대전시 3, 4공단 미세먼지(PM₁₀)의 반복측정 분산분석

Repeated Measures ANOVA for Fine Dust of Industrial Complex 3 and 4 in Daejeon, South Korea

이경수* · 김태형** · 이재오*** · 이용림****

Lee, Kyung Su*, Kim, Tae Hyeong**, Lee, Jae Ou***, and Lee, Yong Lim****

Abstract

This research analyzes the PM₁₀ concentration of fine dust using data provided by the Korea Environment Corporation for the Munpyeong-dong Station located at industrial complexes 3 and 4 of Daejeon, South Korea. The change in PM₁₀ concentration was analyzed by separating seasonal changes of PM₁₀ in 2017 and dividing the day's 24 hours into 4 segments for 5 years from 2013 to 2017. The PM₁₀ concentration was observed to be highest during the morning (6–12 a.m.) followed by dawn (0–6 a.m.), evening (18 p.m.–24 a.m.), and afternoon (12–18 p.m.), which showed the lowest PM₁₀ concentrations. By year, January showed the highest fine dust concentration, except in 2015, and August showed the lowest concentration. Throughout the years, Januarys and Augusts showed a decreasing trend, and April of 2015 showed the highest concentration change. Various results could be obtained by extending and connecting industrial activities related to fine dust. Big data analysis for fine dust is considered necessary. Future research observing the occurrence pattern of fine dust, including ultrafine dust, by year and time segment is also needed.

Key words : Fine Dust, PM₁₀, Repeated Measures of ANOVA, Python, Big Data

요 지

본 연구에서는 대전시의 3, 4공단에 위치한 문평동 측정소를 대상으로 환경관리공단에서 제공하는 미세먼지 데이터를 활용하여 PM₁₀ 농도를 분석 하였다. 분석범위는 2013년부터 2017년까지 최근 5년간 계절에 따른 영향과 하루 24시간을 4개 시간대역으로 나누어 PM₁₀ 농도의 변동 특성을 분석하였다. 분석결과, PM₁₀ 농도는 3월, 4월, 5월, 11월이 가장 높게 나타났으며, 8월이 가장 낮게 나타났다. 시간대별 PM₁₀ 농도는 오전시간(6~12시)이 가장 높은 것으로 나타났다. 다음으로는 새벽시간(0~6시), 저녁시간(18~24시)이 높았고, 오후시간(12~18시)이 전반적으로 가장 낮은 PM₁₀ 농도를 나타냈다. 연도별로는 1월이 2015년도 를 제외하고 가장 높은 미세먼지 농도를 나타냈으며, 8월이 가장 낮은 미세먼지 농도를 나타냈다. 1월과 8월은 전체적으로 감소하는 경향을 나타냈지만 4월은 2015년 가장 높은 미세먼지 농도 변화를 나타냈다. 이와 같은 결과는 미세먼지와 관련되는 산업 활동의 전반적인 데이터를 확장 및 연결하여 다양한 결과를 이끌어 낼 수 있을 것이며, 이런 미세먼지와 관련된 다양한 빅데이터 분석 연구가 필요한 시점이라 사료된다. 향후에는 초미세먼지를 포함하여 연도별, 시간대별로의 미세먼지 발생 패턴에 관한 연구가 필요할 것이다.

핵심용어 : 미세먼지, PM₁₀, 반복측정분산분석, 파이썬, 빅데이터

*정회원, 국립재난안전연구원 시설연구사(E-mail: 39lks@korea.kr)

Member, Research Officer, National Disaster Management Institute, Ministry of the Interior Safety

**교신저자, 정회원, 한국물환경주식회사 대표(Tel: +82-42-670-7143, Fax: +82-504-831-7143, E-mail: rucachi@naver.com)

Corresponding Author, Member, Representative Director, Korea Water & Environment Co., Ltd.

***정회원, 대전대학교 소방방재학과 교수(E-mail: krokasig@dj.kr)

Member, Professor, Department of Fire and Disaster Prevention, Daejeon University

****정회원, 한국물환경주식회사 연구원(E-mail: younglimlee@naver.com)

Member, Researcher, Korea Water & Environment Co., Ltd.

1. 서 론

대기환경보전법에서는 대기오염물질을 아황산가스, 오존, 일산화탄소 등의 가스성 물질과 먼지 등 입자상 물질을 포함하여 총 61종으로 정의하였다. 그 중 미세먼지는 PM_{10} 과 $PM_{2.5}$ 두 가지 종류가 있으며, PM_{10} 은 입경이 $10 \mu g/m^3$ 인 미세먼지 $PM_{2.5}$ 는 입경이 $2.5 \mu g/m^3$ 인 초미세먼지이다. 초미세먼지는 세계보건기구(World Health Organization, WHO)가 2013년 10월에 1군 발암물질로 분류하였다. 이런 미세먼지는 식물의 잎 표면에 침착되어 신진대사를 억제하고 건축물에 퇴적되어 부식을 일으키며, 사람이 많이 흡입하면 호흡기질환과 염증 반응을 일으키고 심혈관계 질환을 증가시킨다고 보고되고 있다.

미세먼지는 단일 물질이 아니라 다양한 원인에 의해 생성되는 미세한 고체와 액체입자이다. 따라서 미세먼지는 발생원인에 따라 1차 미세먼지와 2차 미세먼지로 구분할 수 있다. 1차 미세먼지는 연료연소시설의 굴뚝, 자동차 배기구, 바람에 날리는 먼지 형태로 많이 배출되고, 2차 미세먼지는 대기중에서 이미 배출되어 있는 황산화물, 질산화물, 암모니아가 화학반응을 통하여 생성되는 질산암모늄이나 황산암모늄 같은 에어로졸과 대기 중에서 응축되는 유기물이 여기에 해당된다. 1차 미세먼지에 대해서는 대기오염배출량을 산출할 수 있지만 2차 미세먼지에 대해서는 직접적으로 배출량을 산출하기 어렵다(Jang, 2016).

