 진행 중이었으며 상술한 바와 같은 무기원소들의 농도가 높게 조사된 것으로 보아 토양에서 기원한 것으로 사료된다. 학교의 실내공간에서는 Ca가 2,196.0 ng/m<sup>3</sup>으로 최고농도를 보였으며, Na, Zn, Al, K 등이 높은 농도값을 나타내었다. 사무실의 경우는 Al이 1,096.8 ng/m<sup>3</sup>으로 가장 높은 농도로 조사되었으며, 그 다음으로 B, Fe, Ba, Cr 등이 높은 농도값을 나타내었다. 주거지역의 실내공간에서는 Ca가 4,820.3 ng/m<sup>3</sup>으로 가장 높은 농도를, 그 다음으로는 Zn, Al, Na, Fe 등의 순으로 조사되었으며 이 장소는 자동차 통행이 빈번한 곳으로서 자동차 통행에 의해 발생된 분진이 어느 정도 실내로 유입되어 실내 공기질에 영향을 미친 것으로 사료된다. 마지막으로 지하역사의 경우는 Fe가 16,844.4 ng/m<sup>3</sup>으로 가장 높은 농도를 나타내었으며, Ca, Al 등도 높은 농도로 조사되었다. 지하역사는 실내흡연이나 연소 등과 같은 오염원이 존재하지 않으며, Fe가 다른 장소에 비해 매우 높은 값을 나타낸 것은 전동차 운행에 따른 전선, 레일과의 마찰로 인한 것으로 사료된다.

### 3.3 군집분석 결과

실내공간에서의 PM-10 농도경향과 무기원소의 농도경향을 파악한 후, 군집분석을 수행하여 분석된 PM-10 무기원소 자료를 유사한 특성을 지닌 군집으로 분류하였다. 군집분석을 수행하기 전에 먼저 원자료(raw data)에 대한 자료변환을 수행하였다. 변수의 변량분포가 대칭형을 나타내지 않고 어느 한 쪽으로 치우쳐 있을 경우 통계분석의 결과가 과장되거나 오류가 발생할 가능성이 있기 때문에 본 연

구에서는 먼저 원자료에 대한 빈도분석을 실시하여 자료의 변수별 변량분포를 살펴보았으며, 그 결과 원자료가 한쪽으로 심하게 치우쳐 있어 군집분석에 앞서 자료변환을 실시하여 자료들의 분포를 정규분포화 해야 할 필요성이 제기되었다. 원자료에 대한 자료변환은 로그(logarithmic)변환과 제곱근(square root)변환의 두 가지 방법을 수행하였다. 두 가지 방법의 자료변환 중 군집분석 수행시 로그변환한 자료가 더 좋은 결과를 나타내어 본 연구에서는 이러한 결과를 바탕으로 로그변환한 자료를 입력자료로 이용하여 군집분석을 수행하였으며 또한 z-score를 이용하여 자료의 표준화를 수행하였다(식 3).

$$Z_{ik} = \frac{(X_{ik} - X_i)}{S_i} \quad (3)$$

여기에서  $Z_{ik}$ 는  $X_{ik}$ 의 표준화된 값이며  $X_i$ 와  $S_i$ 는 각각  $i$ 번째 변수의 평균과 표준편차를 의미한다.

본 연구에서는 다양한 군집분석 방법 중 비교적 algorithm이 간단하고 계산이 간편하며, 보편적으로 가장 많이 사용되는 응집위계분석법을 적용하였으며 비유사도의 기준을 결정하기 위해 Euclidian 거리를 이용하여 군집분석을 수행하였다. 그리고 분류된 그룹이나 case 사이의 거리를 표시하는 군집화 방법으로서 Ward's method를 이용하였다. 군집분석에서 비유사도 거리를 정하고 그것을 바탕으로 군집의 수를 정하는 것은 군집분석시 가장 중요한 과정이며, 이 과정에는 연구자의 주관이 개입되기 때문에 최대한 객관적 입장에서 군집의 수를 결정하는 것이 중요하다. 본 연구에서는 비유사도 거리 10을 기준으로 4개의 군집으로 분류하였으며, 그 결과를 그림 3과 같은 수상도(dendrogram)로 나타내었다. 군집 1은 11개의 object (ID : 3, 21, 14, 17, 1, 7, 8, 13, 16, 15, 10), 군집 2는 7개의 object (ID : 37, 40, 12, 20, 36, 38, 39), 군집 3은 10개의 object (ID : 27, 34, 28, 33, 18, 31, 30, 35, 29, 32), 군집 4는 12개의 object (ID : 2, 6, 5, 23, 25, 24, 26, 9, 11, 19, 22, 4)로 분류되었으며 어느 군집에도 포함되지 않는 이상치(outlier)는 분류되지 않았다.

