

ORIGINAL ARTICLE

광양만권 주변지역 주민들의 대기오염 노출추정을 위한 방법론 비교 연구

정순원 · 조용성 · 양원호¹⁾ · 유승도 · 손부순^{2)*}

국립환경과학원 환경보건연구과, ¹⁾대구가톨릭대학교 산업보건학과, ²⁾순천향대학교 환경보건연구과

Comparison of Exposure Estimation Methods on Air Pollution of Residents of Industrial Complexes

Soon-Won Jung, Yong-Sung Cho, Won-Ho Yang¹⁾, Seung Do Yu, Bu-Soon Son^{2)*}

Environmental Health Research Division, National Institute of Environmental Research, Incheon 404-708, Korea

¹⁾*Department of Occupational Health, Catholic University of Daegu, Daegu 712-702, Korea*

²⁾*Department of Environmental Health Science, Soonchunhyang University, Asan 336-745, Korea*

Abstract

The assessment of personal exposure is a critical component in population-based epidemiologic studies of air pollution. This study was conducted to apply and compare the four exposure estimation methods of individual-level to air pollution concentration in a cohort including 2,283 subjects in Gwangyang, Korea. Individual-level exposure of air pollution were estimated using multiple approaches, including average across all monitors, nearest monitor, and spatial interpolation by inverse distance weighting and kriging. The mean concentrations of PM₁₀, NO₂, SO₂, CO, O₃ by four exposure estimation methods were slightly different but not significantly different from each other. Cross-validation showed that kriging was more accurate than other exposure estimation methods because kriging has probably predicted individual exposure levels equivalent to residential locations after estimating the parameters of a model according to the spatial surface of air pollution concentration. These data support that spatial interpolation methods may provide better estimates than selecting the value from the nearest monitor and averaging across values from all monitors by reflecting spatial attributes of air pollution on personal level.

Key words : Air pollution, Cross-validation, Estimation method, Interpolation method

1. 서론

20세기 중반 4000명 이상을 죽음으로 몰아간 런던 스모그 사건(London Ministry of Health, 1954), 미국의 도노라 사건(Schrenk 등, 1949), 벨기에의 뮤즈벨

리 사건(Firket, 1936) 등은 대기오염이 인간의 건강에 심각한 악영향을 줄 수 있다는 역사적 사건으로 잘 알려져 있다. 이러한 대기오염 사건들은 결과적으로 대기오염과 사람의 건강영향에 대한 관련성을 본격적으로 조사·연구하게 된 계기가 되었다. 1990년대 초기

Received 30 July, 2012; Revised 21 November, 2012;

Accepted 31 January, 2013

*Corresponding author : Bu-Soon Son, Department of Environmental Health Science, Soonchunhyang University, Asan 336-745, Korea
Phone: +82-41-530-1270
E-mail: sonbss@sch.ac.kr

© The Korean Environmental Sciences Society. All rights reserved.

© This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

에 시작된 대기오염 노출평가는 역학연구에서 대기오염과 건강영향의 관련성을 규명하기 위한 중요한 요소로, 지난 20년 동안 대기오염으로 인한 인체노출수준을 평가하기 위한 많은 방법들이 이용되었다. 일반적으로 대기오염 역학연구는 인구집단을 대상으로 하기 때문에 유해인자로 인한 수용체의 실제 노출(actual exposure)을 평가하기 위해서는 인구집단이 아닌 연구대상자 개인에 대한 노출평가를 수행하는 것이 가장 바람직한 방법이라고 할 수 있다. 그러나 개인별 직접적 노출평가는 샘플링(sampling)의 어려움뿐만 아니라 비용과 시간 등이 많이 소요되기 때문에 직접적인 개인노출평가 방법을 대체할 수 있는 타당성 있는 간접적 개인노출추정 방법을 이용한 연구들이 이루어졌다. 지금까지 대부분의 대기오염 역학연구들에서는 국가에서 운영하는 대기오염 측정망 자료를 이용하여 공간적으로 평균화된 대기오염 농도가 다양한 위치에 거주하는 사람들의 노출수준을 대표한다는 가정 하에, 측정망이 위치하는 행정구역 단위를 기준으로 하여 하나 혹은 두 지점의 측정망에서 측정된 대기오염 농도를 개인의 노출로 할당하였다(Son 등, 2009). 그러나 이러한 노출평가 방법은 대기오염으로 인한 건강위해 크기를 평가함에 있어 개인의 공간적 이질성(heterogeneity)을 반영하지 못하여 분류오류(exposure misclassification) 등과 같은 불확실성으로 인하여 그 결과가 과소평가될 수 있는 가능성이 제기되어 왔다(Bell, 2006; Kim 등, 2009).

최근 전 세계적으로 GIS(geographic information systems) 발달과 함께 활용 가능한 지리공간자료(geospatial data)가 증가하고 공간분석 방법론들이 발전하면서 환경역학과 노출측정 및 평가 분야에서 새로운 도구로 공간분석 기법이 이용되고 있다(Han과 Hwang, 2011). GIS는 하나의 새로운 방법론으로 대기오염노출의 분류오류를 최소화하고 보다 정확한 노출평가의 필요성이 증가하면서 지리정보를 이용하여 개인의 건강관련 자료와 환경요인들을 공간적으로 표시하고, 이들의 연관성을 기술하고 분석하는 데 사용되고 있다. 또한, GIS를 이용하여 모니터링 자료를 이용할 수 없는 지역이나 기존의 노출평가를 개선하기 위하여 공간분석을 이용해 개인의 노출을 예측하는 모델 및 새로운 방법들을 적용하고 개발하는 다양한

연구들이 수행되고 있다(Finkelstein 등, 2003; Liao 등, 2006; Pikhart 등, 2001). 그러나 국내에서는 대기오염으로 인한 건강영향 평가 시 LUR(land use regression) 모델, Kriging과 같은 내삽법(interpolation method) 등 좀 더 정확하고 세밀한 노출수준을 추정하기 위한 연구들이 많지 않은 실정이다(Eum, 2008; Son 등, 2010). 따라서 본 연구에서는 주요 산업도시인 광양만권에서 특정 요인에 노출된 집단과 노출되지 않은 집단을 추적하여 질병 발생 관계를 추적하는 코호트 연구 참여자를 대상으로 하였다. 2009년 광양만권 코호트 참여자의 개인 단위 거주지 위치정보를 이용하여 대기오염물질의 4가지 개인노출추정 방법을 적용해보고 그 결과를 비교해보고자 한다.

