

환경지표와 도시공간과의 상관관계 분석에 관한 연구 - 교통부문 빅데이터를 중심으로 -

황기연, 김관용

To cite this article : 황기연, 김관용 (2019) 환경지표와 도시공간과의 상관관계 분석에 관한 연구 - 교통부문 빅데이터를 중심으로 - , 대한건축학회
연합논문집, 21:1, 41-47

① earticle에서 제공하는 모든 저작물의 저작권은 원저작자에게 있으며, 학술교육원은 각 저작물의 내용을 보증하거나 책임을 지지 않습니다.

② earticle에서 제공하는 콘텐츠를 무단 복제, 전송, 배포, 기타 저작권법에 위반되는 방법으로 이용할 경우, 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

www.earticle.net

환경지표와 도시공간과의 상관관계 분석에 관한 연구

- 교통부문 빅데이터를 중심으로 -

Analysis of Correlation between Environment Index and Urban Space Usage using Big Data

- Focused on Big Data in Transportation -

황기연*

김관용**

Hwang, Kee-Yeon

Kim, Kwan-Yong

Abstract

With the upcoming era of the 4th Industrial Revolution, big data has been constructed in various field. Especially in transport sector, tremendous passenger and vehicle trips are collected as a navigation big data. Through navigation data, previous spotted traffic volume and speed data contribute to establish traffic volume and speed data in national level. Established transport raw data in national level lead to various processed data such as mobility data, carbon dioxide(CO₂), air pollutants etc. This research utilize geographically weighted regression model in order to analyze correlation between area type (residential and business area) and air pollutants (CO₂, CO, PM-10, NOX, VOCs) in the city of Seoul. The research outcome describes low correlation between environmental indicators corresponding to navigation data related to vehicle trip and area type (residential and business area). Therefore, a following study is required to analyze correlation between environmental indicators corresponding to transport navigation data and socio-economic indicators rather than spatial features. Moreover, a comparative study between general regression model and geographically weighted regression model is necessary for in-depth analysis.

키워드 : 교통, 빅데이터, 환경지표, 도시공간

Keywords : Transportation, BigData, Environmental Index, Urban Space,

1. 서론

1.1 연구의 배경 및 목적

최근 정보통신기술의 발달과 인공지능의 등장으로 인해 제4차 산업혁명 시대를 통해 사회 전반적인 분야에서 변화와 혁신의 시대를 겪고 있다. 제4차 산업혁명 시대의 핵심 키워드는 빅데이터라고 할 수 있다. 빅데이터는 과거 아날로그 환경에서 생성되던 데이터에 비하면 그 규모가 방대하고, 생성주기도 짧으며, 형태 또한 수치 뿐만 아니라 문자와 영상 데이터를 포함하는 대규모 데이터를 뜻한다. 이러한 빅데이터는 기존에 접근 할 수 없었던 다양한 분석과 솔루션을 제공 할 수 있으며, 빅데이터와 빅데이터 간의 융합을 통해

신(新) 비즈니스 창출 등 다방면에서 그 활용성이 높다.

특히, 빅데이터는 가공되지 않은 1차 데이터 자체에도 그 의미가 크지만 1차 데이터를 활용한 2차 가공데이터의 활용성은 더욱더 크다고 할 수 있다.

교통분야에서도 최근 빅데이터를 활용하여 기존에 할 수 없었던 다양한 분석과 솔루션 연구를 통해 비용과 시간을 절감 하는 등 관련 연구가 활발히 진행되고 있다.

대표적인 교통부문 빅데이터는 스마트폰 단말기를 통해 구축되는 이동통신데이터와 차량의 이동경로 정보를 제공하는 내비게이션 데이터를 언급할 수 있다. 내비게이션 데이터는 기존에 한정된 지점 단위의 교통정보를 제공하는 것이 아니라, 경로형 데이터를 제공한다는 의미에서 그 의미가 크다고 할 수 있다.

내비게이션 데이터 이전의 경우, 교통분야에서는 특정한 알고리즘과 주어진 제약식을 통해 차량 통행 이동경로를 분석하였다. 이는 실제 차량이 주행한 경로가 아니므로 그 정확도나 신뢰성에 대해 다소 미흡한 부분이 있었다. 그러나 내비게이션 빅데이터를 활용한 경로형 데이터의 경우 실제

* 홍익대학교 도시공학과 교수

** 홍익대학교 도시계획과 박사수료, 한국교통연구원 국가교통빅데이터사업단

(Corresponding author, E-mail : kykim859@koti.re.kr)

이 논문은 2016학년도 홍익대학교 학술연구진흥비에 의하여 지원되었습니다.

차량이 이동한 경로를 제공하기 때문에 분석에 대한 정확성과 신뢰성 도모에 큰 기여를 할 수 있다.