미세먼지의 원인분석에 관한 연구는 2010년 초반부터 활발히 진행되기 시작했다. Lee et al. (2011)은 제주도 고산지역에서 PM_{10} 과 $PM_{2.5}$ 시료를 채취하여 주요 성분의 화학조성, 오염정도, 배출원 특성 등을 분석하였다. 분석결과, $PM_{2.5}$ 질량농도가 PM_{10} 의 약 61%를 차지하며 미세먼지 조성은 에너지 사용에 따른 인위적 오염원의 영향이 가장 크고, 다음으로 토양 또는 해양의 영향을 많이 받는 것으로 나타났다. 미국 EPA (2014)의 입자상 물질과 건강영향을 종합하여 분석한 연구에서는 $PM_{2.5}$ 보다 PM_{10} 으로 인한 초과사망(excess death)의 발생위험이 높은 것으로 나타났다. Zhou (2014)은 서울시의 연평균 PM_{10} 과 $PM_{2.5}$ 농도를 분석하였으며, 각각 $41 \mu g/m^3$, $23 \mu g/m^3$ 로 PM_{10} 이 연평균 농도가 더 높은 것을 확인하였다. PM_{10} 과 $PM_{2.5}$ 각각의 월 평균 농도는 1월이 $60 \mu g/m^3$, $35 \mu g/m^3$ 로 가장 높았으며, 8월이 $23 \mu g/m^3$, $14 \mu g/m^3$ 로 가장 낮은 농도를 보였다. 또한 시간대별 미세먼지농도는 밤 10시(최고농도)와 아침 6시(최저농도)를 기점으로 변화되었다. 한국국토정보공사(LX, 2016)에서는 빅데이터를 활용하여 전라북도의 미세먼지 원인을 신경망 분석, Support Vector Machine, 앙상블 등 데이터 마이닝 기법을 활용하여 분석하였다. 분석 결과 미세먼지 영향력은 공장 50%, 지형 27%, 도로 18%, 인구 5%로 나타났으며, 공장의 배출시설 관리 강화와 경유차 이동 밀집 지역에 대한 관리 강화 등의 정책적 결과를 제시하였다. Kim (2013)

은 미세먼지(PM_{10} , $PM_{2.5}$)의 발생원인을 공업지역과 상업지역으로 구분하여 분석하였다. 분석결과, PM_{10} 평균농도는 공업지역에서 $64.8 \mu g/m^3$, 상업지역에서 $54.0 \mu g/m^3$ 로 공업지역이 1.2배 정도 높게 나타났으며, 일별 측정값의 농도는 공업지역에서 $44.0 \sim 92.0 \mu g/m^3$, 상업지역에서 $5.3 \sim 91.1 \mu g/m^3$ 의 범위를 나타내었다. $PM_{2.5}$ 의 평균농도는 공업지역에서 $69.4 \mu g/m^3$, 상업지역에서 $32.7 \mu g/m^3$ 로 공업지역이 상업지역보다 2.1배 정도 높게 나타났으며, 일별 측정값의 농도는 공업지역에서 $41.8 \sim 97.0 \mu g/m^3$, 상업지역에서 $1.2 \sim 79.8 \mu g/m^3$ 의 범위를 나타내었다. 이처럼 PM_{10} 과 $PM_{2.5}$ 의 농도가 모두 공업지역에서 높게 나타났다. 이것은 공업지역의 연료사용량이 상업지역보다 많으며, 각종 연소시설과 생산 공정에서 오염물질 배출량 또한 많아 미세먼지 농도가 높게 나타난 것이라고 볼 수 있다. Kim and Seo (2013)는 서해안 대천시 미세먼지(PM_{10})를 채취와 이온, 중금속, 탄소농도 분석을 통하여 미세먼지의 화학성분 특성을 파악하였다. Park and Shin (2017)은 국내 초미세먼지($PM_{2.5}$) 농도에 미치는 영향요인을 측정하기 위해 국내 및 중국 측 요인으로 변수를 구분하여 2015년 2월부터 2016년 5월까지 국내 16개 광역시·도를 대상으로 분석한 결과 중국 산동성 지역 초미세먼지 농도와 서풍계열 풍향비율은 한국의 초미세먼지 농도에 (+)영향을 주는 것으로 나타났다. 그러나 국내 요인으로 설정한 석탄화력 발전 거래량과 경유 소비량, 화학 산업 제조업 생산지수의 유의미한 영향은 발견하지 못하였다. 한편, 국내 시멘트 생산지수는 초미세먼지 농도에 (+)영향을 주는 것으로 나타났다.

2019년부터 공개되고 있는 국립환경과학원 국가오염물질배출량서비스(National Air Pollutants Emission Service)에 의하면 2015년도에 PM_{10} 과 $PM_{2.5}$ 의 주요 배출원은 비산먼지가 47.1%로 가장 높고 제조업 연소가 30.5%이다. $PM_{2.5}$ 는 제조업 연소가 36.9%로 가장 높고 비산먼지가 17.5%, 비도로 이동오염원이 14.2%, 생물성 연소가 12.2% 순으로 나타났다. $PM_{2.5}$ 가 PM_{10} 보다 더 다양한 원인으로 배출되는 것으로 나타났다(Table 1). 배출원의 대표적인 원인을 보면 에너지 산업연소는 공공발전시설, 지역난방시설, 석유정제시설, 민간발전시설 등이며 비산업연소에는 상업 및 공공기관 시설, 주거용 시설, 농업 축산 수산업시설 등이다. 제조업 연소에는 연소시설, 공정로 등이며, 생산 공정은 석유제품, 제철제강업, 무기화학제품제조업, 유기화학제품제조업, 목재, 펄프제조업, 식음료가공 등이다. 에너지수송 및 저장은 휘발유공급 등이며 유기용제사용은 도장시설, 세정시설, 세탁시설, 기타유기용제 사용이다. 도로이동오염원은 승용차, 택시, 승합차, 버스, 화물차, 특수차, 이륜차 등이며 비도로 이동오염원은 철도, 선박, 항공, 농업기계, 건설장비 등이다. 폐기물처리하는 폐기물소각 및 기타 폐기물처리이며, 기타 먼 오염원이 있다.