2. 자료 및 방법

2.1. 연구대상

연구지역은 광양시, 여수시, 하동군이 포함된 광양만권으로 광양제철소, 국가산업단지, 화력발전소 등 철강 산업과 석유화학산업이 집중 배치되어 있는 산업 단지이다. 본 연구에서는 2009년 광양만권 코호트 참여자 총 2,504명 중 주소지 정보가 정확하지 않은 사람을 제외한 총 2,283명을 대상으로 하였다. 대상자의 성별 분포를 살펴보면 여자가 1299명(56.9%)으로 남자 984명(43.1%) 보다 많았으며, 연령별로는 20세 이하 참여자가 1019명(44.7%), 21~60세 472명(20.6%), 61세 이상 792명(34.7%)이 참여하였다. 한편, 대기오염 측정망 및 연구대상자의 위치는 Fig. 1과 같고, 모든 공간위치 정보는 TM(Transverse mercator) 좌표로 변환하여 이용하였다.

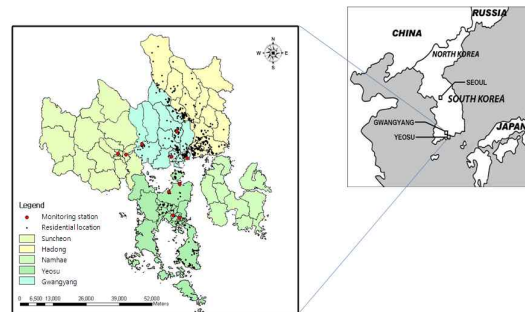


Fig. 1. The location of study areas.

2.2. 대기오염 측정망 자료

대기오염측정망 자료는 2009년 광양시 중동, 태인동, 진상면, 광양읍의 4개 지점과 여수시 광무동, 중흥동, 월내동, 문수동 4개 지점, 그리고 순천시 장천동과 연향동 2개 지점의 측정망에서 측정되는 미세먼지(PM₁₀), 이산화질소(NO₂), 이산화황(SO₂), 오존(O₃), 일산화탄소(CO)의 자료를 이용하여 일일 평균값을 산출하였다. PM₁₀, SO₂, NO₂는 24개의 시간별 평균값을 이용하였으며, O₃, CO는 8시간 이동평균값(이하, O_{3_mean}, CO_mean) 및 최대이동평균값을(이하, O_{3_max}, CO_max) 이용하였다.

2.3. 개인노출추정 방법

2.3.1. Average across all monitor

광양만권에 설치되어 있는 국가 대기오염 측정망 자료를 이용하여 모니터링된 측정값들의 평균치를 개인의 노출로 할당하였다.

2.3.2. Nearest monitor

개인의 거주지 위치정보를 이용하여 가장 가까운 거리에 있는 측정망의 값을 개인의 노출추정치로 할당하는 방법으로 각 위치정보는 TM 좌표로 변환하여 사용하였다.

2.3.3. Inverse distance weighting(IDW)

IDW는 역거리가중치 방법으로 이미 알고 있는 값으로부터 알고자 하는 위치의 값을 보간하는 방법이다. 내삽 하고자 하는 위치에서부터 각 측정지점까지의 거리에 역수를 가중치로 이용하는 방법으로, 예측 하고자 하는 위치와 가까운 거리에 있는 측정 지점일 수록 멀리 떨어져 있는 측정소에 비해 더 많은 가중치를 주는 방법이다. 본 연구에서는 IDW 기법을 이용한 추정치를 구함에 있어서 광양시 및 광양만권 전체 측정소 자료를 모두 포함하도록 하며, 다음과 같은 가중치를 사용하였다(Kinney 등, 1998; Kunzli 등, 1997).

$$W_i = 1/d(x_0, x_i)$$

여기서, W_i는 가중치이며, d_i는 보간 하고자 하는 지점(x₀)과 이미 알고 있는 관측지점(x_i) 쌍 간의 거리를 의미한다.

2.3.4. Kriging

Kriging 기법은 특정 영역 내 관측된 지점들 간의

거리와 관측값의 분산정도를 이용하여 알려져 있지 않은 지점의 공간적 분포를 예측하는 방법이다. 크리깅의 주요 과정은 3가지 단계로 경험적 베리오그램(empirical variogram)의 추정, 추정된 베리오그램에 대한 이론적 베리오그램(theoretical variogram) 결정, 알려지지 않은 위치에 대한 값을 예측하는 방법으로 이루어진다. 베리오그램이란 공간데이터의 상관성을 나타내는 척도로서 지리통계학적 데이터를 분석하는 과정의 첫 번째 단계이면서 매우 중요한 역할을 수행한다. 베리오그램은 거리 h만큼 이격된 두 자료들 간의 차이를 제공하여 계산한 값으로 식은 다음과 같다.

$$2r(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2$$

여기서, 2r(h)는 베리오그램, r(h)은 반베리오그램(semi-variogram), h는 지연거리(lag distance), N(h)은 자료 쌍의 개수, z(x_i)는 임의의 위치 x_i에서의 측정값을 의미한다. 즉, 두 측정값들 간의 거리가 가까우면 측정값들이 유사한 경향을 보이므로 베리오그램은 자료의 거리가 가까울수록 작고 멀어질수록 크다. 이러한 성질의 베리오그램은 이격거리에 따른 자료들의 관계를 분석하여 공간적 자료 분포 구조를 파악한 후 예측하고자 하는 지점의 측정값을 추정하는 데 활용된다. 다음 단계는 주어진 자료로부터 계산된 경험적 베리오그램을 가장 잘 대표하는 이론적 베리오그램을 결정하는 것이다. 이론적 베리오그램 모델로는 구모형(spherical model), 지수모형(exponential model), 가우시안모형(gaussian model) 등이 가장 많이 사용되며 본 연구에서는 구모형을 이용하였다.

