

공간패널모형을 이용한 국내 초미세먼지 농도에 대한 분석[†]

이종현¹ · 김영민² · 김용구³

¹²³ 경북대학교 통계학과

접수 2017년 3월 2일, 수정 2017년 3월 27일, 게재확정 2017년 3월 31일

요약

초미세먼지 (particulate matter 2.5, PM_{2.5})는 분진의 입경이 2.5 이하의 보다 작은 크기의 미세한 입자들을 말하는데, 미세먼지와 달리 대기 중에서 제거가 어렵고 기도나 코 점막에서 걸러지지 않으며, 호흡 시 폐포까지 직접 침투하기 때문에 장기간 노출될 경우, 폐 기능 감소, COPD (chronic obstructive pulmonary disease) 증가, 폐암 발생증가가 있다고 알려져 있다. 현재 국내외에서 초미세먼지에 대해 다양한 연구가 이루어지고 있는데, 초미세먼지의 농도는 기상인자 (풍속, 강우량, 일사량 등)에 영향을 받는 것으로 알려져 있으며, 이산화질소, 오존, 이산화황, 미세먼지 등 대기물질의 농도에도 영향을 받는 것으로 알려져 있다. 특히 우리나라는 점차 증가하고 있는 자동차 수나 오염원으로 인한 초미세먼지 외에도 중국으로부터 유입되는 초미세먼지 또한 고려되어야 하는 대상이므로 기상인자 중 풍향과 풍속 또한 어느 정도 큰 영향을 미칠 것으로 판단되며 인접 지역에 대한 영향 또한 고려되어야 할 것이다. 본 연구에서는 초미세먼지 농도에 영향을 미치는 유의한 대기물질 및 기상자료와 초미세먼지 농도의 지역적 특성을 고려한 공간자기상관 행렬에 기초한 공간패널모형을 소개하였고 이를 서울 25개 구에서 관측된 초미세먼지 자료에 적용하였다. 또한 초미세먼지와 대기오염물질의 농도를 통해 서울시에서 발생한 호흡기 질환 환자 수를 분석하여 그의 위해성을 확인하였다.

주요용어: 공간패널모형, 기상인자, 대기물질, 위해도분석, 초미세먼지.

1. 서론

2016년 9월 27일 세계보건기구 (WHO)의 보고에 의하면, 전 세계 92%의 인구거주환경의 공기질량이 세계보건기구의 기준을 넘어서며 매년 대기오염으로 인한 질병사망이 600만 명을 초과하는 것으로 예측된다. 더불어 우리나라의 초미세먼지 오염이 OECD 34개국 중 터키에 이어 두 번째로 심한 것으로 보고되었다. 또한 대기오염은 경제적 손실 역시 초래한다. OECD는 대기오염으로 인해 노동생산성 저하, 의료비 지출, 농업생산 감소 등으로 전 세계적으로 경제적 손실이 0.3%에 이르고, OECD 회원국 중 우리나라가 가장 큰 손실을 입을 것이라 예상하였다. 최근 우리나라 수도권 지역의 미세먼지 (PM₁₀) 농도는 개선 추세인 반면, 초미세먼지 (PM_{2.5}) 농도는 증가 추세로 전국적으로도 2015년부터 시행된 연간 대기환경기준 (25 $\mu\text{g}/\text{m}^3$)을 대부분의 지역에서 초과하고 있다 (Hwang, 2015).

초미세먼지는 분진의 입경이 2.5 이하의 보다 작은 크기의 미세한 입자들을 말하는데, 미세먼지와 달리 대기 중에서 제거가 어렵고 기도나 코 점막에서 걸러지지 않으며, 호흡 시 폐포까지 직접 침투하기

[†] 본 논문은 제 1저자인 이종현의 석사학위 논문 중 일부를 정리한 것임.

¹ (41566) 대구광역시 북구 대학로 80 경북대학교 통계학과, 석사과정.

² (41566) 대구광역시 북구 대학로 80 경북대학교 통계학과, 조교수.

³ 교신저자: (41566) 대구광역시 북구 대학로 80 경북대학교 통계학과, 부교수.