### 3.4 SIMCA 분석 결과

군집분석의 결과로 생성된 군집 1~군집 4를 각각 training set으로 하고 분산주성분 분석을 수행하



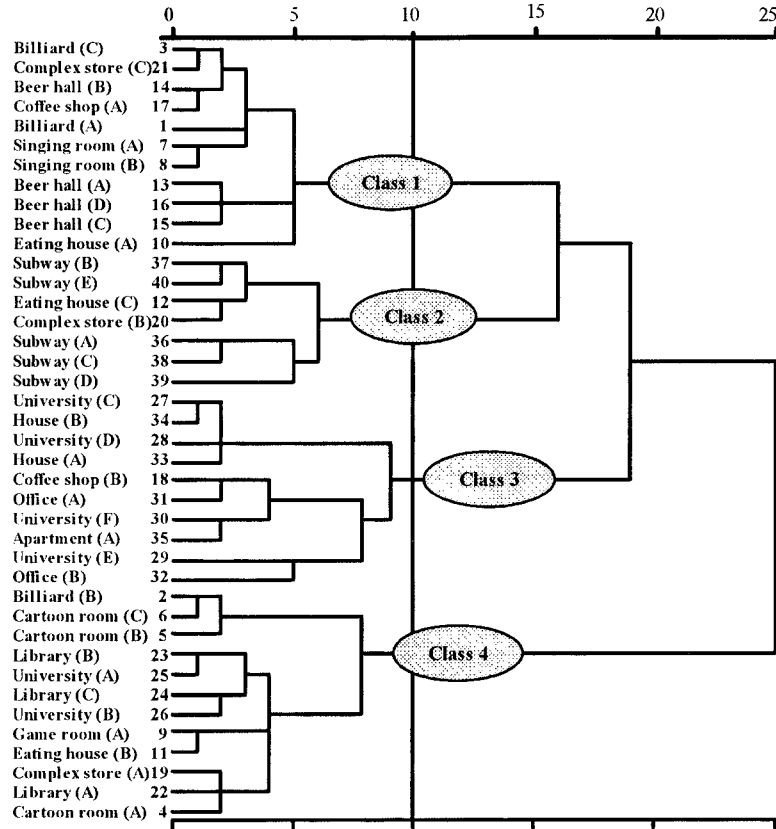


Fig. 3. A dendrogram using the euclidean distance after logrithmic transformation.

였다. 분산주성분 분석의 첫 번째 단계는 횡유효도 (cross validation) 검사를 수행하여 유효한 주성분의 수를 결정하는 것인데, 각 군집에 소속된 object의 수가 3개 이하로 분류되면 유효한 주성분의 결정되지 않았으며 이러한 결과는 군집분석이 올바르게 수행되지 않았음을 의미하는 것이다(유상준과 김동술, 1997). 본 연구의 군집분석 결과에서 각 군집의 object 수는 3개 이상씩 분류되어 유효 주성분의 수를 결정하는데는 문제점이 발생하지 않았다.

각 군집에 대한 유효 주성분의 수가 결정이 되면, 그 다음 단계로 군집과 object에 대한 잔여 표준편차를 구하게 되며, 계산된 값을 이용하여 그림 4와 같이 Cooman's 표준결정도 (standard decision plot)를 작성할 수 있었다. 그림에서 수직과 수평으로 그은 선은 95% 확률의 임계거리 선 (critical distance line)으로서 이 선에 의해 4개의 영역으로 나누어진다.