또한, 앞서 언급한 것과 같이 빅데이터는 2차 가공자료로서의 역할이 중요하며, 내비게이션 데이터 또한 경로형 데이터를 제공할 뿐만 아니라 다양한 교통부문 지표를 산출할 수 있다. 내비게이션 경로형 데이터를 활용하여 전국단위 링크를 중심으로 교통량과 속도를 산출할 수 있으며, 이를 통하여 환경지표, 혼잡지표 등 다양한 지표를 산출한다.

본 연구에서는 최근 미세먼지 및 이산화탄소 등 환경지표에 대한 관심이 높아지고 있는 가운데 내비게이션 데이터를 활용하여 산출된 환경지표를 활용하여 도시공간구조와의 상관관계를 살펴보고자 한다.

최근 미세먼지 배출 원인과 관련하여 사회적 관심이 높아지고 있는 가운데 본 연구에서는 이러한 시대적 배경에 착안하여 도시공간구조와 교통부문과 어떠한 상관관계가 있는지에 대한 연구를 하고자 한다.

이를 위하여 내비게이션 데이터를 활용하여 미세먼지, 이산화탄소, 일산화탄소, 질소산화물, 휘발성 유기화합물 배출량을 읍면동 단위로 산출한다.

환경지표 데이터를 대상으로 서울시 기준 읍면동 주거지 비율 및 상업지 비율 가지고 공간가중회귀 모형 (Geographically Weighted Regression)을 통해 두 변수간의 상관관계를 파악하고자 한다.

1.2 연구의 범위 및 방법

본 연구에서는 5가지 환경지표를 산출하기 위하여, 2015년 기준연도 내비게이션 데이터를 활용하였으며, 기준연도 또한 2015년 기준으로 설정하였다. 공간적 범위의 경우 용도지역 대장이 존재하는 서울시를 기준으로 설정하였으며, 단위는 읍면동 단위로 설정하였다. 분석방법론으로써 공간가중회귀모형을 사용하여 교통부문 환경지표와 도시공간구조와의 관계를 살펴보고자 한다.

1.3 선행연구검토

도시공간구조와 교통, 그리고 이산화탄소 등 환경지표와 관련된 선행연구를 검토한 결과 우선 Seok & Lee(2013)은 도시권별 교통에너지 소비를 유형별로 구분하여 탄소배출량을 분석하였다. 이를 통하여 교통에너지부문 소비를 줄이고 효과적인 탄소저감을 위한 정책을 제시하였다.

또한 도시공간과 이산화탄소 배출량과의 관계에 대한 연구가 활발히 진행되고 있었다. 일반적인 도로의 형태인 연속류와 단속류를 승용차가 주행했을 경우, 승용차에서 배출되는 이산화탄소 배출량의 관계를 규명한 Lee & Oh(2011)은 속도의 가감속이 많은 단속류에서 연속류보다 이산화탄소 배출량이 높아짐을 분석하였다.

교통부문 환경지표와 에너지소비량과의 도시특성 관계분석에 대한 연구로써 Kim, Lee & Ahn(2009)은 도시 형태적 특성을 총량적 특성과 공간구조 특성으로 구분하여 에너지소비량과의 상관관계를 분석하였으며, 이를 통해 콤팩트 도시개발의 타당성을 주장하였다. Kim, Lee & Jung(2002)는 인구분산에 의한 수송 에너지 절감효과 분석을 중심으로 인구의 다극분산화에 의한 에너지 소비 감소효과를 검증하였다. Kim, Lee & Jung(2002)는 위 내용의 연구결과를 토대로 에너지 소비를 최소화 할 수 있는 국토·도시 공간구조 개편안을 도출하여 교통에너지 절약형 도시 공간구조를 제안하였다.

선행연구검토 결과, 교통부문 환경지표와 관련하여 도시공간구조와의 상관성을 살펴보는 연구는 다소 미흡한 것으로 나타났다. 또한 교통부문에서도 통행과 관련한 연구가 주로 이루어졌으며, 도시공간구조와의 연관성은 다소 부족한 것으로 나타났다. 본 연구는 빅데이터 시대를 맞이하여 내비게이션 데이터를 활용하여 교통부문 환경지표를 산출, 이와 관련하여 도시공간구조와 상관관계를 살펴보는 최초의 연구로써, 향후 도시계획 및 도시재생에 있어 정책적, 학술적 기여를 도모 하고자 한다.

2. 본 론

2.1 환경지표의 산출(Kim, Sohn & Kang, 2016)

1) 환경지표

내비게이션 데이터를 활용하여 환경지표를 산정하기 위하여 내비게이션 데이터의 Speed-Profile을 이용하여 링크별 원단위 환경지표를 산정하고, 산정된 원단위 환경지표와 내비게이션 데이터를 활용하여 전수화된 교통량을 이용하여 링크별 총량 환경지표를 산정하였다. 환경지표 분석에 사용된 배출 계수식은 국립환경과학원의 자료를 활용하였다.