이처럼 미세먼지의 배출원인은 다양한 원인에 의해 발생

Table 1. Current Status of PM₁₀ and PM_{2.5} Emissions by Major Sources in 2015

division	PM ₁₀		PM _{2.5}	
	tons/yr	%	tons/yr	%
Combustion of energy industry	4,394	1.8	3,607	3.8
Non-industrial combustion	1,582	0.6	1,025	1.0
Manufacturing combustion	70,893	30.5	36,317	36.9
Production process	6,658	2.9	5,132	5.1
Road transport pollution source	9,583	4.1	8,817	8.9
Non-road moving pollution source	15,320	6.6	14,106	14.2
Waste treatment	340	0.1	209	0.2
Area source	317	0.1	285	0.2
Fugitive dust	109,633	47.1	17,248	17.5
Biological combustion	14,552	6.2	12,060	12.2
total	233,272	100.0	98,806	100.0

되고 있는 것으로 분석되고 있다. 그러나 현재까지 미세먼지의 원인을 분석한 연구는 공간적인 현황과 내부 환경만 고려된 반면, 시간적인 변화에 대한 고려와 외적요인에 의한 영향은 고려는 하지 못하고 있다. 즉, 산업 활동은 과거에도 진행되었고 현재도 진행되고 있으면 앞으로도 진행될 것이다. 특정 년도에 산업활동이 감소하거나 차량의 이동이 더 증가하거나 그렇지 않다는 것이다. 따라서 본 연구에서는 대전광역시 지역의 공단이 밀집되어 있는 지역을 중심으로 외부 유입에 의한 미세먼지의 증가를 판단하기 위해서 계절적인 영향과 시간대별 영향을 고려하여 PM₁₀의 발생 특성을 분석하기 위하여 2013년부터 2017년의 자료를 대상으로 반복측정 분산분석을 실시하였다. PM₁₀만을 분석하는 이유는 PM_{2.5} 자료에 비하여 비교적 풍부하고 충실하기 때문이다.

2. 대상지역의 특성

2.1 일반현황

대전은 크게 4개의 공단을 가지고 있다. 각 공단은 1, 2공단과 3, 4공단으로 두 개씩 묶여 있다. Fig. 1은 대전시의 대기환경오염 측정소 위치를 보여주고 있다. 1공단은 1969년부터 1973년 그리고 2공단은 1975년부터 1979년 사이에 각각 조성되었다. 3공단은 1989년부터 1992년 그리고 4공단은 1993년부터 1996년까지 조성되었다. 관리 기관은 대전광역시이고 1, 2공단은 금속, 석유화학, 섬유, 식품, 제지 등의 업종이 주를 이루고 있다. 3공단은 기계금속, 섬유, 화학, 제지 등의 업종이 주를 이루고 있으며, 4공단은 의약, 전자, 제지, 식품가공업이 주를 이루고 있다. 대기오염물질 자료는 전국 실시간 대기오염도 공개 시스템인 한국환경공단의 에어코리아(airkorea) 웹사이트를 통해 수집하였고, 측정소는 3, 4공단에 위치한 문평동 측정소를 선정하였다. 문평동 측정소는 3공단에 위치하였으나 4공단과 경계 지점에 있다.

3, 4공단에는 식품품 판매업장, 자동차 부품 및 기계·산업용품 생산·판매 사업장, 전보통신업, 중공업/건설업, 원지제조업, 섬유/무역업, 산업자재 및 화학용품 생산·판매업장이 위치하고 있다. 측정소 근처에는 대전시를 관통하여 흐르는 갑천과 대청댐에서 시작되는 금강 상류가 위치한다. 또 높은 산은 아니지만 세종시 방향으로 불무산(160.8 m), 오봉산(240.6 m) 등 계룡산 산맥이 이어져 있고 3, 4공단의 끝 쪽인 신탄진에는 자동차 부품 제조·판매업장, 담배제조공장, 대전철도차량정비단 등이 넓게 분포되어 있다. 그리고 그 위로 계족산 줄기가 옥천과 경계를 이루고 있는 분지 형태이다. 여기서, 문평동 측정소를 연구 대상으로 선정한 이유는 최초 대전 지역의 다른 측정소와 PM₁₀ 농도를 일원배치분산 분석을 통해 비교하였더니 가장 높게 나타났기 때문이다(Table 2). Table 2는 Fig. 1에서 표시되어 있는 대기환경오염 측정소의 PM₁₀ 평균농도를 비교한 결과이며, 기준년도는 자료를 확보할 수 있는 가장최근 년도인 2017년이다. 이 중에서 문평동 측정소가 52.70 $\mu\text{g}/\text{m}^3$ 로 가장 높게 나타났으며, 대흥동1측정소가 51.63 $\mu\text{g}/\text{m}^3$, 노은동측정소가 47.26 $\mu\text{g}/\text{m}^3$ 순이며, 둔산동측정소의 PM₁₀ 평균농도가 41.39 $\mu\text{g}/\text{m}^3$ 로 가장 낮은 것으로 나타났다. 사후검정(Dunnett's t3) 결과 구성동측정소와 성남동1측정소 그리고 읍내동측정소의 PM₁₀ 평균농도가 동일하게 나타났으며, 그 보다 낮은 농도로 문창동측정소가 월평동측정소로 나타났다. 가장 낮은 농도를 나타낸 곳은 정림동측정소와 둔산동측정소 이다.

2.2 배출량 특성

대전시 미세먼지 연도별 배출현황은 에어코리아에서 공개되어있는 2015년도까지의 현황을 분석하였다. 분석결과 2010년 552 ton에서 2014년까지 384 ton으로 감소하다가 2015년 1,912 ton으로 급격히 증가하는 것으로 나타났다(Fig. 2). 그러나 다른 년도와 비교하여 주요 배출원인인 비산먼지

Table 2. Average Concentration of PM₁₀ in Daejeon Area (2017)

No.	Measuring station	Address/location	N	Average concentration ($\mu\text{g}/\text{m}^3$)	Standard Deviation	F(p)	Dunnett's t3*
No.1	Eupnae-dong	75, Daejeon-ro 1331 beon-gil, Daedeok-gu	8,363	45.29	26.508	1207.809 (0.0***)	4
No.2	Munpyeong-dong	34, Munpyeongdong-ro 18 beon-gil, Daedeok-gu	8,568	52.70	29.695		1
No.3	Munchang-dong	38, Bomun-ro 20 beon-gil, Jung-gu	8,566	43.41	25.092		5
No.4	Guseong-dong	407, Daehak-ro, Yuseong-gu	8,479	45.67	27.538		4
No.5	Noeun-dong	89, Noeundong-ro 87 beon-gil, Yuseong-gu	8,589	47.26	27.944		3
No.6	Daheung-dong1	58, Juangang-ro, Jung-gu	8,457	51.63	28.097		2
No.7	Seongnam-dong1	70, Gyejog-ro 368 beon-gil, Dong-gu	8,440	45.47	22.613		4
No.8	Jeongnim-dong	10, Jeongnimdong-ro, Seo-gu	8,569	41.69	23.634		6
No.9	Dunsan-dong	84, Dunsanse-ro, Seo-gu	8,238	41.39	27.309		6
No.10	Woalpyeong-dong	160-5, Woalpyeong-dong, Seo-gu	8,554	42.96	25.602		5
total			84,823	45.76	26.734		