그림 4의 a)는 군집 1과 군집 4에 대한 표준결정도를 도식한 것으로서, 95% 확률의 임계거리 선에 의해 4개의 영역으로 나누어지는 것을 알 수 있다. 그림 왼쪽 상단 영역은 군집 1의 object들만이 존재하는 영역이고, 오른쪽 하단 영역은 군집 4의 object들만이 존재하는 영역이다. 왼쪽 하단 영역은 두 군집의 object가 공존하는 영역이고, 오른쪽 상단 영역은 군집 1과 군집 4를 제외한 그 밖의 군집 object가 존재하는 영역이다. 이 그림에서는 군집 2와 군집 3의 object들이 오른쪽 상단에 위치한다. 따라서 좌측 상단 및 우측 하단 영역에 object가 존재하면 95% 확률로 해당 군집에 속하게 되는 것이고 그 이외의 영역에 object가 존재하면 해당 군집에서 배제된다. 이와 같은 방법으로 그림 4의 a)~f)와 같은 표준결정도를 이용하여 군집분석 수행시 주관적인 군집수의 결정, 즉 비유사도의 임의적 결정을 객관화하

여 순수한 군집 (homogeneous cluster)을 도출하였으며 이를 표 5에 나타내었다. 또한, 각 군집별 object에 대한 PM-10 농도와 무기원소들의 농도를 평균하여 표 6에 나타내었다. 순수군집의 수는 5개(군집 99 포함)로 결정되었으며 이들 순수군집들은 다양한 실내공간에서의 오염형태와 특성을 나타낸다고 할 수 있다.

군집 1의 경우는 PM-10 농도와 무기원소 중 Na, Mg, Ca 등의 농도값이 다른 장소에 비해 높게 조사된 특징을 보이고 있다. 군집 1로 분류된 각 object들의 특징은 사람들의 이용빈도와 왕래가 빈번한 곳(당구장 A와 C, 호프집 A, B, C, D, 복합상가 C)으로서, PM-10 평균농도(239.5 µg/m<sup>3</sup>)와 Na 평균농도(1,375.5 ng/m<sup>3</sup>)가 4개 군집 중 가장 높은 값을 나타내었다. 또한 건물자체의 증축공사(복합상가 C), 시료 채취장소 주변에서의 건축공사, 굴착기와 포크레인을 이용한 토목공사(당구장 A, C, 호프집 A, C) 등의 영향으로 Ca, Mg, Al, K 등의 토양기원 무기원소들의 농도값이 비교적 높게 조사된 것으로 사료된다.

군집 2의 object들은 식당 C, 복합상가 B, 그리고 지하역사 A, B, C, D, E가 포함되어 분류된 object 대부분이 지하 공간(지하철의 승강장, 대합실)이라는 특징을 보이고 있다. 군집 2는 PM-10 농도가 165.8 µg/m<sup>3</sup>으로서 비교적 높고 특히, B 808.0 ng/m<sup>3</sup>, Mg 672.2 ng/m<sup>3</sup>, Al 3,177.9 ng/m<sup>3</sup>, Ca 4,458.3 ng/m<sup>3</sup>, Fe 14,162.8 ng/m<sup>3</sup>, Cu 554.3 ng/m<sup>3</sup>, Ba가 601.5 ng/m<sup>3</sup>으로서 4개 군집 중 가장 높은 농도값을 나타내었다. 이는 지하철 이용 승객들이 옷, 신발 등에 묻혀 외부에서 실내로 유입되는 토양성분과 전동차 운행에 따른 전동차와 전선, 전동차와 레일과의 마찰로 의해 Fe, Cu 등이 다른 군집에 비해 높은 농도를 나타내는 것으로 사료된다.

군집 3으로 분류된 object들은 coffee shop B, 학교 C, D, E, F, 사무실 A, 가정집 A, B, C 등으로 조사되었다. 군집 3의 PM-10 농도는 4개의 군집 중 가장 낮은 농도값인 29.2 µg/m<sup>3</sup>으로 조사되어 가장 양호한 공기질을 나타내는 것으로 조사되었다. 또한 Na가 570.6 ng/m<sup>3</sup>, Al 572.4 ng/m<sup>3</sup>, Ca 4,231.4 ng/m<sup>3</sup> 등이 비교적 높은 농도값을 나타내었는데 이러한 사실로 외부의 토양성분들이 내부로 유입되었을 가능성이 존재한다고 사료된다. 특히, Zn의 경우는

Table 5. Classified objects after disjoint principal component analysis.