Table 1. References used for estimating environmental indices

Division	Contents
Carbon Dioxide	The system for greenhouse gas and climate change response in transport sector(II) - development of bottom-up emissions coefficient for automobiles
Carbon Monoxide Nitrogen Oxide Volatile Organic Compounds	Manual for estimating the amount of air pollutant emissions at the national level (III)

환경지표 산정방법은 기 개발된 교통 환경지표와 마찬가지로 내비게이션의 개별 차량의 실제 속도자료를 통해 Speed-Profile을 구축하고, 각 오염물질의 배출 계수식을 이용하여 링크별 원단위 혼잡지표를 산출하였다. 산정된 원단위 혼잡지표와 추정된 전수화 교통량을 이용하여 링크별 총량지표를 산정하도록 하였다.

2) 산정방법의 개요

환경지표의 차종별 원단위 산정은 기본적으로 Network Occupancy Rate를 고려하여 각 차종별·연료별 배출 계수식을 산정하였다. 화물차와 버스의 경우 고속으로 주행하는 것이 한계가 있으므로, 내비게이션 데이터를 통해 산출된 혼잡 경계속도를 기준으로 혼잡교통류, 정상 교통류로 구분하여 배출량을 산출하였다.

환경지표 배출량 산정을 위한 방법은 기본적으로 IPCC (Intergovernmental Panel on Climate Change) 가이드 라인에서 제시하고 있으며, 배출계수도 국내에서 개발된 내용을 제외하고는 IPCC에서 제시된 것을 사용하고 있다. IPCC 가이드에서 배출량 산정을 위한 방법으로는 Tier1, Tier2, Tier3 등 총 3가지 방법론을 제시하고 있으며, 본 연구에서는 연료 소비량을 기본으로 하는 Tier1, Tier2 방법론을 배제하고 차량 주행거리를 기본으로 하는 Tier3 방법을 기반으로 배출량을 산정하였다.

3) 환경지표의 산출 : 차종별·연료별 Network Occupancy Rate 산출

환경지표의 배출계수는 차종별, 규모별, 연료별로 각각 개발되어 있으나, 실제로 해당 구간에 주행하는 차량의 구성비를 모두 파악하는 것은 한계가 있다. 기존의 연구들에서는 차종별로 대표차종을 선정하여 배출량을 산정하는 연구가 대부분이나, 이는 신뢰성 있는 배출량 산정의 결과로 판단하는데 한계가 있다.

따라서, 본 연구에서는 차종별 대표차종을 선정하지 않고, 국립환경과학원에서 개발된 차종별·연료별 배출계수를 이용하여 환경지표의 원단위를 산정하고자 한다. 또한, 기존의 연구에서 대표차종을 선정하여 배출량을 산정하는 방법을 적용하지 않고, 교통안전공단에서 조사하는 자동차 주행거리 자료를 바탕으로 도로상의 해당 차종이 얼마나 주행하였는지를 산정하여 각각의 비율을 산출하였다.

차종별 연간주행거리 산정은 관용, 자가용차량을 기준으로 산정하였으며, 각 차종별·규모별로 연간주행거리를 살펴보면 승용차는 중·대형차량이 80% 이상을 점유하고 있으며, 승합차는 소형이 90% 이상, 화물차도 소형이 87%로 가장 높은 점유율을 보이고 있는 것으로 나타났다.

도로이동 오염원의 사용연료는 경우, 무연휘발유, LPG, CNG 등이 사용되며, 교통안전공단의 2012년도 자동차 주행거리 실태분석 연구(Korea Transportation Safety Authority, 2013)자료에서 연료별·차종별 총 주행거리 자료를 이용하여 각 비율을 산정하였다.

4) 이산화탄소 배출량 산정

「수송부문 온실가스 기후변화대응 시스템 구축(III) - 자동차 온실가스 Bottom-Up」 배출계수 개발에서는 모든 차종에 대한 이산화탄소 배출 계수를 제시하고 있다. 제시되는

배출계수는 속도의 함수로 이루어져 있으며, 차종 및 연료에 따라 다른 배출계수가 적용되도록 설정되어 있다.

이산화탄소 배출량 산정의 정밀도 향상을 위하여 차종 및 연료별 도로상의 점유율을 고려하여 각각의 배출 계수식을 적용, 배출량 원단위를 산정하였다.