2.4. 교차검증(cross-validation)

본 논문에서는 Nearest monitor, IDW, Kriging으로 추정한 값의 정확성을 평가하기 위하여 교차검증을 수행하였다. 노출추정에 이용된 10개 측정망 중에서 각각 1개씩 제외한 나머지 9개의 측정망 자료를 이용하여 제외된 측정망의 농도를 추정하였다. 또한, 교차검증을 통하여 예측결과의 정확성을 정량적으로 평가하기 위해 비교 통계량을 사용하였다. 비교통계량으로는 예측값의 정확성을 평가하는 평균제곱오차(mean square error, 이하 MSE)와 표준화된 평균제곱오차(normalized mean square error, 이하 NMSE)를 계산

하여 비교하였고, 측정값과 예측값의 일관성은 변화 계수(coefficient of divergence, 이하 COD)를 계산하여 비교하였다. COD는 0에 가까울수록 두 지점의 농도 차이가 없는 것을 의미하고, 1에 가까울수록 공간적 이질성(heterogeneity)이 확실한 것을 의미한다 (Wilson 등, 2006).

- MSE= average(observed-predicted)²
- MMSE=average((observed-predicted)²/observed)×100%
- COD(coefficient of divergence) =

$$\sqrt{\frac{1}{p} \sum_{i=1}^p [(x_{ij} - x_{ik}) / (x_{ij} + x_{ik})]^2}$$

여기서, P는 측정수이고 X_{ij}는 i날에 측정된 평균 농도이고 X_{ik}는 i날에 예측된 평균 농도이다.

2.5. Passive sampler 측정치와 추정치와의 비교

코호트 참여자 중 150명을 대상으로 NO₂ passive sampler를 이용하여 월요일부터 금요일까지(9/14-27)

4-5일 동안 참여자의 주택 실외에서 측정하였다. 이중 측정이 잘 이루어지지 않은 sample을 제외한 132개의 실외 측정치와 본 연구에서 이용된 4가지 노출추정 방법으로 추정된 예측치와 비교하였다.

2.6. 통계분석

자료의 통계학적 분석은 SAS version 9.1(SAS Institute, Cary, NC)과 R 2.12.1을 이용하였다. 범주화된 변수의 관련성은 chi-square 검정을 사용하였으며, 평균의 비교는 t 검정 또는 분산분석을 사용하였다.

3. 결과

3.1. 연구지역의 대기오염물질 특성

Fig. 2는 2009년도에 연구지역에서 측정된 NO₂, SO₂, O₃, CO, PM₁₀의 시간에 따른 농도 변화 추이를 나타낸 그림이다. 본 연구지역의 시간에 따른 대기오염물질의 일평균 농도를 보여주고 있다. 오존은 계절에 따른 변화를 가장 잘 보여주고 있으며, 2월부터 농

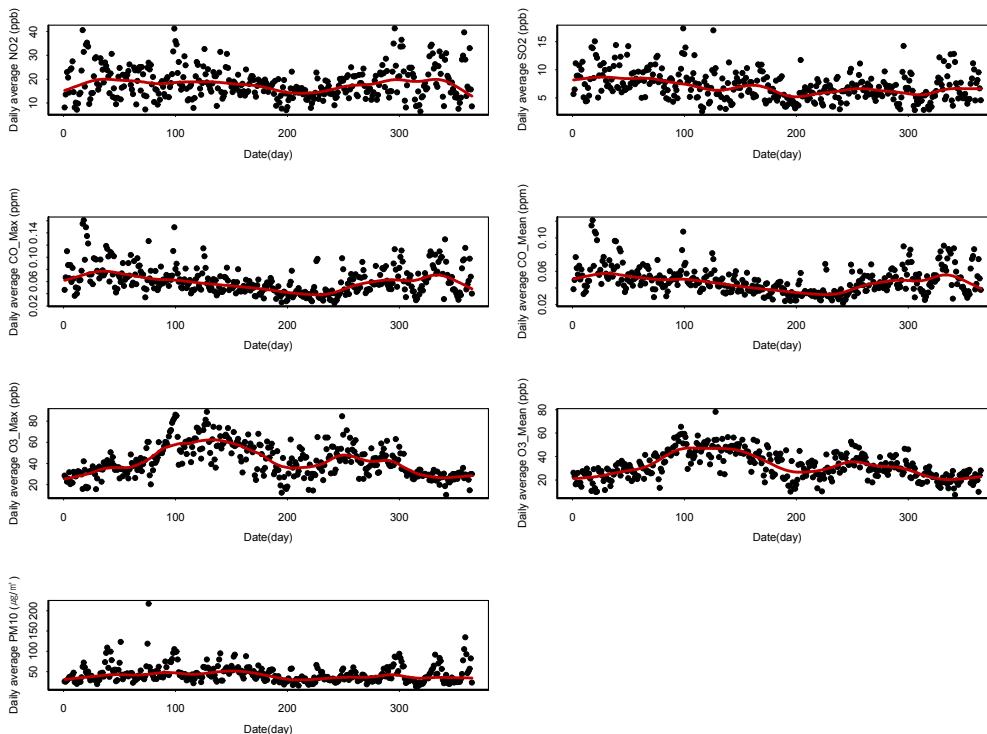


Fig. 2. Time trends of daily average concentration of NO₂, SO₂, CO, O₃ and PM₁₀ in Gwangyang Korea, 2009.

도가 점차 높아지다 4~5월에 농도가 가장 높은 경향을 나타냈으며 겨울철에 낮은 농도를 보이는 것으로 나타났다. 이산화황과 이산화질소 및 일산화탄소도 겨울철에 농도가 약간 높은 경향을 보였으나, PM₁₀은 계절적인 변화 경향은 없는 것으로 나타났다.

Fig. 3은 각 오염물질의 공간적 이질성을 조사하기 위하여 각 오염물질별 측정망 쌍의 상관계수를 계산하여 거리에 따른 상관성을 보여주고 있다. SO₂는 측정망간 거리가 증가함에 따라 다른 물질들에 비해 상관계수가 가장 크게 감소하는 것으로 나타나 공간적 이질성이 가장 큰 것으로 나타났다. NO₂와 CO도 거리가 증가함에 따라 측정망간 상관계수가 감소하는 경향을 보였고, PM₁₀과 O₃은 측정망간 상관계수가 높게 나타났으며, 거리가 증가함에 따른 감소 경향이 다른 물질들에 비해 낮게 나타났다.