Class	No. of samples	Objects
1	7	1, 3, 13, 14, 15, 16, 21
2	7	12, 20, 36, 37, 38, 39, 40
3	8	18, 27, 28, 30, 31, 33, 34, 35
4	11	2, 4, 5, 6, 9, 11, 19, 22, 23, 25, 26
99	7	7, 8, 10, 17, 24, 29, 32

Table 6. Average concentrations of PM-10 and inorganic compounds in each class after disjoint principal component analysis. (Unit : ng/m<sup>3</sup>)

	Class 1	Class 2	Class 3	Class 4
PM-10*	239.5	165.8	29.2	32.4
B	19.2	808.0	25.9	19.4
Na	1,375.5	1,311.8	570.6	351.3
Mg	341.7	672.2	95.1	24.5
Al	229.0	3,177.9	572.4	98.7
K	114.2	83.3	202.5	25.0
Ca	2,482.1	4,458.3	4,231.4	161.1
Ti	2.4	107.1	14.3	5.1
V	2.0	5.7	0.7	0.2
Cr	0.3	49.9	12.7	N.D
Fe	14.1	14,162.8	165.5	N.D
Ni	11.8	21.5	8.9	8.0
Cu	48.1	554.3	5.0	11.7
Zn	114.3	45.8	1,506.7	N.D
As	2.9	75.9	N.D	1.2
Se	2.1	3.5	0.6	2.3
Cd	5.0	3.9	16.1	0.6
Ba	27.5	601.5	1.8	N.D
Ce	39.4	9.1	N.D	9.2
Pb	58.2	121.3	21.6	6.5

\* : µg/m<sup>3</sup>

1,506.7 ng/m<sup>3</sup>으로서 다른 군집에 비해 월등히 높은 값을 나타내는 것으로 조사되었는데, 이는 가정집 A와 가정집 B에서 측정된 자료의 영향으로 이 장소는 한남대교 부근에 위치하고 있으며 한남대교의 많은 자동차 통행량과 자동차 정체에 따른 타이어 마모, 브레이크 마모 등에 의한 영향으로 사료된다.

군집 4의 object들은 당구장 B, 만화방 A, B, C, game room, 식당 B, 복합상가 A, 도서관 A, B, 학교 A, B 등으로 조사되었다. 군집 4의 PM-10 농도는 32.4 µg/m<sup>3</sup>으로 조사되어 비교적 양호한 공기질을 나타내는 것으로 조사되었다. 무기원소들의 농도값은 다른 군집에 비해 전체적으로 낮은 값을 나타내는 것으로 조사되었다.