이산화탄소 배출량은 배출량 원단위와 해당링크에 대하여 추정된 교통량과 차종구성비, 연장으로 이루어진 함수식을 이용하였다.

$$Emissions(CO_2) = \sum_{a=1}^3 (\le ngth \times EU_a \times Volume_a)$$

$$EU_a = \sum_b (NOR \times EF_a) \dots\dots\dots (1)$$

Emissions : 배출량(ton), *≤ ngth* : 연장(km),

EU : 배출량 원단위, *Volume* : 교통량(대)

NOR : 네트워크 점유율(%),

EF : 배출계수식(g/대·km),

a : 차종(1=승용차, 2=버스, 3=화물차),

b : 연료별/차종별 구분

5) 대기오염물질(미세먼지, 일산화탄소, 휘발성 유기화합물, 질소산화물) 산정

미세먼지, 일산화탄소, 휘발성 유기화합물, 질소산화물 등 대기오염물질의 배출량 산정은 「국가 대기오염물질 산정방법 편람(III)」에 모든 차종에 대한 배출계수를 제시하고 있으며 대기오염물질 또한, 이산화탄소 배출량과 동일하게 산정방법을 선정하여 적용하였다.

$$Emissions(CO_2) = \sum_{a=1}^3 (\le ngth \times EU_a \times Volume_a)$$

$$EU_a = \sum_b (NOR \times EF_a) \dots\dots\dots (2)$$

Emissions : 배출량(ton), *≤ ngth* : 연장(km),

EU : 배출량 원단위, *Volume* : 교통량(대)

NOR : 네트워크 점유율(%),

EF : 배출계수식(g/대·km),

a : 차종(1=승용차, 2=버스, 3=화물차),

b : 연료별/차종별 구분

차량의 속도와 연료별 도로상 점유율을 고려하여 이산화탄소 및 대기오염물질 배출량 산정식은 다음과 같으며 차종별, 속도별, 도로 점유율, 연료별 배출 계수식을 분리하여 산정하였다. (Table 2, Table 3 참조)

Table 2. Carbon Dioxide(CO₂) & Particulate Matter(PM) Emission Factor

Type		Fuel	Network Occupancy Rate	Carbon Dioxide(CO ₂) Emission Factor		Particulate Matter(PM) Emission Factor	
				Speed	Calculation Formula	Speed	Calculation Formula
Vehicle	Light weight	Gasoline	6.8%	< 65.4km/h	$Y=887.12 \times V^{(-0.5703)}$		
				≥ 65.4km/h	$Y=0.9303 \times V+30.821$		
	Small	Gasoline	6.7%	< 665.4km/h	$Y=1313.7 \times V^{(-0.6)}$		
				≥ 65.4km/h	$Y=0.5447 \times V+78.746$		
		Diesel	3.2%	< 665.4km/h	$Y=1133.1 \times V^{(-0.587)}$		$Y=0.0420 \times V^{(-0.3420)}$
				≥ 65.4km/h	$Y=0.6175 \times V+62.478$		
	Mid-sized	Gasoline	27.3%	< 665.4km/h	$Y=1555.5 \times V^{(-0.578)}$		
				≥ 65.4km/h	$Y=0.0797 \times V+144.19$		
		Diesel	13.1%	< 65.4km/h	$Y=1818.1 \times V^{(-0.6643)}$		$Y=0.0396 \times V^{(-0.3420)}$
				≥ 65.4km/h	$Y=0.3184 \times V+95.66$		
		LPG	12.3%	< 65.4km/h	$Y=1539.4 \times V^{(-0.5748)}$		
				≥ 65.4km/h	$Y=0.5056 \times V+117.39$		
	Large	Gasoline	21.0%	< 65.4km/h	$Y=1970.1 \times V^{(-0.6187)}$		
				≥ 65.4km/h	$Y=0.1791 \times V+145.07$		
		LPG	9.5%	< 65.4km/h	$Y=1849.8 \times V^{(-0.6164)}$		
				≥ 65.4km/h	$Y=0.1348 \times V+159.9$		
Bus	Large (City Bus)	CNG	78.4%	-	$Y=5019.8 \times V^{(-0.5582)}$		
	Large (Intercity Bus)	Diesel	21.6%	< 65.4km/h	$Y=2676.7 \times V^{(-0.3344)}$	< 80.0km/h	$Y=0.2986 \times V^{(-0.5711)}$
				≥ 65.4km/h	$Y=1.3034 \times V+548.56$	≥ 80.0km/h	$Y=0.0001 \times V^{(1.2263)}$
Truck	Small	Diesel	87.0%	< 65.4km/h	$Y=1135.2 \times V^{(-0.4668)}$		$Y=0.1759 \times V^{(-0.5357)}$
				≥ 65.4km/h	$Y=2.2307 \times V+25.76$		
	Mid-sized	Diesel	9.1%	$Y=0.1029 \times V^2-14.937V+798.9$			$Y=0.2979 \times V^{(-0.4008)}$
	Large	Diesel	4.0%	$Y=6240.3 \times V^{(-0.3829)}$			$Y=2.6847 \times V^{(-0.6112)}$

Table 3. Carbon Monoxide(CO), Volatile Organic Compounds(VOC) & Nitrogen Oxide(NO_x) Emission Factor