p<0.001

Fig. 1. Current Status of Air Pollution Measurement Stations in Daejeon City

Fig. 2. Emissions of Fine Dust by Year in Daejeon

나 제조업 연소 활동이 증가한 것은 아니다. 또한 이런 현상은 다른 도시에서도 유사하다. 따라서 국내의 산업 활동의 증가로 인해서 발생했다고 보기에는 어렵고 외부의 요인에 의해 증가한 것으로 판단할 수 있을 것이다. 외부 유입에 의한 미세먼지의 증가를 판단하기 위해서는 더 세부적인 분석이 필요하다.

3. 분석방법

3.1 분석방법

본 연구에서는 2017년 기준으로 대전광역시 내 위치한 관측소 중 미세먼지 농도가 가장 높은 것으로 나타난 문평동 측정소를 대상으로 외부 유입에 의한 미세먼지의 증가를 판단하기 위해 2013년부터 2017년까지 최근 5년간 계절에 따른 영향과 하루 24시간의 미세먼지 변동 특성을 분석하였다. 분석방법은 다변량 분산분석을 통해 정량화하기 어려운 내용을 신뢰성 있는 통계로 객관적 접근이 가능한 반복측정 분산분석을 이용하였다. 동일한 대기환경 속에서 미세먼지 농도의 변화를 비교분석하기 위하여, 귀무가설은 “문평동 관측소는 시기별, 연도별, 시간대별 미세먼지 농도에 변화가 없을 것이다”이며 대립가설은 “문평동 관측소는 시기별, 연도별, 시간대별 미세먼지 농도에 변화가 있을 것이다”로 설정하였다.

계절에 따른 영향을 분석하기 위하여 기상특성이 뚜렷하였던 특정시기를 선정하였으며, 연평균 최저 기온을 나타낸 1월과 최고 기온을 나타낸 8월 그리고 평균 풍속이 가장 높은 4월을 비교 월로 선정하였다(Table 3). 그리고 현재

에어코리아에서 제공하는 미세먼지 데이터는 1시간에 1회, 24시간을 측정된 데이터를 제공한다. 연구의 편의성을 위해 24시간을 4개의 시간대역으로 구분하였다. 0시부터 6시까지는 새벽 시간대, 6시부터 12시까지는 오전 시간대, 12시부터 18시까지는 오후 시간대, 18시부터 24시까지는 저녁 시간대로 구분하여 비교분석 하였다(Fig. 3).

모든 데이터는 고유의 ID를 가지고 있어야 하지만 본 연구에서 사용된 자료는 별도의 고유 ID를 가지고 있지 않았다. 년도와 분기로 나누어진 원본 데이터를 하나의 행렬로 묶기 위해서는 각 데이터 별로 별도의 고유 ID를 부여하는 작업을 먼저 선행해야 한다. 따라서 측정소 식별 번호와 시간 데이터를 활용하였다. 시간 데이터는 년도, 월, 일, 시간(분과 초는 제외)까지 포함하여 YYYYMMDDTT으로 작성되어 있었다. 이를 파이썬(Python)의 `split` 함수를 활용하여 년도, 월, 일 그리고 시간으로 나누었다. 이를 다시 측정소와 고유번호와 월, 일 그리고 시간이 일치하게 데이터를 정렬 시켰다. 따라서 년도는 다르지만 나머지 데이터는 서로 일치 되게 데이터가 정렬이 되게 된다. 또한 각 시간을 4개의 집단으로 나누었다.

3.2 반복측정 분산분석

반복측정 분산분석(Repeated Measured Analysis of Variance, RMANOVA)은 다변량 분산분석(Multivariate Analysis of Variance, MANOVA)을 확장한 개념으로 동일한 집단에 대하여 셋 이상의 조건에 노출이 되어 측정된 변수의 결과를 분석하거나 하나 이상의 집단이 동일한 조건에 반복적으로 노출되는 경우 시간에 따라 종속변수의 변화량을 측정하기 위하여 적용

Table 3. 1981~2010 Daejeon Area Weather Information Statistics (KMA data release portal)

division	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
average air temperature	-1.0	1.5	6.5	13.0	18.2	22.4	25.0	25.6	21.3	14.7	7.5	1.2
the highest temperature	4.0	7.0	12.6	19.2	24.1	27.5	29.0	29.8	26.4	20.9	13.3	6.4
lowest temperature	-5.4	-3.5	1.0	7.0	12.6	17.9	21.8	22.2	17.1	9.4	2.5	-3.4
rainfall	29.6	34.2	55.6	81.7	103.7	206.3	333.9	329.5	169.7	47.4	41.1	25.9
average wind velocity	1.7	1.8	2.2	2.3	2.2	1.9	2.1	1.9	1.8	1.4	1.5	1.5
average relatively humidity	64.8	59.5	56.5	55.3	62.1	68.7	77.7	77.5	74.4	70.4	66.9	66.4
Fog time continued	9.71	7.68	3.22	2.17	1.44	0.83	0.89	1.65	5.85	11.9	11.1	6.81
Ground-surface temperature	-0.4	2.1	7.7	15.3	21.6	26.1	27.5	28.2	23.9	16.3	7.6	1.2

Fig. 3. Procedures and Methods of Research

되는 기법이다. 반복측정은 몇 가지 처리(treatment)를 적용함에 있어 실험에서 충분한 수의 실험개체를 확보하기 어려울 경우나 실험개체 간의 차이가 커서 전체 결과에 영향을 주는 것을 방지하기 위해 실험개체를 블록(block)으로 처리된다. 그러나 한 개체를 반복적으로 측정함으로써 많은 시간이 소요되며, 처음의 처리가 이후 실험에 영향을 가져올 수 있는 단점이 있다. 동일한 대상을 여러 번 측정한 이와 같은 자료는 잠재성장모형이 적용되는 자료와 자료구조가 동일하며 의·약학 분야에서 치료 효과의 추이를 파악하거나 사회과학분야의 시간에 따른 변화 연구에서는 오히려 반복측정의 단점이 크게 문제가 되지 않는다. 잠재성장모형이 변화 추이를 가진 모형에 대한 전체 적합도에 치중한 분석이라면 반복측정 분산분석은 변화의 추이보다는 각 단계별로 차이여부를 중점적으로 해석한다(Lee and Kang, 2013).