3.2. 개인노출 추정치

4가지 개인노출 추정방법을 적용하여 연구대상자

들의 대기오염노출 예측치의 기술통계량을 Table 1에 나타내었다. PM₁₀ 농도는 Nearest monitor가 평균 45.38 $\mu\text{g}/\text{m}^3$ 로 가장 높은 것으로 나타났으며, IDW 45.02 $\mu\text{g}/\text{m}^3$, Kriging 44.77 $\mu\text{g}/\text{m}^3$, All monitors 43.92 $\mu\text{g}/\text{m}^3$ 순으로 나타났으나 통계적으로 유의한 차이를 보이지는 않았다. NO₂와 SO₂ 및 O₃, CO의 경우도 노출추정방법별 평균농도는 약간씩 차이가 나타났으며, NO₂는 All monitors가 18.47 ppb, SO₂는 IDW가 7.56 ppb, O₃은 All monitors가 8시간 이동평균 및 최대이동평균이 각각 31.92 ppb, 42.22 ppb, CO는 Nearest monitor가 8시간 이동평균 및 최대이동평균이 각각 0.542 ppm, 0.683 ppm으로 다른 추정방법보다 높았으나 통계적으로 유의한 차이는 나타나지 않았다. 각 노출추정 방법별 오염물질의 예측치는 어느 특정 방법에 치우친 결과가 아닌 오염물질에 따라 양상이 다른 것으로 나타났다. 한편, 각 노출추정 방법별 표준편차는 모든 물질에서 All monitors 방법이 가장 작은 것으로 나타났으며, Nearest monitor 방법은 다른 방법

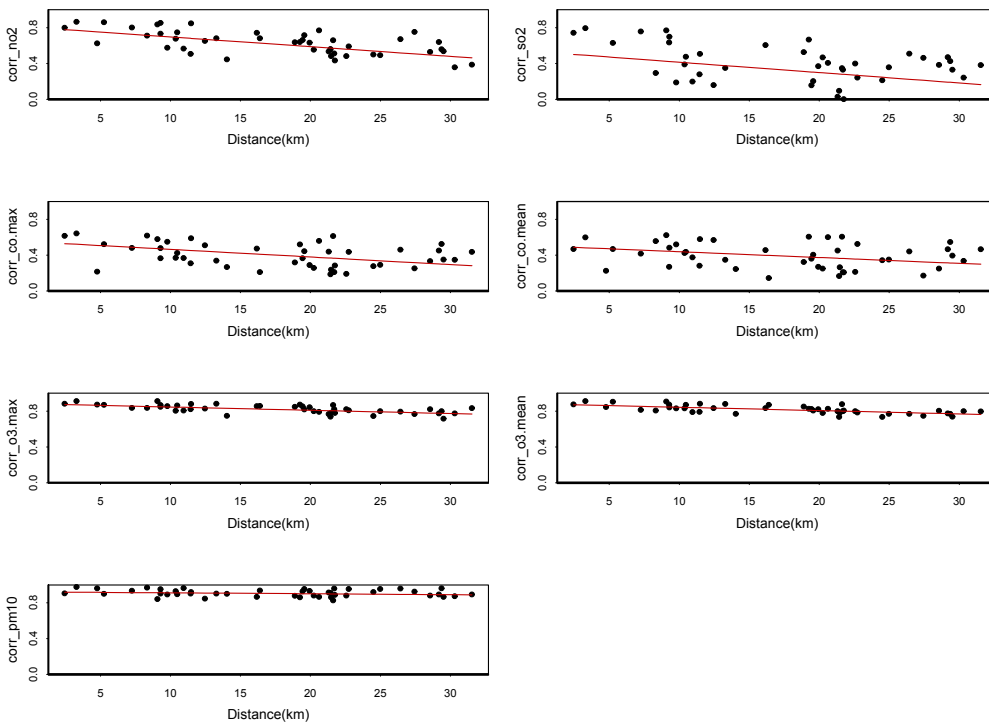


Fig. 3. Correlation coefficients by distance between monitor pairs during study period(2009).

Table 3. Summary statistics comparing the result of the four different methods for daily air pollutant concentrations in Gwangyang, 2009(N=2,283)

Air pollutant	Method	Mean	S.D.	Min	25%	Median	75%	Max
Daily average PM ₁₀ ($\mu\text{g}/\text{m}^3$)	All monitors	43.86	21.98	12.92	29.13	39.09	52.07	217.28
	Nearest monitor	45.38	24.07	4.71	28.59	40.29	55.83	326.29
	IDW	45.02	22.48	5.00	29.69	39.91	54.62	325.22
	Kriging	44.77	22.16	5.67	29.80	39.49	54.22	265.44
Daily max 8-h moving average O ₃ (ppb)	All monitors	42.22	15.46	10.44	30.08	39.39	53.98	88.63
	Nearest monitor	40.99	16.59	3.71	28.88	38.88	51.75	113.75
	IDW	41.40	15.63	3.80	29.74	39.00	52.16	113.13
	Kriging	41.68	15.85	5.90	29.55	39.23	52.58	113.49
Daily 8-h moving average O ₃ (ppb)	All monitors	31.92	11.86	7.25	23.41	29.75	40.89	78.00
	Nearest monitor	31.40	12.82	2.26	22.07	30.10	39.88	89.62
	IDW	31.58	12.03	2.30	22.97	30.14	40.00	89.27
	Kriging	31.78	12.16	4.61	22.94	30.54	40.21	94.40
Daily average NO ₂ (ppb)	All monitors	18.47	6.70	6.43	13.71	17.25	22.02	41.28
	Nearest monitor	17.99	9.30	1.39	11.14	16.67	23.18	58.71
	IDW	18.11	7.85	1.39	12.69	17.14	22.52	58.65
	Kriging	17.70	7.86	1.31	12.21	16.73	22.14	58.61
Daily average SO ₂ (ppb)	All monitors	7.14	2.60	2.62	5.16	6.76	8.47	17.31
	Nearest monitor	7.39	5.01	1.00	4.04	6.00	9.29	51.88
	IDW	7.56	3.95	1.03	4.74	6.67	9.43	51.87
	Kriging	7.50	3.96	0.10	4.74	6.56	9.10	51.71
Daily max 8-h moving average CO(ppm)	All monitors	0.62	0.24	0.25	0.45	0.57	0.73	1.61
	Nearest monitor	0.68	0.40	0.10	0.40	0.60	0.88	3.28
	IDW	0.66	0.32	0.10	0.44	0.59	0.81	3.17
	Kriging	0.65	0.31	0.09	0.44	0.58	0.79	3.07
Daily 8-h moving average CO(ppm)	All monitors	0.49	0.17	0.22	0.36	0.46	0.60	1.21
	Nearest monitor	0.54	0.29	0.05	0.32	0.49	0.69	2.69
	IDW	0.52	0.23	0.05	0.36	0.47	0.63	2.58
	Kriging	0.51	0.22	0.06	0.36	0.47	0.61	2.31

들에 비해 표준편차가 큰 것으로 나타났다.