Type		Fuel	Network Occupancy Rate	Carbon Monoxide(CO) Emission Factor		Volatile Organic Compounds(VOC) Emission Factor		Nitrogen Oxide(NO _x) Emission Factor	
				Speed	Calculation Formula	Speed	Calculation Formula	Speed	Calculation Formula
Vehicle	Light weight	Gasoline	6.8%	< 45.0km/h	$Y=4.9952 \times V^{(-0.8461)}$	-	$Y=0.2958 \times V^{(-0.7830)}$	-	$Y=0.4819 \times V^{(-0.9198)}$
				≥ 45.0km/h	$Y=-0.0001 \times V^2+0.0229 \times V-0.5701$				
	Small	Gasoline	6.7%	-	$Y=19.681 \times V^{(-1.0085)}$	-	$Y=2.4151 \times V^{(-1.3927)}$	-	$Y=1.2279 \times V^{(-0.7864)}$
		Diesel	3.2%	-	$Y=0.5775 \times V^{(-0.7524)}$	-	$Y=0.0825 \times V^{(-0.6848)}$	< 65.4km/h	$Y=1.1849 \times V^{(-0.5476)}$
								≥ 65.4km/h	$Y=-0.000002 \times V^2+0.0021 \times V-0.0110$
		Gasoline	27.3%	-	$Y=36.511 \times V^{(-1.2078)}$	-	$Y=3.6805 \times V^{(-1.5356)}$	-	$Y=1.3520 \times V^{(-0.7978)}$
	Mid-sized	Diesel	13.1%	-	$Y=0.5414 \times V^{(-0.7524)}$	-	$Y=0.0927 \times V^{(-0.6848)}$	< 65.4km/h	$Y=1.1281 \times V^{(-0.5476)}$
								≥ 65.4km/h	$Y=-0.000002 \times V^2+0.0020 \times V-0.0105$
		LPG	12.3%	-	$Y=73.022 \times V^{(-1.2078)}$	-	$Y=4.4166 \times V^{(-1.5356)}$	-	$Y=2.0280 \times V^{(-0.7978)}$
	Large	Gasoline	21.0%	-	$Y=36.511 \times V^{(-1.2078)}$	-	$Y=3.6805 \times V^{(-1.5356)}$	-	$Y=1.3520 \times V^{(-0.7978)}$
		LPG	9.5%	-	$Y=73.022 \times V^{(-1.2078)}$	-	$Y=4.4166 \times V^{(-1.5356)}$	-	$Y=2.0280 \times V^{(-0.7978)}$
Bus	Large (City Bus)	CNG	78.4%	-	$Y=18.235 \times V^{(-0.3767)}$	-	$Y=8.0544 \times \exp(-0.0174V)$	-	$Y=8.6972 \times \exp(-0.0130V)$
	Large (Intercity Bus)	Diesel	21.6%	-	$Y=8.3966 \times V^{(-0.8759)}$	-	$Y=1.2191 \times V^{(-0.5266)}$	-	$Y=40.9398 \times V^{(-0.5611)}$
Truck	Small	Diesel	87.0%	-	$Y=1.2211 \times V^{(-0.6083)}$	-	$Y=0.1224 \times V^{(-0.5684)}$	< 65.4km/h	$Y=2.0832 \times V^{(-0.6485)}$
								≥ 65.4km/h	$Y=-0.00003 \times V^2+0.0060 \times V-0.1339$
	Mid-sized	Diesel	9.1%	-	$Y=15.256 \times V^{(-0.7448)}$	-	$Y=2.0502 \times V^{(-0.6504)}$	-	$Y=15.001 \times V^{(-0.4528)}$
	Large	Diesel	4.0%	-	$Y=52.136 \times V^{(-0.8618)}$	-	$Y=3.7878 \times V^{(-0.5425)}$	-	$Y=107.5 \times V^{(-0.5679)}$

6) 산출결과

서울특별시를 대상으로 이산화탄소 및 대기오염물질 산출 결과를 살펴보면 다음과 같다. 우선, 일산화탄소의 경우, 강남구, 영등포구, 용산구, 중구 등 대표적인 도심지역에서 높은 산출결과를 나타내고 있었다. 또한, 이산화탄소와 미세먼지, 질소산화물은 강남구, 강서구, 영등포구, 강동구 지역에서 높은 산출량을 보이고 있었다. 휘발성 유기화합물은 강남 지역을 제외하고는 특별히 높은 산출지역은 없는 것으로 산출되었다. 공통적으로 서울특별시 내 상대적으로 높은 산지 지역을 포함하고 있는 서울 북부 지역, 도봉구, 강북구, 노원구에서는 5가지 환경지표가 낮게 산출된 것으로 나타났다. (Figure 1, Figure 2, Figure 3, Figure 4, Figure 5 참조)