E-mail: kim.1252@knu.ac.kr

때문에 장기간 노출될 경우, 폐 기능 감소, COPD (chronic obstructive pulmonary disease) 증가, 폐암 발생증가가 있다고 알려져 있다 (Jung, 2014; Kim, 2014). 현재 국내외에서 초미세먼지에 대해 다양한 연구가 이루어지고 있는데 (Liu 등, 2012; Karppinen 등, 2004; Lee, 2010), 초미세먼지의 농도는 풍속, 강우량, 일사량 등 기상인자에 영향을 받는 것으로 알려져 있으며 (Cho, 2014), 이산화질소 (NO_2), 오존 (O_3), 이산화황 (SO_2), 미세먼지 등 대기물질의 농도에도 영향을 받는 것으로 알려져 있다. 또한 중국에서 한국으로 이동되는 황산화물의 양은 매년 94,000톤을 넘고 있으며, 이는 한국의 총 황산화물 양의 20%에 해당하는 양이며, 중국의 사막화가 가속화되면서 이로 인한 우리나라의 미세먼지 양은 매년 증가할 것으로 예상되고 있다 (Jung, 2009). Kang 등 (2004a)은 서울지역과 충남, 아산지역에서 봄철 중국의 황사 유입 전, 중, 후로 시기를 구분하여 많은 양의 황사먼지가 중국에서 한반도로 이동함을 연구하였고, 측정지역의 농도에 크게 기여하였음을 밝힌바 있다. 따라서 우리나라는 점차 증가하고 있는 자동차 수나 오염원으로 인한 초미세먼지 외에도 중국으로부터 유입되는 초미세먼지 또한 고려되어야 하는 대상이므로 기상인자 중 풍향과 풍속 또한 어느 정도 큰 영향을 미칠 것으로 판단되며 인접 지역에 대한 영향 또한 고려되어야 할 것이다. 그리고 서울시 초미세먼지 농도가 외국의 주요 대도시에 비해서 매우 높은 수준으로 나타난 것으로 알려져 있기 때문에 (Kang 등, 2004b) 우리나라의 초미세먼지 농도에 따른 호흡기 및 심혈관계 질환으로 인한 사망률이 매우 높을 것으로 예상된다. 따라서 초미세먼지 농도와 그에 따른 호흡기 환자 수를 예측하는 연구 역시 병행되어야 한다고 예상되어진다 (Yang 등, 2015).

대기오염물질의 관리가 배출기준, 대기환경기준을 강화함으로써 되고 있음에도 불구하고 수도권 및 대도시 지역을 중심으로 자동차 대수 증가, 기온상승 등의 이유로 선진국형 오염물질으로 알려져 있는 초미세먼지, 이산화질소 (NO_2), 오존 (O_3), 이산화황 (SO_2) 등의 오염도는 기대만큼 감소되지 않고 있어 우리나라의 대기질은 아직 후진국 또는 개발도상국의 대기질과 유사한 것이 현실이다 (Lee, 2015; Kwon과 Kim, 2016). 국내에서도 초미세먼지에 대한 다양한 연구가 진행되고 있지만 대기물질과 기상자료, 지역적 특성을 함께 고려한 연구는 부족한 실정이다. 따라서 본 연구에서는 초미세먼지 농도에 영향을 미치는 대기물질과 기상자료를 통해 일반화선형모형을 통해 유의한 인자들을 확인하고 초미세먼지 농도의 지역적 특성을 설명하기 위해 서울 25개 구에 기초한 공간자기상관 행렬을 고려하였다 (Na와 Kim, 2016). 또한 초미세먼지와 대기오염물질의 농도를 통해 서울시에서 발생한 호흡기 질환 환자 수를 분석하여 그의 위해성을 확인하려 한다.

2. 관측 자료

$PM_{2.5}$ 농도에 영향을 주는 변수로 먼저 기상인자를 고려하였다. 기상인자 자료는 2015년 8월 24일 ~ 2016년 8월 23일, 총 366일 중 결측 6일을 제외한 360일의 서울시 25개 구 (마포구, 용산구, 서대문구 등) 자료이며, 이 중 풍향은 더미변수로 $0^\circ \sim 90^\circ$ 는 0 그리고 $90^\circ \sim 180^\circ$ 는 1로 고려하였다. 이 자료는 서울시 열린데이터 광장 및 기상청 전자민원에서 제공한 것이며, 자세한 변수는 Table 2.1에 정리하였다. 그리고 다른 설명변수로서 대기오염물질을 고려하였는데, 기상인자와 같은 시간대, 장소의 서울시 열린데이터 광장에서 제공한 자료를 이용하였다. 자세한 변수는 Table 2.2에 정리하였다.