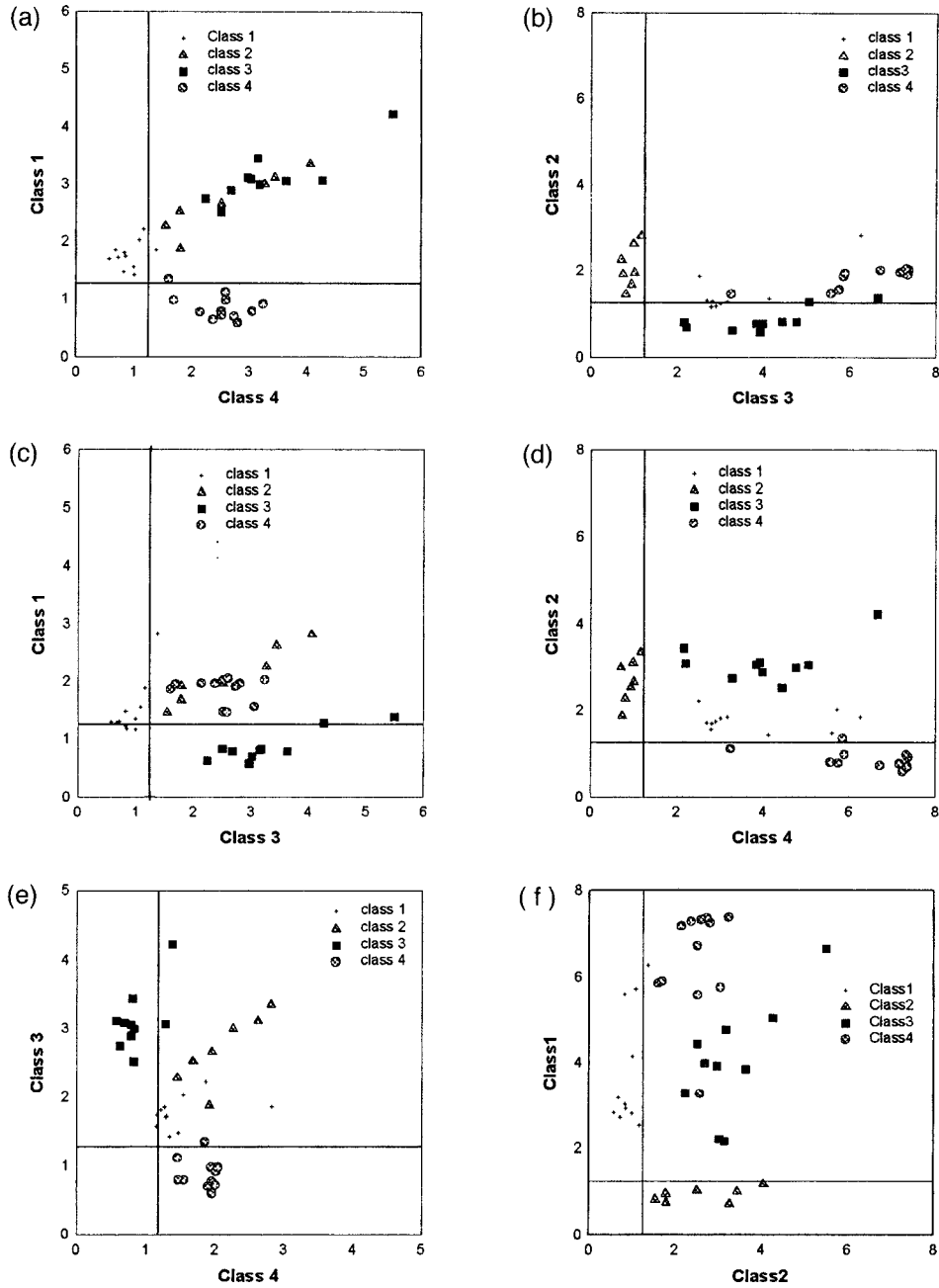


Fig. 4. Cooman's standard decision plot between two specific classes.

4. 결 론

다양한 실내공간(가정, 사무실, 학교, 지하시설물,

상가, 음식점, 지하철 등) 중 총 40곳의 측정지점에서 2000년 4월부터 2000년 10월까지 mini-vol portable sampler에 채취된 PM-10 입자의 무기원소를

분석하여 다양한 실내공간에서의 PM-10 오염현황을 파악하였으며, 군집분석법과 분산주성분 분석법을 이용하여 PM-10 무기원소 자료를 유사한 특성을 지닌 군집으로 분류하고 분류된 군집을 확률적으로 검증한 뒤, 최종적으로 산출된 순수군집의 특징과 실내환경 중 PM-10 오염의 패턴을 파악하고자 하였다.

본 연구에서는 군집분석법과 분산주성분 분석을 동시에 응용한 형태인식기법을 적용하여 다양한 실내공간에서 채취한 PM-10에 대하여 총 4개의 군집 pattern으로 분류할 수 있었다. 이들 군집들에 대해서 95% 확률로 검증을 한 후 비유사도의 임의적 결정을 객관화하고 통계적으로 양호한 4개의 순수군집을 도출할 수 있었다. 먼저, 군집 1은 PM-10 농도와 무기원소 중 Na, Mg, Ca 등의 농도값이 다른 장소에 비해 높게 조사된 특징을 보이고 있었다. 군집 2의 경우는 분류된 object 대부분이 지하 공간(지하철의 승강장, 대합실)이라는 특징을 보이고 있었다. 군집 3의 PM-10 농도는 4개의 군집 중 가장 낮은 농도값인  $29.2 \mu\text{g}/\text{m}^3$ 으로 조사되어 가장 양호한 공기질을 나타내는 것으로 조사되었으며, Zn의 경우 다른 군집에 비해 월등히 높은 값을 나타내는 것으로 조사되었다. 마지막으로 군집 4의 PM-10 농도는  $32.4 \mu\text{g}/\text{m}^3$ 으로 조사되어 비교적 양호한 공기질을 나타내는 것으로 조사되었으며 무기원소들의 농도값은 다른 군집에 비해 전체적으로 낮은 값을 나타내는 것으로 조사되었다. 이들 순수군집들은 다양한 실내공간에서의 오염형태와 오염특성을 나타낸다고 할 수 있다. 즉, 이러한 각각의 순수군집들에 포함된 object(시료)들은 동일 오염원 또는 유사 오염원에서 기인한 오염물질이 영향을 미쳤을 가능성이 크다고 할 수 있다.