3.3. 오염물질 추정방법별 상관성 분석

대기오염물질 농도간의 관련성을 분석하기 위하여 4가지 노출추정 방법별로 Pearson 상관분석을 실시하였다. PM₁₀은 다른 대기오염물질들과 모두 상관성이 있는 것으로 나타났으나($p < 0.001$), 상관계수는 높지 않았으며($R = 0.292 \sim 0.532$), 나머지 추정방법들도 비슷한 경향을 나타내었다. 한편, O₃은 CO에서만 4가지 방법 모두 음의 상관성이 있는 것으로 나타났으며($p < 0.001$), 다른 오염물질들과는 상관성이 없는 것으로 나타났다. Fig. 4는 각 노출추정 방법별 추정값을 비교하기 위하여 산점도 및 상관계수를 계산하여 나타내었다. 상관분석 결과, Kriging, IDW, Nearest

monitor 추정방법 간에는 모든 물질에서 상관관계가 높은 것으로 나타났고($0.81 \sim 0.99$), All monitors 방법은 SO₂를 제외한 모든 물질에서 Kriging과 가장 높은 상관관계가 보였으며($0.86 \sim 0.99$), Nearest monitor와 가장 낮은 상관관계를 보였다($0.68 \sim 0.94$). 한편, PM₁₀의 경우 4가지 노출추정 방법 모두 상관성이 높은 것으로 나타났으며($0.94 \sim 0.99$), NO₂는 Kriging과 IDW의 상관계수가 0.96으로 상관성이 가장 높았으며, All monitors와 Nearest monitor는 0.68로 가장 낮은 것으로 나타났다. SO₂의 경우는 All monitors 방법과 다른 방법들과의 상관성은 Nearest monitor가 0.33, Kriging이 0.58, IDW가 0.61로 상관관계가 낮은 것으로 나타났다. O₃과 CO는 8시간 이동평균 및 최대이동평균 모두 Kriging과 IDW가 상관관계가 높은 것으로 나타났다.

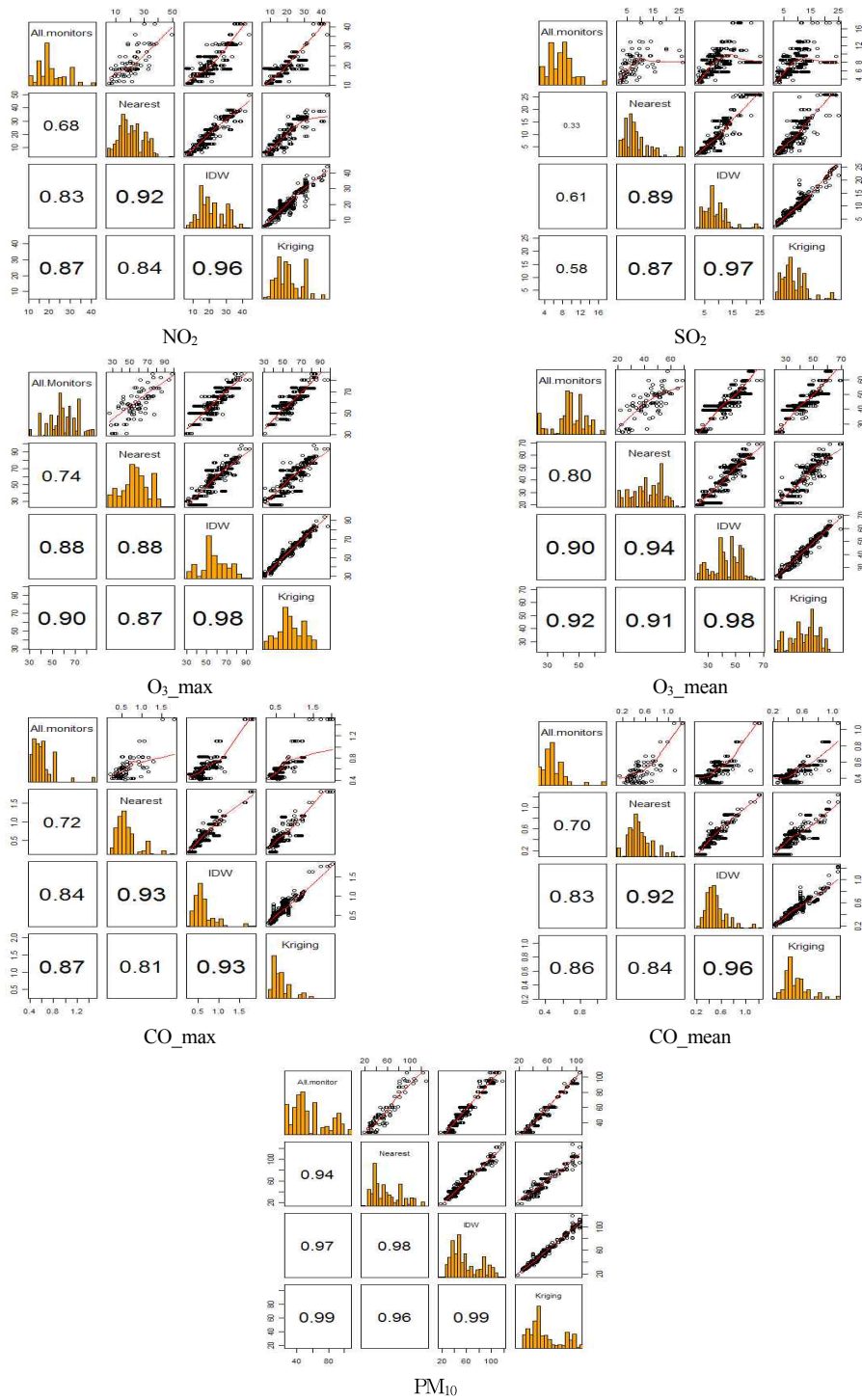


Fig. 4. Pairwise scatter plots and pearson correlation coefficients for estimated daily PM₁₀ concentration.