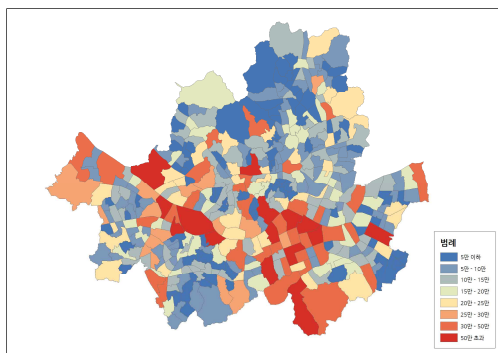


Figure 1. Carbon Monoxide Emission

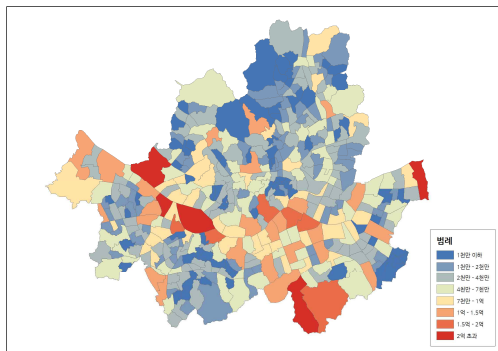


Figure 2. Carbon Dioxide Emission

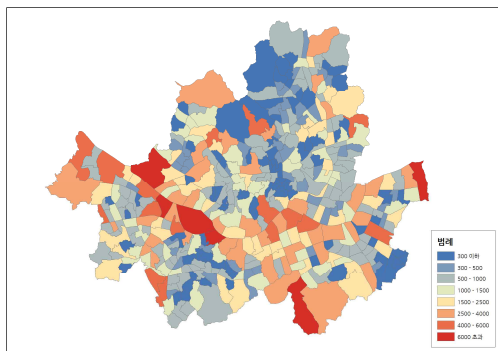


Figure 3. Particulate Matter Emission

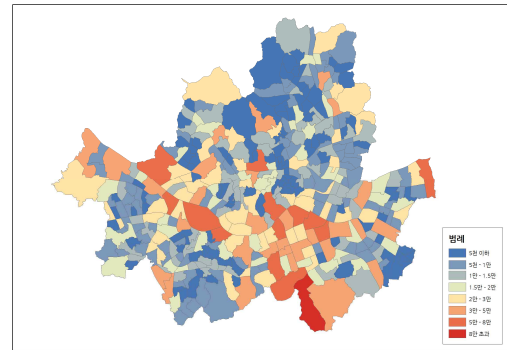


Figure 4. Volatile Organic Compounds Emission

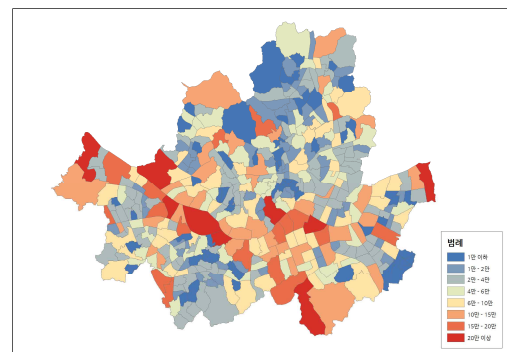


Figure 5. Nitrogen Oxide Emission

2.2 분석방법론

공간가중회귀모형은 공간상 서로 다른 위치에 존재하는 데이터들에 대하여 서로 다른 회귀모형을 적용함으로써 공간적 이질성을 탐색하는 방법이다. 일반적으로 선형회귀 분석은 분석 대상 전체 지역에 대한 전역적인 효과를 전제로 하지만, 공간가중회귀모형은 더 세분화된 공간적 단위의 국지적 효과를 전제로 한다.

이는 분석대상 지역 내에서 개별 관측치에 대한 회귀 분석 기준점을 선정하여 그 중심점으로부터 기준거리 내에 있는 관측들을 대상으로 회귀 분석을 실시한다. 따라서 공간가중회귀모형을 사용하여 에너지 소비에 따른 이산화탄소 배출 모형을 추정할 경우, 각 행정동 수만금의 회귀식이 도출된다. 아래 식은 공간가중회귀모형을 나타낸다. 공간가중회귀모형은 다중선형회귀모형에 좌표를 부여하여 공간분석이 가능하도록 해준다.

$$y_i = \sum_k \beta_k(u_i, v_i)x_{ik} + \epsilon_i \dots\dots\dots(3)$$

Equation (3)에서 y_i 는 행정동에서 종속변수 값을, x_{ik} 는 행정동 i에서 k번째 설명변수를 나타낸다. (u_i, v_i) 는 행정동 i의 중심점의 좌표를 의미한다. β_k 는 행정동의 k번째 독립변수의 계수를 의미한다. 그리고 ϵ_i 는 각 행정동별 오차항을

의미한다. 일반회귀모형과 달리 설명변수의 계수에 각 공간 별 위치 좌표가 붙기 때문에 각 행정동의 중심점에 대하여 독립변수 별 서로 다른 계수를 얻게 된다. 이때, $\beta(u_i, v_i)$ 의 추정식은 다음과 같다.

$$\hat{\beta}(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) Y \dots \dots \dots (4)$$

Equation (4)에서 $W(u_i, v_i)$ 는 다음과 같다.