본 기초연구는 다양한 실내환경에서 오염도의 패턴분류를 시도한 것으로 다양한 실내환경을 분류하고 각 환경에 대한 기준설정 및 오염항목선정에 큰 도움을 줄 것으로 사료된다. 또한 추후 시료의 실내 채취장소를 좀 더 다양화하고, 시료의 채취시간을 계절별로 확대한다면, 국내에 실내공기질법이 도입될 때 대상시설물의 기준 및 항목선정에 도움을 줄 것으로 사료된다.

## 사 사

본 연구의 일부는 1999년 한국학술진흥재단 대학부설연구소 지원과제(과제번호 : 99-005-E00025)의 일환으로 수행되었으며, 이에 감사드립니다.

## 참 고 문 헌

- 김기영, 전명식(1997) 다변량 통계자료분석, 자유아카데미.  
 김동술, 김형석(1990) Pattern Recognition을 이용한 지하상가에서의 대기오염물질의 농도 분석에 관한 연구, 한국대기보전학회지, 6(1), 1-10.  
 대기환경연구회(2000) 대기환경개론, 동화기술.  
 유상준, 김동술(1997) 군집분석과 분산 주성분분석법을 이용한 대기분진시료의 분류, 한국대기보전학회지, 13(1), 51-63.  
 환경부(1998) 지하생활공간 공기질 관리법.  
 황인조, 김동술(1998) Submicron 부유분진의 화학적 조성 및 분포에 관한 연구, 한국대기보전학회지, 14(1), 11-23.  
 황인조, 김태오, 김동술(2001) PMF 방법론을 이용한 수원 지역 PM-10의 오염원 확인, 한국대기환경학회지, 17(2), 133-145.  
 Chapin, F.S.(1974) *Human Activity Patterns in the City*, John Wiley & Sons, New York.  
 Dyremark, A., R. Westerholm, E. Övervik, and J. Gustavsson (1995) Polycyclic aromatic hydrocarbon (PAH) emissions from charcoal grilling, Atmospheric Environment, 29(13), 1553-1558.  
 Gomez, M.L.S. and M.C.R. Martin (1987) Application of cluster analysis to identify sources of air-borne particles, Atmospheric Environment, 21(7), 1521-1527.  
 Hopke, P.K. (1985) *Receptor Modeling in Environmental Chemistry*, Wiley Interscience, New York.  
 Hopke, P.K. (1991) *Receptor Modeling for Air Quality Management*, Elsevier Science Publishing Company Inc., New York.  
 Szalazi, A. (1972) *The Use of Time, Daily Activities of Urban and suburban Population in Twelve Countries*, Mougton, Mougton, The Hague, Paris.  
 SIMCA-3B Manual (1984) A pattern recognition program for CPM and MS-DOS based microcomputers, principal components models, 2505 Shepard Blvd, Columbia, MO.

- Wiley, J.A., P.J. Robinson, T. Piazza, K. Garrett, Y. Cheng, and G. Marti (1991) *Activity Patterns of California Residents*, Final Report for the Research Division, California Air Resources Board. Contact N. Sacramento, CA, A6, 173.
- Wold, S. (1976) Pattern recognition by means of disjoint principal components models, *Pattern Recognition*, 8, 127-139.
- Wold, S. and M. Sjöström (1977) SIMCA : a method for analyzing chemical data in terms of similarity and analogy, *Chemometrics theory and application*, edited by B.R. Kowalski, ACS symposium series 52, 243-282.
- Wold, S., C. Albano, W.J. Dunn III, U. Edlund, K. Esbensen, P. Geladi, S. Hollberg, E. Johansson, W. Lindberg, and M. Sjöström (1984) Multivariate data analysis in chemistry, *Chemometrics : mathematics and statistics in chemistry*, edited by B.R. Kowalski, NATO ASI series.