3.4. 교차 검증(Cross-validation)

본 논문에서는 Nearest monitor, IDW, Kriging으로 추정된 값의 정확성을 평가하기 위하여 교차검증을 수행하였다. 노출추정에 이용된 10개의 측정망 중에서 각각 1개씩 제외한 측정망의 측정값을 나머지 9개의 측정망 자료를 이용하여 추정된 값과 비교하였다. 평균제곱오차(mean square error, MSE)와 표준화된 평균제곱오차(normalized mean square error, NMSE)는 모든 물질에 있어서 Kriging이 오차가 가장 낮게 나타났으며, Nearest monitor가 가장 오차가 큰 것으로 나타났다(Table 2). 변화계수(coefficient of divergence, COD)는 PM₁₀의 경우 Kriging이 0.115로 IDW 0.134, Nearest monitor 0.171보다 낮은 것으로 나타나 본 연구에서는 Kriging 방법이 다른 방법들보다 공간적 이질성이 가장 낮은 것으로 나타났다.

Table 2. Results from cross-validation data for 2009

Pollutant	Method	MSE	Normalized MSE	COD
PM ₁₀	Nearest monitor	192.50	421.09	0.171
	IDW	124.11	264.56	0.134
	Kriging	102.35	214.36	0.115
NO ₂	Nearest monitor	58.42	360.95	0.207
	IDW	23.65	141.27	0.131
	Kriging	19.75	101.22	0.119
SO ₂	Nearest monitor	46.12	653.70	0.298
	IDW	20.17	265.41	0.217
	Kriging	15.36	208.57	0.188
O ₃ _max	Nearest monitor	110.99	289.86	0.139
	IDW	76.61	184.61	0.113
	Kriging	55.96	148.99	0.099
O ₃ _mean	Nearest monitor	63.91	226.82	0.144
	IDW	45.26	144.23	0.116
	Kriging	30.19	136.25	0.101
CO_max	Nearest monitor	0.193	35.880	0.304
	IDW	0.082	14.151	0.212
	Kriging	0.069	10.457	0.156
CO_mean	Nearest monitor	0.116	28.536	0.303
	IDW	0.051	11.735	0.215
	Kriging	0.041	8.459	0.159

3.5. 개인노출 예측치와 passive sampler 실측치의 비교

광양만권 코호트 참여자 중 132명을 대상으로 거주지 실외에서 passive sampler로 측정된 NO₂ 실측치와

본 연구에 이용된 각 노출추정 방법의 예측치를 비교하였다(Fig. 5). 실측치와 예측치의 관련성을 분석하기 위하여 상관분석을 실시하였다(Table 3). 상관분석 결과, 4가지 노출추정 방법 모두 예측치와 유의한 상관관계가 있는 것으로 나타났으며($p < 0.001$), 상관계수는 Kriging이 0.488로 상관계수가 가장 높았으며, Nearest monitor 0.478, IDW 0.369, All monitors 0.367 순으로 나타났으나 상관관계가 높지는 않았다.

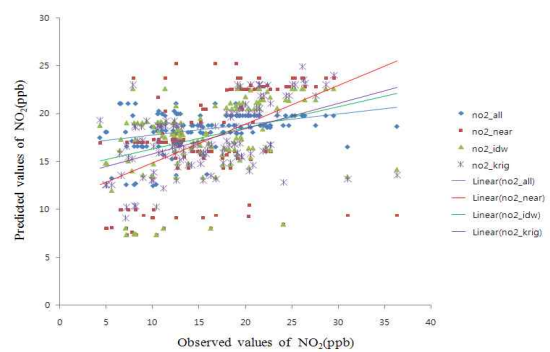


Fig. 5. The Distribution of between predicted values by prediction methods and observed values by passive sampler for NO₂.

Table 3. Pearson correlation coefficients between prediction methods and observed values by passive sampler for NO₂

Pollutant	Passive	All monitor	Nearest monitor	IDW	Kriging
Passive	1	0.367**	0.478**	0.369**	0.488**
All monitor		1	0.528**	0.583**	0.536**
Nearest			1	0.879**	0.827**
IDW				1	0.916**
Kriging					1

** : $p < 0.001$, * : $p < 0.05$

4. 고찰

본 연구에서는 다양한 공간분석 방법 중 선행 연구들에서 많이 이용된 개인노출 추정방법인 average across all monitors, nearest monitor, idw, kriging 방법을 적용하여 직접적인 개인노출 평가방법을 대체할

수 있는 개인노출수준을 추정하였다. 이용된 각 노출 추정 방법의 예측치를 비교한 결과, 추정방법별 PM₁₀, NO₂, SO₂, CO, O₃의 평균 농도는 약간씩 차이가 있었으나, 통계적으로 유의한 차이는 나타나지 않았으며, 각 노출추정 방법별 상관분석을 한 결과 Kriging, IDW, Nearest monitor 노출추정 방법 간에는 모든 물질에서 상관관계가 높은 것으로 나타났다(0.81~0.99). 한편, All monitors 방법은 SO₂를 제외한 모든 물질에서 Kriging과 가장 높은 상관관계를 보였으며(0.86~0.99), Nearest monitor와 가장 낮은 상관관계를 보였다(0.68~0.94). Son 등(2009)의 연구에서는 PM₁₀, NO₂, O₃의 노출추정 방법 간(kriging, idw, nearest monitor) 상관계수가 0.84~0.98로 높은 상관관계를 나타내는 것으로 보고하고 있어 본 연구결과와 유사한 결과를 나타내었지만 SO₂와 CO는 양상이 조금 다르게 나타났다. 이는 오염물질이 측정된 시기, 측정망의 위치와 지역적 차이 등 시·공간적 차이에 의한 것으로 생각된다.