또한 반복측정 분산분석에서는 구형성(sphericity) 가정이 중요하다. 구형성을 가정할 수 있으나 없느냐에 따라 분석 통계량이 달라지기 때문이다. 독립표본 t 검정과 마찬가지로 각 처리마다 변수의 분산이 동일하다는 등분산 검정과 각 처리 쌍 간에 동일한 상관관계가 존재한다는 복합대칭성 가정을 말한다. 그러나 실제로 구형성 가정을 만족하기는 매우 어려운 일이다. Girden (1992)에 의하면 반복측정 분산분석은 구형성에 대한 가정이나 구형성의 특별한 형태인 복합대칭성(compound symmetry)에 대한 가정이 만족되어야 한다. 그러나 최근에는 복합대칭성의 가정이 분석에서 충분조건이지 필요한 가정이 아니라고 하여 가정을 만족하지 못한 경우 자유도를 수정한 방법을 사용하는 것이 가능하다고 하였다.

본 연구에서는 모든 항목에서 구형성 가정을 만족하지 않는 것으로 나타났다. 구형성 가정이 만족되지 않는 경우 다변량 분석을 하게 된다. 다변량 통계량은 Pillai's Trace, Wilk's Lambda, Hotelling-Lawley Trace, Roy의 최대근(Maximum Root) 등을 통하여 1종 오류와 통계적 검증력을 확인한다. 다변량 분석에 사용되는 4가지 통계량은 $E^{-1}H$ 행렬의 고유값(λ_i)과 고유벡터와 관련이 있으며 4가지 통계량은 다음과 같다.

$$\text{Pillai's Trace: } \text{Trace}[H(H+E)^{-1}] = \sum_i \frac{\lambda_i}{1+\lambda_i} \quad (1)$$

$$\text{Wilk's Lambda: } \frac{E}{H+E} = H\left(\frac{1}{1+\lambda_i}\right) \quad (2)$$

$$\text{Hotelling-Lawley Trace: } \text{Trace}[E^{-1}H] = \sum_i \lambda_i \quad (3)$$

$$\text{Roy's Maximum Root: } \text{Max } \lambda_i \quad (4)$$

각 통계량들은 합리적인 이론적 근거를 가지고 있으므로 통계량 간에 우열을 비교할 수 없으며, 일반적으로 개체수가 클 때는 4가지 통계량들이 거의 동일한 결과를 나타낸다.

Pillai's Trace는 양의 값을 갖는 통계량이며, 이 값의 증가는 처리 효과가 모형에 미치는 영향이 큰 다는 것을 의미한다. 또한 Box's M test에서 동질성 가정을 위배하였을 때 이용된다. Wilks's Lambda는 0과 1사이의 값을 가지는 통계량이며, 이 통계량 값이 작다는 것은 처리 효과가 모형에 미치는 영향력이 더 크다는 것을 의미한다. Hotelling's Trace는 검정행렬의 고유값의 합으로 양의 값을 가진 통계량으로 이 값이 높으면 처리 효과가 모형에 미치는 영향이 크다는 것을 의미한다. Hotelling-Lawley Trace는 항상 Pillai's Trace보다 큰 값을 가지지만, 검정행렬의 고유값이 작은 경우 이 두 통계량은 거의 같아지며 이는 종속변수들 간의 강한 상관관계가 있다는 것을 암시한다. Roy's Maximum Root 검정행렬의 가장 큰 고유값을 의미하기 때문에 양의 값을 가진 통계량이며, 이 값이 증가하면 처리 효과가 모형에 미치는 영향이 더 크다는 것을 의미한다. 따라서 본 연구에서는 가장 강력한 통계적 검증력을 가지고 있는 것으로 알려진 Pillai's Trace, Wilk's Lambda 통계량만을 분석 하였다.

4. 분석결과

4.1 월별 PM₁₀ 농도 분석

동질한 대기환경과 산업활동에서는 미세먼지의 발생도 동일 할 것으로 예상하였지만 분석 결과 귀무가설을 기각하고 대립가설이 채택되는 것으로 나타났다(Tables 5에서 8). Airkorea의 측정자료를 이용하여 문평동 측정소의 2017년도 월별 PM₁₀ 농도를 비교한 결과, 3월이 66.889 $\mu\text{g}/\text{m}^3$ 으로 가장 높게 나타났으며, 11월이 65.605 $\mu\text{g}/\text{m}^3$, 5월이 64.99 $\mu\text{g}/\text{m}^3$, 4월이 64.36 $\mu\text{g}/\text{m}^3$, 1월이 63.548 $\mu\text{g}/\text{m}^3$ 로 PM₁₀ 농도가 60 $\mu\text{g}/\text{m}^3$ 이상으로 다른 시기에 비해 상대적으로 높게 나타났다. 그 다음으로는 12월이 59.492 $\mu\text{g}/\text{m}^3$, 2월이 56.031 $\mu\text{g}/\text{m}^3$, 9월이 45.451 $\mu\text{g}/\text{m}^3$, 6월이 44.896 $\mu\text{g}/\text{m}^3$, 10월이 44.172 $\mu\text{g}/\text{m}^3$, 7월이 32.297 $\mu\text{g}/\text{m}^3$, 8월이 23.580 $\mu\text{g}/\text{m}^3$ 순으로 나타났다. 이런 월별 차이는 통계적($F=216.512$, $p<0.001$)으로 유의한 차이를 나타냈다. 사후검정 결과(Dunnett's t3) 1월, 3월, 4월, 5월, 11월이 가장 높은 PM₁₀ 농도를 나타냈으며, 8월이 가장 낮은 PM₁₀ 농도를 나타냈다(Table 4). 여기서, N은 측정횟수 이다.