$$W(u_i, v_i) = \begin{Bmatrix} W_{i,1} & 0 & \dots & 0 & 0 \\ 0 & W_{i,2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & W_{i,n-1} & 0 \\ 0 & 0 & \dots & 0 & W_{i,n} \end{Bmatrix} \dots \dots \dots (5)$$

Equation 4와 Equation 5에서 X는 (n×k) 설명변수 행렬, Y는 (n×1) 종속변수 벡터이고 $W(u_i, v_i)$ 는 (n×n) 공간가중 행렬이다. $W(u_i, v_i)$ 의 대각선의 값들은 좌표 (u_i, v_i) 를 갖는 행정동 I에 대한 전체 행정동의 개수인 관측치의 지리적 가중치를 의미한다. 즉, 행정동I와 다른 행정동 사이의 거리(가중치)가 고려되며 일반적인 공간가중 행렬과는 다른 대각 행렬 형태이다. 여기서 $W(u_i, v_i)$ 는 Kernel function으로 정의되며 다음과 같이 계산된다.

$$W_{ij} = \exp(-\gamma d_{ij}^2) \dots \dots \dots (6)$$

Equation (6)에서 d_{ij} 는 행정동 I의 중심점에서 행정동 j의 중심점까지의 거리를 의미한다. kernel function은 일반적으로 Gaussian 형태를 띄며 여러 회귀점을 기준으로 일정한 거리 내의 사례에 대해 가중치를 준다. γ 는 추정이 필요한 값으로서 Kernel band-width이며 γ 가 큰 경우 공간적 변이를 고려하지 않은 일반선형 회귀모형과 동일하게 된다. γ 를 추정하기 위한 방법론으로 Least square CV(Cross-Validation)와

Least AIC(Akaike Information Criterion)방법론이 있다. Least square CV 방법론은 γ 를 변화시키면서 관측치와 추정치의 차이의 제곱합의 오목 볼록 변곡점에서의 γ 를 결정하게 된다. AIC 방법론은 AIC값을 최소화하는 γ 를 선택하는 방법이다.

이처럼 공간가중회귀모형은 전역적 영향을 세분화하여 국지적인 영향을 분석함으로써, 각 공간 혹은 지역별 서로 다른 영향을 확인할 수 있다는데 의의가 있다.

2.3 분석결과

환경지표와 주거지역 및 상업지역 비율 간의 공간가중회귀 모형 결과는 다음과 같다. 일반적으로 공간가중회귀모형은 전통적인 회귀모형과 달리 국지적 변이를 파악하기 위하여 사용되는 분석기법으로, 수정된 결정계수 (R^2 Adjusted) 및 AIC(Akaike Information Criterion) 값을 살펴보고 모형의 적합도를 판단하게 된다. 수정된 결정계수는 일반적으로 값이 높을수록 설명력이 높다고 판단하며, AIC는 값이 낮을수록 추정된 모형이 현실을 보다 잘 설명하는 것으로 해석할 수 있다.

우선, 상업지역과 환경지표간의 모형 결과를 살펴보면 수정된 결정계수 (R^2 Adjusted)는 미세먼지 약 0.06, 이산화탄소 0.09, 일산화탄소 0.12, 질소산화물 0.08, 휘발성 유기화합물 0.11 정도로 모형적합도가 상대적으로 떨어지는 것으로 분석되었다.

주거지역 또한, 수정된 결정계수 (R^2 Adjusted)가 미세먼지 0.32, 이산화탄소 0.48, 일산화탄소 0.39, 질소산화물 0.47, 휘발성 유기화합물 0.37로 상업지역에 비해 상대적으로 높은 모형 적합도를 나타내고 있었다.

AIC(Akaike Information Criterion) 결과를 살펴보면, 상업지역의 경우 약 2,965~2,978 정도로 나타났으며, 주거지역의 경우 이보다 높은 약 3,546~3,593로 분석되었다. 모형의 적합도는 상대적으로 주거지역이 높은 반면, 현실성은 상업지역이 높은 것으로 도출 되었다.

Table 4. Commercial Area - Environmental Indices

VARNAME	VARIABLE				
	Commercial Area-Particulate Matter	Commercial Area-Carbon Dioxide	Commercial Area-Carbon Monoxide	Commercial Area-Nitrogen Oxide	Commercial Area-Volatile Organic Compounds
Bandwidth	4374.00391851000	3390.84460156000	2963.05185124000	3748.91906193000	3055.86294354000
ResidualSquares	24742.02920960000	23091.76073270000	21491.65569490000	23837.58681500000	22025.51705820000
EffectiveNumber	32.01914644560	48.28325056270	60.17460938040	40.87842262880	57.08785233760
Sigma	7.94484137326	7.83967837996	7.68578860270	7.88792618465	7.74786439444
AICc	2978.14725441000	2975.38016808000	2965.11609462000	2976.68743940000	2970.31257131000
R2	0.13693127020	0.19449708690	0.25031306750	0.16848066100	0.23169054286
R2Adjusted	0.06863289522	0.09312605107	0.12837976507	0.08192933739	0.11424328019
Dependent Field	0.00000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000
Explanatory Field	1.00000000000	1.00000000000	1.00000000000	1.00000000000	1.00000000000