본 연구에서는 각 오염물질의 공간적 이질성을 조사하기 위하여 측정망간(pairs) 거리 및 측정값의 상관계수와와의 관련성을 조사하였다. 측정망간 거리는 약 2.4 km~31.5 km이었으며, SO₂ 및 NO₂는 다른 물질들에 비해 거리가 증가할수록 상관계수가 감소하는 경향이 큰 것으로 나타났으며, 측정망간 거리가 증가함에 따라 상관계수가 가장 큰 감소를 보인 물질은 SO₂로 나타났다. Son 등(2010)은 측정망간 거리가 약 20 km 떨어져 있는 모니터 측정값의 상관계수 사이의 관련성을 조사한 결과, 거리가 증가함에 따라 SO₂의 공간적 패턴이 가장 큰 것으로 보고하였다. 또한, Sarnat 등(2010)은 측정망과 연구대상에 있어 측정망의 위치와 거리는 인구 밀집지역에서 20마일(약 32 km) 안에 위치한 도시지역으로부터의 자료를 이용했을 때 어떤 오염물질도 유의하게 영향을 끼치지 않는 것으로 나타났지만, 공간적으로 이질적인 오염물질(CO, NO₂)은 인구 밀집지역으로부터 30마일(약 48 km) 이상 떨어져 있는 자료를 이용했을 때는 노출추정 오차가 있는 것으로 보고하였다. 따라서 공간적 이질성이 큰 오염물질을 대상으로 공간 내삽법을 적용할 경우에는 측정망 및 연구대상의 위치와 거리 등을 고려하여 분석할 필요가 있다. 또한, 노출추정에 사용

되는 방법 및 모델의 검증 등도 중요하게 고려되어야 한다.

본 연구에 이용된 노출방법의 정확성을 평가하기 위하여 교차검증을 수행한 결과, Kriging이 모든 물질에서 다른 공간 내삽방법(interpolation methods)보다 MSE(mean square error), NMSE(normalized mean square error), COD(coefficiency of divergence)가 가장 낮게 나타나 다른 추정방법보다 공간적 이질성이 작고 오차가 적은 것으로 나타났다. 또한, passive sampler를 이용한 NO₂ 실측치와 예측치를 비교한 결과, 상관계수는 Kriging이 0.488, Nearest monitor 0.478, IDW 0.369, All monitors 0.367로 실측치와 유의한 상관관계가 나타났으나 상관관계가 높지는 않았다. 이는 예측치와 비교하기 위한 실측치의 샘플수가 132개로 샘플수가 많지 않았고, passive sampler의 측정 위치, 능동식 채취기가 아닌 passive sampler로 인한 측정오차가 ±20%인 점 등이 상관계수에 영향을 준 것으로 생각된다. 한편, NO₂ 실측치와의 상관성은 Kriging 등 공간내삽 방법들이 All monitors 방법보다 상관관계가 높은 것으로 나타났다(p<0.001). All monitor가 다른 방법들과 상관계수가 낮은 이유는 본 연구에서 NO₂가 측정망간 거리가 증가함에 따라 공간적 이질성이 크게 나타났기 때문에 거리에 따른 가중치를 적용한 다른 추정 방법들과는 비교적 상관관계가 낮게 나타난 것으로 생각된다.

본 연구대상인 광양만권은 주요 산업도시로 광양시에 4개 여수시에 4개의 측정망만이 운용되고 있어 공간분석을 이용한 노출추정 방법에 이용할 수 있는 측정망이 상대적으로 적기 때문에 공간 내삽을 할 때 인근 지역으로부터의 측정망을 사용할 경우 연구대상이 측정망으로부터 멀리 떨어져 있거나 측정망 바깥쪽에 위치하는 경우 노출수준을 과대평가할 수 있는 가능성이 있다. 그러므로 측정망의 위치가 균등히 분산되고, 측정망의 수가 많아진다면 인구집단을 대상으로 하는 역학연구 등에서 노출추정에 있어 더 정확한 예측치를 얻을 수 있을 것으로 생각된다.

노출평가는 유해인자의 발생원에서 수용체(인간)의 실제 또는 잠재적 노출정도를 평가하는 중요한 분석적 도구로 건강 위해성평가와 역학연구에서 필수적인 부분이라고 할 수 있다. 본 연구에서는 대기오염으

로 인한 역학연구의 노출평가에 있어 공간분석 기법을 이용하여 보다 정확한 노출수준을 추정하고자 하였다. 이러한 공간분석 기법을 이용한 대기오염의 노출수준을 추정함에 있어서 몇 가지 제한점을 고려해야 할 필요가 있다. 우선 개인단위 위치정보에 기반을 둔 자료의 공간위치정보의 정확성은 중요한 의미를 가지고 있다. 주소매칭 방법의 사용에 따른 공간위치정보의 오차는 노출추정 분류오류로 인한 연구의 검정력을 낮출 수 있고, 위험도 추정치를 편향시킬 가능성이 있기 때문이다. 따라서 GPS나 원격탐사를 통해 얻은 실 측정값을 기준값(gold standard)으로 사용하여 비교·분석이 고려될 필요가 있다(Han과 Hwang, 2011). 두 번째로 노출추정시 이용되는 모형의 적합성 문제가 있을 수 있다. IDW와 all monitor의 경우 search radius의 범위를 어떻게 설정하는 것이 중요하지만 이에 대한 기준이 없어 연구결과의 객관성이 부족할 수 있다. 그러나 Wong 등(2004)은 search radius의 범위를 어떻게 설정하는지에 따라 연구결과에 큰 영향을 미치지 않는 것으로 보고했다. IDW는 단순히 거리에 따른 가중치를 적용하여 예측하고자 하는 지점의 값을 추정하는 방법이지만 Kriging은 관측된 지점들 간의 거리와 관측값의 분산정도를 이용하여 알려져 있지 않은 지점의 공간적 분포를 예측하는 방법으로 더 좋은 추정치를 제공할 수 있다. 그러나 본 연구결과는 광양만권이라는 특정 지역에 적용된 결과로서 Kriging 방법이 다른 노출추정방법보다 나은지 판단하기는 어려우며 공간적 특성 및 오염물질에 따라 각기 다른 방법이 적용되어야 할 것으로 생각 된다. 마지막으로 연구대상자의 건강영향을 평가하기 위한 오염도는 연구대상자들이 지역의 대표적인 오염도에 유사한 수준으로 노출되는 것으로 가정하고 있지만, 실제로는 한 장소에 머물지 않고 계속적 또는 주기적으로 이동하고 있고, 실외보다 실내에서 보내는 시간이 더 길기 때문에 개인의 시간활동양상(time activity pattern) 등을 반영한 개인노출수준 모니터링 방안들에 대한 연구가 필요하다.