4.2 반복측정 분산분석

Table 5와 Fig. 4는 새벽시간(0~6시)의 미세먼지 발생 농도의 변화를 보여주고 있다. PM₁₀ 농도가 가장 높았던 시기는 2016년 4월(77.82 $\mu\text{g}/\text{m}^3$)이었으며, PM₁₀ 농도가 가장 낮았던 시기는 2014년 8월(14.58 $\mu\text{g}/\text{m}^3$)이었다. 1월에서는 2013년이 73.28 $\mu\text{g}/\text{m}^3$ 로 가장 높고, 4월에서는 2016년이 77.82 $\mu\text{g}/\text{m}^3$ 로 가장 높게 나타났다. 그리고 8월에서는 2013년이 38.25 $\mu\text{g}/\text{m}^3$ 로 가장 높게 나타났다. 오전시간(6~12시)에 PM₁₀ 농도가 가장 높았던 시기는 2016년 4월(80.78

Table 4. Comparison of Monthly PM₁₀ Concentration at Moonpyeong-dong Station (2017)

division	N	Average ($\mu\text{g}/\text{m}^3$)	Standard Deviation	F(p)	Dunnett's t3
Jan.	742	63.548	29.428	216.512 (0.0***)	2
Feb.	670	56.031	25.786		4
Mar.	742	66.889	22.615		1
Apr.	718	64.367	29.311		1
May	741	64.995	45.334		1
Jun.	712	44.896	16.847		5
Jul.	742	32.297	17.882		6
Aug.	708	23.580	13.959		7
Sept.	635	45.451	24.784		5
Oct.	741	44.172	21.966		5
Nov.	718	65.605	25.246		1
Dec.	696	59.492	27.528		3
total	8,565	52.707	29.700		

p<0.001

Table 5. Results of PM₁₀ Concentration Analysis at Dawn Time (0~6 AM)

division	Jan.			Apr.			Aug.		
	M ₁₀ ($\mu\text{g}/\text{m}^3$)	SD	N	M ₁₀ ($\mu\text{g}/\text{m}^3$)	SD	N	M ₁₀ ($\mu\text{g}/\text{m}^3$)	SD	N
2013	73.28	37.397	171	56.07	34.313	180	38.25	28.089	24
2014	68.70	32.484	171	48.96	23.535	180	14.58	3.078	24
2015	58.68	22.401	171	44.54	22.312	180	37.50	8.262	24
2016	52.31	23.443	171	77.82	38.048	180	32.50	3.514	24
2017	61.28	29.082	171	62.90	26.125	180	17.67	3.199	24
Statistical verification		Value	F	Hypothesis df		Error df	Significance level		
Factor	Pillai's Trace	0.050	4.836c	4.000		369.000	0.001***		
	Wilks' Lambda	0.950	4.836c	4.000		369.000	0.001***		
Factor group	Pillai's Trace	0.364	20.606	8.000		740.000	0.000***		
	Wilks' Lambda	0.654	21.782c	8.000		738.000	0.000***		

***p<0.001, **p<0.01, *p<0.05

Fig. 4. Change of PM₁₀ Concentration at Dawn Time (0~6 AM) **Fig. 5.** Change of PM₁₀ Concentration at Morning Time (6~12 AM)

Table 6. Results of PM₁₀ Concentration Analysis at Morning Time (6~12 AM)

division	Jan.			Apr.			Aug.		
	M ₁₀ (μg/m ³)	SD	N	M ₁₀ (μg/m ³)	SD	N	M ₁₀ (μg/m ³)	SD	N
2013	79.48	39.830	166	59.87	30.629	172	44.38	24.657	24
2014	72.28	27.111	166	53.30	25.490	172	13.83	4.697	24
2015	62.40	22.699	166	42.54	20.044	172	36.04	10.796	24
2016	63.11	28.296	166	80.78	53.432	172	34.00	7.907	24
2017	71.05	32.166	166	69.73	33.468	172	18.25	4.110	24
Statistical verification		Value	F	Hypothesis df		Error df	Significance level		
Factor	Pillai's Trace	0.097	9.512c	4.000		356.000	0.000***		
	Wilks' Lambda	0.903	9.512c	4.000		356.000	0.000***		
Factor group	Pillai's Trace	0.267	13.780	8.000		714.000	0.000***		
	Wilks' Lambda	0.747	13.945c	8.000		712.000	0.000***		

***p<0.001, **p<0.01, *p<0.05

μg/m³)이었으며, PM₁₀ 농도가 가장 낮았던 시기는 2014년 8월(13.83 μg/m³)이었다. 1월에서는 2013년이 79.48 μg/m³로 가장 높고, 4월에서는 2016년이 80.78 μg/m³로 가장 높게 나타났다. 그리고, 8월에서는 2013년이 44.38 μg/m³로 가장 높게 나타났다(Table 6과 Fig. 5).

Table 7과 Fig. 6은 오후시간(12~18시)의 미세먼지 발생 농도의 변화를 보여주고 있다. PM₁₀ 농도가 가장 높았던 시기는 2016년 4월(75.51 μg/m³)이었으며, PM₁₀ 농도가 가장 낮았던 시기는 2014년 8월(14.22 μg/m³)이었다. 1월에서는 2013년이 67.89 μg/m³로 가장 높고, 4월에서는 2016년이 75.51 μg/m³로 가장 높게 나타났다. 그리고, 8월에서는 2013년이 31.26 μg/m³로 가장 높게 나타났다. 저녁시간(18~24시)에 PM₁₀ 농도가 가장 높았던 시기는 2016년 4월(79.47 μg/m³)이었으며, PM₁₀ 농도가 가장 낮았던 시기는 2014년 8월

(14.22 μg/m³)이었다. 1월에서는 2013년이 70.44 μg/m³로 가장 높고, 4월에서는 2016년이 79.47 μg/m³로 가장 높게 나타났다. 그리고, 8월에서는 2015년이 32.0 μg/m³로 가장 높게 나타났다(Table 8과 Fig. 7).