추가로 $PM_{2.5}$ 농도가 건강에 미치는 영향을 알아보기 위해서 서울지역 호흡기 환자수를 2011년 1월 1일 ~ 2015년 12월 31일까지 월별 자료를 조사하였으며 사용된 상병코드는 J43, J430, J431, J432, J438, J439, J44, J440, J441, J448, J449 이다. 이 자료는 국민건강보험공단에서 제공한 것이며, 자세한 변수는 Table 2.3에 정리하였다.

Table 2.1 Data description of weather factor

Variable	type	Description
mean.temp	Cont.	daily mean temperature (°C)
mn.temp	Cont.	daily minimum temperature (°C)
mx.temp	Cont.	daily maximum temperature (°C)
mean.humid	Cont.	daily mean humidity (%)
mn.humid	Cont.	daily minimum humidity (%)
mx.humid	Cont.	daily maximum humidity (%)
mean.wdspeed	Cont.	daily mean wind speed (m/s)
max.wdspeed	Cont.	daily maximum wind speed (m/s)
percips	Cont.	daily precipitation (mm)
wd.dir	Disc.	wind direction (0 : Northeast, 1 : Southeast, 2 : Southwest, 3 : Northwest)

Table 2.2 Data description of air pollutants

Variable	type	Description
PM _{2.5}	Cont.	daily mean PM _{2.5} concentration (µg/m ³)
PM ₁₀	Cont.	daily mean PM ₁₀ concentration (µg/m ³)
NO ₂	Cont.	daily mean NO ₂ concentration (ppm)
O ₃	Cont.	daily mean O ₃ concentration (ppm)
CO	Cont.	daily mean CO concentration (ppm)
SO ₂	Cont.	daily mean SO ₂ concentration (ppm)

3. PM_{2.5} 농도에 대한 통계모형

일반적으로 PM_{2.5}의 농도에 대하여 감마분포를 가정할 수 있는데, 다양한 기상인자 및 대기오염물질과 PM_{2.5} 농도와의 관련성을 알아보기 위해서 다음과 같은 일반화 선형모형을 고려할 수 있다.

$$E(Y_t^c) = \mu_t^c, \quad Var(Y_t^c) = \phi^c \mu_t^c, \tag{3.1}$$

$$\log \mu_t^c = \mu^c + \gamma_1^c C_{m_t} + \gamma_2^c S_{m_t} + \mathbf{x}_t^{c'} \boldsymbol{\beta}_1^c + \mathbf{z}_t^{c'} \boldsymbol{\beta}_2^c. \tag{3.2}$$

여기에서

- Y_t^c 는 지역 c 에서 t 일($t = 1, \dots, 360$)에 측정된 PM_{2.5}의 농도
- μ_t^c 는 c 지역 t 일의 PM_{2.5}의 평균 농도
- $C_{m_t} = \cos(2\pi(m_t/12))$, $S_{m_t} = \sin(2\pi(m_t/12))$. 여기에서, m_t ($m_t = 1, \dots, 12$)는 t 일의 월
- \mathbf{x}_t^c 는 c 지역, t 일의 PM_{2.5} 농도에 영향을 미치는 대기오염물질 (NO₂, O₃, CO, SO₂, PM₁₀)
- \mathbf{z}_t^c 는 c 지역, t 일의 기상변수 (평균기온, 최저기온, 최고기온, 평균습도, 최저습도, 최고습도, 평균풍속, 최대풍속, 강수량, 풍향)

참고로 위의 모형에서 일반적인 기상인자 이외에 중국의 황사먼지의 영향을 알아보기 위해서 풍향을 설명변수로 추가하였다.