5. 결론

본 연구는 2009년 광양만권 코호트 참여자 2,283명

을 대상으로 개인의 거주지 정보를 이용하여 Average across all monitors, Nearest monitor, Inverse distance weighting, Kriging 노출추정 방법을 적용해보고, 각 노출추정 기법의 추정값을 비교하기 위해 수행되었다. 각 개인노출 추정방법을 적용하여 예측치를 비교한 결과, 추정방법별 PM₁₀, NO₂, SO₂, CO, O₃의 평균 농도는 약간씩 차이가 나타났으나 통계적으로 유의한 차이는 나타나지 않았다. 대기오염 물질의 공간적 이질성을 조사하기 위하여 측정망간(pairs) 거리 및 모니터 측정값의 상관관계 사이의 관련성을 조사한 결과, 거리가 증가함에 따라 SO₂ 및 NO₂는 다른 물질들에 비해 공간적 이질성이 큰 것으로 나타났다. 각 노출추정 방법의 정확성을 평가하기 위하여 교차검증을 수행한 결과, Kriging이 모든 물질에서 오차가 가장 낮게 나타나 다른 추정방법보다 정확한 것으로 나타났으며, passive sampler를 이용한 NO₂ 실측치와 예측치를 비교한 결과에서도 Kriging이 실측치와 가장 상관성이 높은 것으로 나타났다($p < 0.001$). 본 연구에서는 Kriging 방법이 다른 노출추정 방법들보다 개인수준에서의 공간적 변이를 반영함으로써 더 좋은 추정치를 제공할 수 있었다. 이러한 접근방법은 환경역학과 노출평가 분야에 새로운 방법과 도구를 적용함으로써 대기오염과 건강영향의 복잡한 관계를 설명하는 데 기초자료로 이용될 수 있을 것으로 생각된다. 그러나 앞서 언급한 오염물질별 공간적 이질성, 개인의 시간활동양상 등을 반영하지 못한다면 개인노출 추정의 분류오류 및 불확실성은 여전히 남아있기 때문에 이러한 문제를 해결하기 위한 노력이 필요할 것으로 사료된다.

감사의 글

본 연구는 국립환경과학원 “지역주민 환경오염 노출수준 및 생체지표 모니터링(광양만권)” 연구사업의 일환으로 수행되었으며, 지원에 감사드립니다.

참고 문헌

Bell, M. L., 2006, The use of ambient air quality modeling to estimate individual and population exposure for human health research: A case study

- of ozone in the Northern Georgia Region of the United States, *Environ. Int.*, 32, 586-593.
- Eum, K. D., 2008, Spatial and temporal assessment of air pollution exposure and its health effects, Ph. D. Dissertation, Seoul National University.
- Finkelstein, M. M., Jerrett, M., Deluca, P., Finkelstein, N., Verma, D. K., Chapman, K., Sears, M. R., 2003, Relation between income, air pollution and mortality : a cohort study, *Can. Med. Assoc. J.*, 169, 397-402.
- Firket, J., 1936, Fog Along the Meuse Valley, *Trans. Faraday Soc.*, 32, 1192-1197.
- Han, D. K., Hwang, S. S., 2011, Spatial epidemiology and environmental health : On the use of spatially referenced health and environment data, *J. Env. Health Sci.*, 1, 1-11.
- Kim, S. Y., Sheppard, L., Kim, H., 2009, Health effects of long-term air pollution, influence of exposure prediction methods, *Epidemiology*, 20, 442-450.
- Kinney, P. L., Aggarwal, M., Nikiforov, S. V., Nadas, A., 1998, Methods development for epidemiologic investigations of the health effects of prolonged ozone exposure. Part III: an approach to retrospective estimation of lifetime ozone exposure using a questionnaire and ambient monitoring data (US sites), *Health Effects Inst. Res. Rep.*, 81, 79-107.
- Kunzli, N., Lurman, F., Segal, M., Ngo, L., Balmes, J., Tager, I., 1997, Association between lifetime ambient ozone exposure and pulmonary function in college freshman- results of a pilot study, *Environ. Res.*, 72, 8-23.
- Liao, D., Peuquet, D. J., Duan, Y., Whitsel, E. A., Dou, J., Smith, R. L., Lin, H. M., Chen, J. C., Heiss, G., 2006, GIS approaches for the estimation of residential- level ambient pm concentrations, *Environ. Health Perspect.*, 114, 1374-1380.
- Ministry of Health, 1954, Mortality and Morbidity during the London fog of December 1952, Reports on Public Health and Medical Subjects No. 95, London Ministry of Health, London.
- Pikhart, H., Bobak, M., Gorynski, P., Wojtyniak, B., Danova, J., Celko, M. A., Kriz, B., Briggs, D., Elliott, P., 2001, Outdoor sulphur dioxide and respiratory symptoms in Czech and Polish school children: a small-area study, *Int. Arch. Occup. Environ. Health*, 74(8), 574-578.
- Sarnat, S. E., Klein, M., Sarnat, J. A., Flanders, W. D., Waller, L. A., Mulholland, J. A., Russell, A. G., Tolbert, P. E., 2010, An examination of exposure measurement error from air pollutant spatial variability in time-series studies, *J. Expo. Sci. Environ. Epidemiol.*, 20, 135-146.
- Schrenk, H., Heimann, H., Clayton, G., Gafafer, W., Wexler, H., 1949, Air Pollution in Donora, PA: Epidemiology of the Unusual Smog Episode of October 1948, Preliminary Report, Public Health Bulletin No. 306, U.S. Public Health Service, Washington, DC.
- Son, J. Y., Kim, Y. S., Cho, Y. S., Lee, J. T., 2009, Prediction Approaches of Personal Exposure from Ambient Air Pollution Using Spatial Analysis : A Pilot Study Using Ulsan Cohort Data, *J. Kor. Soc. Atmos. Environ.*, 4, 339-346.
- Son, J. Y., Bell, M. L., Lee, J. T., 2010, Individual exposure to air pollution and lung function in Korea: Spatial analysis using multiple exposure approaches, *Environ. Res.*, 110, 739-749.
- Wilson, J. G., Kingham, S., Sturman, A. P., 2006, Intraurban variations of PM₁₀ air pollution in Christchurch, New Zealand: implications for epidemiological studies, *Sci. Total Environ.*, 367, 559-572.
- Wong, D. W., Yuan L., Perlin S. A., 2004, Comparison of spatial interpolation methods for the estimation of air quality data, *J. of Ex. Anal. and Env. Epi.*, 14, 404-415.