시간대별 PM₁₀ 농도를 비교 결과, 1월의 경우 오전시간(6~12시)의 PM₁₀ 농도가 79.48 μg/m³로 가장 높게 나타났으며, 새벽시간(0~6시)이 73.28 μg/m³, 저녁시간(18~24시)이 70.44 μg/m³, 오후시간(12~18시)이 67.89 μg/m³순으로 나타났다. 4월의 경우 오전시간(6~12시)의 PM₁₀ 농도가 80.78 μg/m³로 가장 높게 나타났으며, 저녁시간(18~24시)이 79.47 μg/m³, 새벽시간(0~6시)이 77.82 μg/m³, 오후시간(12~18시)이 75.51 μg/m³순으로 나타났다. 8월의 경우에는 오전시간(6~12시)의 PM₁₀ 농도가 44.38 μg/m³로 가장 높게 나타났으며, 새벽시간(0~6시)이 38.25 μg/m³, 오후시간(12~18시)

Table 7. Results of PM₁₀ Concentration Analysis at Afternoon Time (12~18 PM)

division	Jan.			Apr.			Aug.		
	M ₁₀ (μg/m ³)	SD	N	M ₁₀ (μg/m ³)	SD	N	M ₁₀ (μg/m ³)	SD	N
2013	67.89	43.734	176	47.32	22.052	175	31.26	15.841	23
2014	66.71	33.298	176	49.94	20.888	175	15.78	4.889	23
2015	57.63	24.589	176	40.45	19.943	175	30.39	6.841	23
2016	54.78	24.197	176	75.51	48.763	175	30.91	5.736	23
2017	60.88	30.907	176	64.50	32.381	175	14.22	4.472	23
Statistical verification		Value	F	Hypothesis df		Error df	Significance level		
Factor	Pillai's Trace	0.039	3.715c	4.000		368.000	0.006**		
	Wilks' Lambda	0.961	3.715c	4.000		368.000	0.006**		
Factor group	Pillai's Trace	0.238	12.489	8.000		738.000	0.000***		
	Wilks' Lambda	0.769	12.892c	8.000		736.000	0.000***		

***p<0.001, **p<0.01, *p<0.05

Fig. 6. Change of PM₁₀ Concentration at Afternoon Time (12~18 PM)

Fig. 7. Change of PM₁₀ Concentration at Evening Time (18 PM~24 AM)

Table 8. Results of PM₁₀ Concentration Analysis at Evening Time (18 PM~24 AM)

division	Jan.			Apr.			Aug.		
	M ₁₀ (µg/m³)	SD	N	M ₁₀ (µg/m³)	SD	N	M ₁₀ (µg/m³)	SD	N
2013	70.44	39.233	179	48.18	31.317	177	30.00	20.196	18
2014	67.54	33.091	179	49.36	23.155	177	14.39	3.238	18
2015	58.37	27.233	179	45.87	23.477	177	32.00	8.080	18
2016	56.45	25.897	179	79.47	43.461	177	28.39	6.031	18
2017	59.84	26.608	179	60.72	24.899	177	16.72	3.801	18
Statistical verification		Value	F	Hypothesis df		Error df	Significance level		
Factor	Pillai's Trace	0.032	3.045c	4.000		368.000	0.017*		
	Wilks' Lambda	0.968	3.045c	4.000		368.000	0.017*		
Factor group	Pillai's Trace	0.266	14.169	8.000		738.000	0.000***		
	Wilks' Lambda	0.740	14.931c	8.000		736.000	0.000***		

***p<0.001, **p<0.01, *p<0.05

이 31.26 µg/m³, 저녁시간(18~24시)이 30.00 µg/m³순으로 나타났다. 따라서, 시간대별 PM₁₀ 농도는 오전시간(6~12시)이 전반적으로 가장 높은 것으로 나타났다. 다음으로 새벽시간(0~6시), 저녁시간(18~24시)이 높았고, 오후시간(12~18시)이 전반적으로 가장 낮은 PM₁₀ 농도를 나타냈다.

1월의 PM₁₀ 농도는 2013년부터 2016년까지 감소하다가 2017년도에 증가하는 것으로 나타났다. 농도 범위는 52.31~79.48 µg/m³으로 평균 64.16 µg/m³이었으며 평균 미세먼지 농도가 가장 높았다. 4월은 2013년부터 2015년까지 감소하다가 2016년에 급격하게 증가하고 다시 감소하는 것으로 나타났다. 농도 범위는 40.45~80.78 µg/m³으로 평균 57.89 µg/m³이었다. 8월은 2014년도에 최저 농도를 나타냈고 2015년도에 다시 증가하였다가 2016년에 감소하는 것으로 나타났다. 농도 범위는 13.83~44.38 µg/m³로 평균 26.55 µg/m³이었으며 가장 낮은 미세먼지 농도를 나타냈다. PM₁₀ 농도는 1월과 8월에 전체적으로 감소하는 경향을 나타냈지만 4월은

2016년에 80.78 µg/m³로 가장 높은 미세먼지 농도를 기록했다. 전체적으로 분석 요인간의 변화는 통계적으로 유의한 차이를 나타냈다. 또한 요인과 집단 간의 상호작용도 통계적으로 유의한 것으로 나타났다. 즉, 월별과 년도에 따라 시간대 미세먼지 농도의 변화는 다른 것으로 나타났다(Tables 5에서 8, Fig. 8).

이때, 측정횟수 N값이 각각 다른 이유를 Table 5의 1월의 N값 171인 것을 사례로 설명하면 다음과 같다. 1월은 31일까지 있어 1시부터 6시까지 6번 측정하여 총 1개월 동안 186번 측정하게 된다. 그러나, 센서의 기계적 이상(점검)과 전체 년도의 데이터 중 일부가 누락이 되어 본 연구에서는 171개의 관측 데이터만을 활용한 것이다. 8월 같은 경우 18~24개의 데이터를 가지고 분석하였다. 8월은 2013년도가 다른 월에 비해 표준편차가 매우 높은 것으로 나타났다. 하지만 그 편차 정도는 다른 월의 표준편차 값과 비교하면 높지 않다. 따라서 센서의 기계적 문제 보다는 특정 시간에 미세먼지 농도가 급격히 증가한다는 것을 의미한다.

Fig. 8. Change in PM₁₀ Concentration for Each Month

5. 결 론

본 연구에서는 대전시 3, 4공단을 중심으로 미세먼지의 계절적인 영향과 시간대별 영향을 고려하기 위하여 두 가지 방식으로 구분하여 분석하였다. 첫 번째는 2017년 기준으로 대전 지역 측정소 간의 미세먼지 농도 비교와 문평동 측정소를 기준으로 월별 미세먼지 농도를 비교하였다. 두 번째는 문평동 측정소 기준으로 2013년부터 2017년까지 미세먼지 농도를 시간대와 주요 월별로 비교하였다.