관측된 각 지역별 PM_{2.5}의 농도는 시계열 자료와 지역별 자료의 정보를 가지므로 시계열 분석과 횡단면 분석을 동시에 수행할 수 있는 패널모형을 고려할 수 있는데, 패널분석은 개별적 특이성을 통제하면

Table 2.3 Data description of patients

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
8704	10450	10940	10970	11560	12870

서 다양한 정보를 제공해주면서 다중공선성의 문제를 줄일 수 있고, 시계열이나 횡단면 자료에서 포착하기 힘든 효과를 잘 보여준다. 관측된 $PM_{2.5}$ 의 농도에 대한 패널모형은 다음과 같이 표현할 수 있다.

$$\log Y_t^c = \beta_0 + \beta_1 C_{m_t} + \beta_2 S_{m_t} + \mathbf{x}_t^{c'} \boldsymbol{\beta}_3 + \mathbf{z}_t^{c'} \boldsymbol{\beta}_4 + \epsilon_t^c \quad (3.3)$$

$$\epsilon_t^c = \mu^c + \lambda_t + \nu_t^c. \quad (3.4)$$

여기에서

- ϵ_t^c 는 Y_t^c 에 대한 기본 선형모형의 오차항
- μ^c 는 관측되지 않은 c 지역 특성 효과
- λ_t 는 관측되지 않은 시간 효과
- ν_t^c 는 순수 오차항

일반적으로 패널모형은 오차항의 형태에 따라 Random effect 모형과 Fixed effect 모형으로 구분할 수 있는데 하우스만 검정 (Hausman specification test)을 통해서 어떤 모형을 사용할 지를 결정한다. 참고로 시간불변의 개별특성효과와 독립변수들간의 관련성에 대한 확실한 정보가 없는 경우에는 Fixed effect 모형을 더 선호하는 경향이 있다.

마지막으로 각 지역의 $PM_{2.5}$ 의 농도가 주변지역에 어떠한 영향을 미치는 지를 알아보기 위해서 기존의 패널모형에 공간적인 효과를 추가하여 다음과 같은 공간패널모형으로 표현할 수 있다.

$$\mathbf{y} = \lambda(\mathbf{I} \otimes \mathbf{W})\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad (3.5)$$

- \mathbf{y} 는 $NT \times 1$ 크기의 2015년 8월 24일부터 2016년 8월 23일까지 일 평균 $PM_{2.5}$ 농도.
- \mathbf{X} 는 대기오염물질, 기상인자들로 구성된 $NT \times k$ 크기의 독립변수 행렬.
- \mathbf{I} 는 T 크기의 단위행렬.
- \mathbf{W} 는 $N \times N$ 크기의 공간 구조 행렬.
- λ 는 관측값의 자기상관을 설명하는 공간모수.

참고로 본 연구에서 사용된 패널자료는 2015년 8월 24일부터 2016년 8월 23일까지 25개 구별 자료이며 $T = 360$ 이고 $N = 25$ 이다. 오차항인 \mathbf{u} 에 대하여 다음과 같은 자기상관 공간모형을 가정할 수 있다.

$$\mathbf{u} = \rho(\mathbf{I} \otimes \mathbf{W})\mathbf{u} + \mathbf{e}, \quad (3.6)$$

- ρ ($|\rho| < 1$)는 오차항에 대한 자기상관 공간 모수
- $\mathbf{e} \sim MVN(\mathbf{0}, \sigma^2 \mathbf{I})$.

지역별 인접성이 $PM_{2.5}$ 농도에 어떠한 영향을 미쳤는지를 알아보기 위해서 공간 구조 행렬 \mathbf{W} 는 인접지역으로 연결되어 있는지 유무에 따라 공간 자기상관이 존재하도록 정의할 수 있다. 따라서 공간 자기상관 모수인 λ 와 ρ 은 모형에서 인접한 지역들의 영향을 설명한다. 일반적으로 λ 가 유의한 모형을 spatial lag model (SLM)이라고 하고, ρ 가 유의한 모형을 spatial error model (SEM)이라고 하며, λ 와 ρ 가 모두 유의한 모형을 spatial durbin model (SDM)이라고 한다 (Giovanni와 Gianfranco, 2012; Jushan과 Kungpeng, 2013).

Table 4.1 Estimated coefficients and corresponding standard errors of GLM model

	Estimate	Std. Error	<i>t</i> value	<i>Pr</i> (> <i>t</i>)	
(Intercept)	1.419	0.040	35.825	< 2e-16	***
<i>SO</i> ₂	50.470	3.437	14.685	< 2e-16	***
<i>NO</i> ₂	20.569	0.434	47.344	< 2e-16	***
<i>O</i> ₃	11.806	0.510	23.165	< 2e-16	***
<i>mx.humid</i>	0.009	0.000	27.130	< 2e-16	***
<i>max.wdspeed</i>	-0.027	0.005	-5.225	1.78E-07	***
<i>mean.wdspeed</i>	-0.023	0.010	-2.252	0.024	*
<i>precips</i>	-0.008	0.000	-17.490	< 2e-16	***
<i>wd.dir1</i>	0.004	0.015	0.237	0.813	
<i>wd.dir2</i>	0.091	0.013	7.237	4.98E-13	***
<i>wd.dir3</i>	0.099	0.0121	8.183	3.14E-16	***
<i>m.sin</i>	0.132	0.007	19.347	< 2e-16	***
<i>m.cos</i>	-0.004	0.009	-0.493	0.622	

4. 분석결과

일반화 선형모형을 통해서 지역과 시간에 상관없이 기상인자와 대기오염물질이 PM_{2.5} 농도에 어떠한 영향을 주는지를 알아보았다. 분석 결과, 아황산가스, 이산화질소, 오존, 최고습도, 최고풍속, 평균풍속, 강수량, 풍향, 계절성이 유의하게 나타났고 다중공선성은 나타나지 않았다 (Table 4.1).

일반화 선형모형에서 유의한 변수들이 패널모형에서도 유의한 변수인지 확인하고 이 변수들에 의해서 설명되지 못하는 공간적 특성이 존재하는지 알아보았다. 패널 분석에 앞서 단위근 검정을 통해 정상성 (stationarity)을 확인하기 위해 Im-Pesaran-Shin (IPS) 검정의 결과, 귀무가설을 기각 ($\chi^2=-47.098$, p -value < 2.2e-16)하여 정상적인 시계열분포를 보이고 있는 것을 확인하였다. 또한 강건성 (robustness) 검증을 위해 설명변수 중 아황산가스와 이산화질소를 먼저 모형에 추가하고 그 다음에 오존, 최고습도, 최고풍속, 평균풍속, 강수량, 풍향, 계절성 순으로 순차적으로 모형에 추가하였는데, 모형의 유의성과 부호의 방향이 동일한 것을 확인할 수 있었다. 또한 하우스만 검정 (Hausman specification test)을 통해서 Random effect 모형과 Fixed effect 모형 중 어느 모형이 더 적절한지 확인해본 결과 귀무가설을 기각 ($\chi^2=128.72$, p -value < 2.2e-16)하여 Fixed effect 모형이 더 적절한 것을 확인하였다. 그 결과를 Table 4.2에서 정리하였다.

Table 4.2 Test on the fixed effect and random effect panel model

Parameter	Pooled OLS		Random Effect		Fixed Effect	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
(Intercept)	1.3728	0.0402	1.2772	0.0449		
<i>SO</i> ₂	51.3948***	3.4902	69.4845***	3.9594	71.1322***	4.0036
<i>NO</i> ₂	21.7312***	0.4412	20.1381***	0.4614	19.9379***	0.4639
<i>O</i> ₃	12.1760***	0.5176	14.3179***	0.5222	14.5039***	0.5232
<i>mx.humid</i>	0.0075***	0.0003	0.0090***	0.0003	0.0091***	0.0003
<i>max.wdspeed</i>	-0.0262***	0.0052	-0.0364***	0.0052	-0.0374***	0.0052
<i>mean.wdspeed</i>	-0.0259*	0.0104	-0.0650***	0.0110	-0.0697***	0.0110
<i>precips</i>	-0.0083***	0.0005	-0.0076***	0.0004	-0.0076***	0.0004
<i>wd.dir1</i>	0.0206	0.0156	0.0361*	0.0154	0.0377*	0.0153
<i>wd.dir2</i>	0.1124***	0.0128	0.0924***	0.0125	0.0908***	0.0125
<i>wd.dir3</i>	0.1207***	0.0123	0.1342***	0.0120	0.1354***	0.0120
<i>m.sin</i>	0.1381***	0.0070	0.1484***	0.0069	0.1497***	0.0069
<i>m.cos</i>	-0.0182*	0.0090	0.0079	0.0090	0.0104	0.0090