- (1) 대전 지역의 3, 4공단(문평동 측정소)이 다른 지역보다 PM₁₀ 농도가 높은 것으로 나타났다. 그 이유는 미세먼지의 주요 발생 원인이 제조업 연소이고 대전지역의 주요한 생산 공장이 3, 4공단에 위치하기 때문으로

판단된다. 문평동 측정소의 월별 PM₁₀ 농도는 3월이 가장 높았으며, 8월이 가장 낮은 것으로 나타났다. 봄철인 1, 3, 4, 5, 11월의 PM₁₀ 농도가 $60 \mu\text{g}/\text{m}^3$ 이상으로 다른 시기에 비해 상대적으로 높게 나타났다.

- (2) 반복측정 분산분석 결과, 시간대별 PM₁₀ 농도는 오전 시간(6~12시)이 전반적으로 가장 높은 것으로 나타났다. 다음으로 새벽시간(0~6시), 저녁시간(18~24시)이 높았고, 오후시간(12~18시)이 가장 낮은 PM₁₀ 농도를 나타냈다.
- (3) PM₁₀ 평균 농도는 $64.16 \mu\text{g}/\text{m}^3$ 로 1월이 가장 높았으며, 그 다음으로 4월이 $57.89 \mu\text{g}/\text{m}^3$, 8월이 $26.55 \mu\text{g}/\text{m}^3$ 로 낮았다. 연도별로 보면 모든 시간대에서 1월과 8월은 2017년까지 감소하지만 4월은 2015년까지 감소를 하다가 2016년에도 급격하게 증가했다가 2017년 다시

감소하는 것으로 나타났다.

- (4) 기존의 연구결과와 마찬가지로 대전시도 공단지역에 위치한 측정소가 다른 지역에 위치한 측정소 보다 미세먼지 농도가 높은 것으로 나타났다. 그러나 특정 년도에 급격히 증가하는 것은 공단지역이라는 요인으로 해명되어지지 않는다. 그 이유는 특정 년도에 대전 지역의 경제활동이 급격히 증가하지 않았기 때문이다. 결과적으로 외부의 영향 요인으로 특정 년도에 미세먼지가 증가했다고 판단되어진다. 따라서 외부의 영향 요인에 대해서 더 세밀한 연구가 필요하다.

동질한 기상환경과 산업활동에서는 미세먼지의 발생에도 동일 할 것으로 예상하였지만 각기 다른 것으로 나타났다. 따라서 미세먼지의 발생을 저감하기 위해서는 미세먼지 저감에 관한 기술 개발도 중요하지만 근본적으로 미세먼지의 발생 원인과 이동 경로 등에 대한 연구가 필요하다. 그러나, 현재의 연구 동향을 보면 미세먼지가 발생 후 저감하는 장비 및 기술 개발 연구에 집중되어 있다. 현재의 인공지능과 딥러닝, 머신러닝, 빅데이터 분석의 기술을 활용한다면 미세먼지의 발생 원인에 대한 심도있는 연구가 진행될 수 있을 것이다. 또한 미세먼지와 관련되는 산업 활동의 전반적인 데이터를 확장 및 연결하여 다양한 결과를 이끌어 낼 수 있을 것이며, 이런 미세먼지와 관련된 다양한 빅데이터 분석 연구가 필요한 시점이라 사료된다. 향후에는 초미세먼지를 포함하여 연도별, 시간대별로의 미세먼지 발생 패턴에 관한 연구가 필요할 것이다.

감사의 글

본 연구는 대전테크노파크의 지원(뉴비즈니스 모델 일 자리창출사업)과 행정안전부 국립재난안전연구원의 지원(재난원인 현장감식 기술개발, NDMI-주요-2019-05-01)에 의해 수행되었습니다.

References

Environmental Protection Agency. (2014). *Intergrated science assessment for particulate matter*. Retrieved from http://www.sinia.cl/1292/articles-51242_Integrated_Science.pdf

Girden, E.R. (1992). *ANOVA: Repeated measures*. Sage University paper series on quantitative applications in the social sciences, 07-084. Newbury Park, CA: SAGE.

Jang, Y.K. (2016). Status and problems of fine dust pollution.

Journal of Environmental Studies, Graduate School of Environmental Studies, Seoul National University, Vol. 58, pp. 4-13.

Kim, Y.S. (2013). *PM₁₀ and PM_{2.5} source identification study in industrial and commercial region*. Ph.D. dissertation, Dong-A University.

Kim, C.K., and Seo, Y.H. (2013). Chemical characteristics of west coast regional fine aerosol of Daechon city. *Journal of Korea Society of Environmental Administration*, Vol. 19, No. 1, pp. 21-28.

Korea Land and Geospatial Informatix Corporation (LX). (2016). *Analysis of the cause of fine dust in Jeollabuk-do using Big Data*. Jeollabukdo Institute of Health & Environment Research.

Korea Environment Corporation (K-eco). *AirKorea homepage*. Retrieved from www.airkorea.or.kr

Korea Meteorological Administration (KMA). *Data release portal*. Retrieved from <https://data.kma.go.kr>

Lee, S.B., Jung, D.S., Cho, E.K., Kim, H.A., Hwang, E.Y., and Kang, C.H. (2011). Composition and pollution characteristics of PM₁₀ and PM_{2.5} particles at Gosan site of Jeju Island in 2008. *Analytical Science & Technology*, The Korean Society of Analytical Science, Vol. 24, No. 4, pp. 310-318.

Lee, H.J., and Kang, S.B. (2013). Analysis of latent growth model using repeated measures ANOVA in the data from KYPS. *Journal of the Korean Data & Information Science Society*, Vol. 24, No. 6, pp. 1409-1419.

National Institute of Environmental Research (NIER). *National air pollutants emission service*. Retrieved from <http://airemiss.nier.go.kr>

Park, S., and Shin, H. (2017). Analysis of the factors influencing PM_{2.5} in Korea: Focusing on seasonal factors. *Journal of Environmental Policy and Administration*, Vol. 25, No. 1, pp. 227-248.

Zhou, X. (2014). *Statistical characteristics of meteorological factors and particulate matters (PM₁₀, PM_{2.5}) in Seoul*. Master's thesis, Seoul National University of Science and Technology.

Received	April 11, 2019
Revised	April 16, 2019
Accepted	May 31, 2019