서울시 구별 인접성이 PM_{2.5} 농도에 어떠한 영향을 미쳤는지를 알아보기 위해서 공간 구조 행렬

W는 구끼리 연결되어 있는지 유무에 따라 공간 자기상관이 존재하도록 정의하였다. 이를 위해서 서울시 25개 구 (1.종로구, 2.중구, 3.용산구, 4.성동구, 5.광진구, 6.동대문구, 7.중랑구, 8.성북구, 9.강북구, 10.도봉구, 11.노원구, 12.은평구, 13.서대문구, 14.마포구, 15.양천구, 16.강서구, 17.구로구, 18.금천구, 19.영등포구, 20.동작구, 21.관악구, 22.서초구, 23.강남구, 24.송파구, 25.강동구)에 대해 인접한 구는 1, 인접하지 않은 구는 0으로 정의하였다 (Figure 4.1).

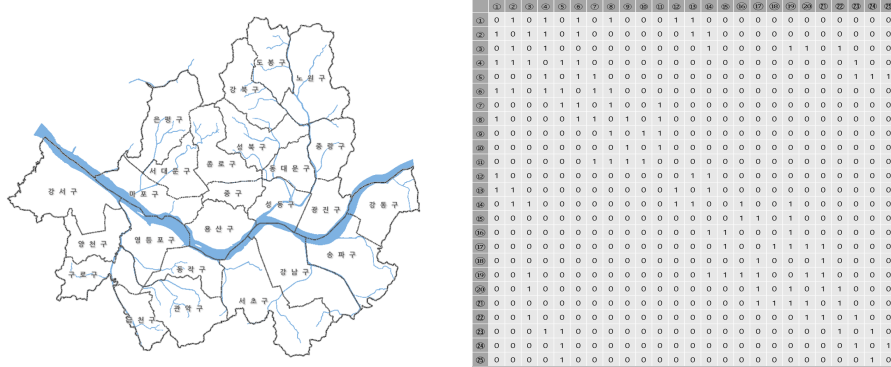


Figure 4.1 Locations and corresponding proximity matrix of the sampling sites

분석 결과를 통해서 아황산가스, 이산화질소, 오존, 최고습도, 최고풍속, 평균풍속, 평균 강수량, 평균 풍속, 계절성 등 일반화선형모형에서 유의했던 변수 모두 공간패널모형에서도 유의함을 알 수 있었다. 서울시 구별 인접성이 지역별 PM_{2.5}에 어떠한 영향을 보여주는 두 공간 자기상관 중 오차항에 대한 공간 자기상관 모수인 ρ 와 λ 모두 유의하게 나타났다. 또한 공간 패널 모형에 대한 Spatial Hausman 검정을 통해서 SEM 모형과 SDM 모형이 모두 유의함 (SEM 모형에 대한 χ^2 값은 99.332이고 p-값은 7.524e-16이며 SDM 모형에 대한 χ^2 값은 459.99이고 p-값은 2.2e-16)을 알 수 있었다. 두 모형 중 SDM 모형의 logLikelihood가 SEM 모형의 logLikelihood보다 크므로 (SDM : 333.1018, SEM : 318.2397) SDM 모형을 최종 모형으로 선택하였다. 참고로 Akaike information criterion (AIC)하에서도 유사한 결과를 얻을 수 있었다. Table 4.3에서 SDM 모형 분석 결과를 정리하였다. 참고로 분석을 위해 R 프로그램과 splm 패키지를 사용하였다.

Table 4.3 Estimated coefficients and corresponding standard errors of SDM model

	(Intercept)	SO ₂	NO ₂	O ₃	max. humid	max. wdspeed	mean. wdspeed	precips
Est.	1.900	6.526	11.922	5.403	0.006	-0.014	-0.042	-0.007
S.E.	0.045	2.538	0.384	0.458	0.000	0.003	0.007	0.000
T	42.031	2.572	31.060	11.808	18.722	-4.077	-6.311	-13.668
	wd.dir1	wd.dir2	wd.dir3	m.sin	m.cos	λ	ρ	
Est.	0.018	0.020	0.028	0.208	-0.090	0.022	0.164	
S.E.	0.008	0.008	0.007	0.009	0.010	0.006	0.002	
T	2.109	2.581	3.765	22.902	-8.615	3.810	76.752	

PM_{2.5} 농도가 건강에 미치는 영향을 알아보기 위해서 서울지역 호흡기 환자수에 대하여 다음과 같은 포아송 회귀모형을 고려할 수 있다.

$$E(Y_t) = \mu_t, \quad Var(Y_t) = \phi\mu_t, \tag{4.1}$$

$$\log\mu_t = \mu + \gamma_1 C_t + \gamma_2 S_t + \mathbf{x}'_t \beta_1. \tag{4.2}$$

여기에서

- Y_t 는 t 월($t = 1, \dots, 60$)에 집계된 총 호흡기 환자 수.
- μ_t 는 t 월의 집계된 총 호흡기 환자 수.
- $C_t = \cos(2\pi(t/12))$, $S_t = \sin(2\pi(t/12))$.
- \mathbf{x}_t 는 t 월의 호흡기 환자 수에 영향을 미치는 대기오염물질 (NO_2 , O_3 , CO , SO_2 , PM_{10} , $PM_{2.5}$).

위의 모형 적합결과, 호흡기 환자수가 계절성을 가지며 PM_{2.5} 농도와도 유의한 관련성이 있음을 알 수 있다 (Table 4.4).

Table 4.4 Estimated coefficients and corresponding standard errors of GLM model

	Estimate	Std. Error	t value	$Pr(> t)$	
(Intercept)	9.290	0.008	1110.313	< 2e-16	***
PM _{2.5}	0.001	0.000	2.916	< 0.004	**
$m.sin$	0.056	0.003	20.721	< 2e-16	***
$m.cos$	0.023	0.002	11.456	< 2e-16	***

5. 결론

본 논문에서는 서울시 25개 구 간의 인접성이 PM_{2.5} 농도에 어떠한 영향을 미치는지 알아본 결과, 일반화선형모형을 통해 선택된 변수들 모두 공간 패널 모형에서 유의하였으며 지역특수효과 추정계수로부터 PM_{2.5} 농도의 변화가 설명되었고, 25개 구 별 PM_{2.5} 농도가 서로간의 유의한 영향을 주는 것을 확인하였다. 유의한 변수들을 살펴보면, 풍속과 강수량은 PM_{2.5} 농도에 음의 영향을 가지는데 이는 풍속이 빠를수록, 강수량이 많을수록 대기 중 PM_{2.5}가 빠르게 제거됨을 알 수 있었다. 또한 풍향이 유의하게 나타났는데 이는 사막화가 가속화되고 있는 중국의 영향이 가장 컸고 (북서풍), 국내 PM_{2.5} 농도의 약 33%를 차지하고 있는 질소산화물과 황산화물의 주요 원인으로 나타나고 있는 자동차의 수가 전국에서 경기지역이 가장 많은 것으로 나타나 설명 가능하였다. 그리고 계절성의 유의성 역시 확인할 수 있었고, 각 구 별 인접성이 유의한 것으로 나타나 주변 지역의 PM_{2.5} 농도에도 영향을 받는 것을 알 수 있었다. 그리고 서울시 호흡기 환자 수와 대기오염물질간의 유의성을 일반화선형모형으로 확인해본 결과 PM_{2.5}와 계절성이 유의한 것으로 나타나 호흡기 환자수를 파악하는 것을 위해서도 PM_{2.5}에 대한 예측이 필요한 것으로 알 수 있었다.

References

- Cho, W. S. (2014) *Statistical characteristics of meteorological factors and particulate matters (PM₁₀, PM_{2.5}) in Seoul*, Master Thesis, Seoul National University, Seoul.
- Giovanni, M., Gianfranco, P. (2012) splm: Spatial panel data models in R. *Journal of Statistical Software*, **47**, 1-38.
- Hwang, S. H. (2015) *The effect of concentration of ultrafine particles (PM_{2.5}) in urban area on respiratory and circulatory diseases*, Master Thesis, Yonsei University, Seoul.
- Jung, J. H. (2009) *Study on the characteristics of PM_{2.5} and its major constituents collected in Chuncheon*, Master Thesis, Kangwon National University, Chuncheon.
- Jushan, B. and Kungpeng, L. (2013) Spatial panel data models with common shocks. *MPRA Paper*, University Library of Munich, Germany.
- Kang, C. M., Lee, H. S., Kang, B. W., Lee, S. K. and Sunwoo, Y. (2004a) Chemical characteristics of acidic gas pollutants and PM_{2.5} species during hazy episodes in Seoul, South Korea. *Atmospheric Environments*, **38**, 4749-4760.

- Kang, C. M., Sunwoo, Y., Lee, H. S., Kang, B. W. and Lee, S. K. (2004b) Atmospheric concentrations of $PM_{2.5}$ trace elements in the Seoul urban area of South Korea. *Journal of the Air & Waste Management Association*, **54**, 432-439.
- Karppinen, A., Harkonen, J., Kukkonen, J., Aarnio, P. and Koskentalo, T. (2004) Statistical model for assessing the portion of fine particulate matter transported regionally and long range to urban air. *Scandinavian Journal of Work, Environment & Health*, **30**, 47-53.
- Kim, C. W. (2014) *Development of guideline for the prevention and management of particulate matter/Asian dust-induced health damage (Allergic diseases)*, Technical report, Centers for Disease Control & Prevention.
- Kwon, H. J. and Kim, Y. (2016) A statistical prediction for concentrations of Manganese in the ambient air. *Journal of the Korean Data & Information Science Society*, **27**, 577-586.
- Lee, H. J. (2010) Analysis of time series models for PM_{10} concentrations at the Suwon city in Korea. *Journal of the Korean Data & Information Science Society*, **21**, 1117-1124.
- Lee, K. B. (2015) *Estimation of unknown components in PM_{10} and $PM_{2.5}$ in Seoul and evaluation of contribution of pollution source*, Ph.D Thesis, University of Seoul, Seoul.
- Liu, Y., He, K., Li, S., Wang, Z., Christiani, D. C. and Koutrakis, P. (2012) A statistical model to evaluate the effectiveness of $PM_{2.5}$ emissions control during the Beijing 2008 Olympic Games. *Environment International*, **44**, 100-105.
- Na, Y. and Kim, Y. (2016) A spatial panel regression model for household final consumption expenditure based on KTX effects. *Journal of the Korean Data & Information Science Society*, **27**, 1147-1154.
- Yang, H. j., Kim, S. H., Jang, A. S., Kim, S. H., Song, W. J., Kim, T. B., Ye, Y. M., Yoo, Y., Yu, J., Yoon, J. S. and Jee, H. M. (2015) Guideline for the prevention and management of particulate matter/yellow dust-induced adverse health effects on the patients with bronchial asthma. *Journal of the Korean Medical Association*, **58**, 1034-1043.

Spatial panel analysis for PM_{2.5} concentrations in Korea[†]

Jong Hyun Lee¹ · Young Min Kim² · Yongku Kim³

¹²³Department of Statistics, Kyungpook National University

Received 2 March 2017, revised 27 March 2017, accepted 31 March 2017

Abstract

It is well known that the air quality of 92% of the world is known to exceed the standard of WHO and the death caused by air pollution is almost 6 million per year. The PM_{2.5} concentration in Korea is the second most serious among the OECD countries following Turkey. Since the PM_{2.5} has a direct effect on the respiratory system, it has been actively studied in domestic and foreign countries. But current research on the PM_{2.5} is limited in weather factor or air pollutants. In this paper, we consider the influence of spatial neighbor with weather factor or air pollutants using spatial panel model. We applied the proposed method to 25 borough of Seoul in Korea. The result shows a significant effect of spatial neighbor on the PM_{2.5} concentration fields.

Keywords: Air pollution, particulate matter 2.5, risk assessment, spatial analysis, spatial panel model.

[†] This paper is based on part of Jong Hyun Lee's Master thesis.

¹ Graduate student, Department of Statistics, Kyungpook National University, Daegu 41566, Korea.

² Assistant professor, Department of Statistics, Kyungpook National University, Daegu 41566, Korea.

³ Corresponding author: Associate professor, Department of Statistics, Kyungpook National University, Daegu 41566, Korea. E-mail: kim.1252@knu.ac.kr