Table 5. Residential Area – Environmental Indices

VARNAME	VARIABLE				
	Residential Area-Particulate Matter	Residential Area-Carbon Dioxide	Residential Area-Carbon Monoxide	Residential Area-Nitrogen Oxide	Residential Area-Volatile Organic Compounds
Bandwidth	2361.83412508000	2361.83412508000	2361.83412508000	2361.83412508000	2361.83412508000
ResidualSquares	84432.48373600000	80748.66334600000	75015.30946770000	82283.19543590000	77692.34317000000
EffectiveNumber	84.81572160540	86.24802563080	87.00482163560	85.76318561920	86.88936550600
Sigma	15.77745410490	15.46210872670	14.91980369620	15.59714597750	15.18108749860
AICc	3593.44211489000	3577.02284877000	3546.49863872000	3584.15166713000	3561.38192569000
R2	0.45772781006	0.48138734561	0.51821012075	0.47153173033	0.50101672712
R2Adjusted	0.32372709776	0.35049039101	0.39525212227	0.33909595713	0.37388529808
Dependent Field	0.00000000000	0.00000000000	0.00000000000	0.00000000000	0.00000000000
Explanatory Field	1.00000000000	1.00000000000	1.00000000000	1.00000000000	1.00000000000

전반적으로 주거지역 및 사업지역과 환경지표와의 연관성이 다소 떨어지는 것은 산출된 환경지표가 교통부문과 빅데이터를 이용하여 산출한 지표이기 때문에 공간적 상관성이 다소 떨어지는 것으로 사료된다. (Table 4, Table 5)

3. 결 론

본 연구는 제4차 산업혁명 시대를 맞이하여 증가하는 교통부문 빅데이터인 내비게이션 데이터를 활용하여 환경지표를 산출, 공간적 상관관계를 살펴보는 최초의 연구이다. 내비게이션 데이터를 활용하여 환경지표를 산출하기 위해, 국립환경과학원에서 제공하는 배출계수를 이용하여 이산화탄소, 미세먼지, 일산화탄소, 질소산화물, 휘발성 유기화합물 등 총 5가지의 환경지표를 산출하였다.

5가지 환경지표와 도시 공간적 상관성을 살펴보기 위하여 공간가중회귀모형을 활용하여 주거지역과 상업지역간의 모형을 도출하였다.

모형결과 교통부문 환경지표와 주거지역 및 상업지역 간의 상관성은 다소 낮은 것으로 나타났으며, 특히 상업지역의 상관성은 거의 없는 것으로 분석되었다. 이는 교통부문 데이터만을 이용하여 산출된 환경지표이기 때문에 공간적 상관성이 다소 떨어지는 것으로 판단된다.

따라서, 교통부문 환경지표는 공간적 상관성을 살펴보기 보다는 지역별 사회경제지표와의 관계를 살펴봐야 할 것으로 판단된다.

REFERENCES

- Seok, J.Y., Lee, H.Y. (2013). Analysis of CO₂ Emissions Using Transportation Energy Consumption in the City-regional Level. *Journal of the Korean Urban Geographical Society*, 16(3), pp.41~54.
- Lee, Y.S., Oh, H.U. (2011). Comparison of Vehicle Carbon Emissions in Expressway and National Highway.

International Journal of Highway Engineering, 13(3), pp.177~184.

- Kim, S.N., Lee, K.H., Ahn, K.H. (2009). The Effects of Compact City Characteristics on Transportation Energy Consumption and Air Quality. *Korea Planning Association*, 44(2), pp.233~248.
- Kim, S.H., Lee, M.H., Jung, J. (2002). A Study on National Spatial Structure for Energy Saving – Focusing on Transportation Energy Saving by Population Redistribution. *Korea Research Institute for Human for Human Settlements*, 92.
- The Korea Transport Institute (2015). Transport Congestion Map DataBase. Final report for KOTI National Transport DataBase, Book 14, pp.124~134.
- Korea Transportation Safety Authority (2013). Study on current status of vehicle kilometers traveled. Final report for Korea Transportation Safety Authority, pp.37~156.
- National Institute of Environment Research (2016). Manual for estimating air pollutant emissions(III). Final report for National Institute of Environment Research, pp.130~164.
- <https://ViewT.ktdb.go.kr>, Retrieved November 20, 2017
- Kim, D.H., Sohn, S.Y., Kang, K.Y. (2016). Spatial Pattern Analysis of CO₂ Emission in Seoul Metropolitan City Based on a Geographically Weighted Regression. *Journal of the Korea Institute of Industrial Engineers*, 42(2), pp.96~111.

■ (Received:2018.11.19, Revised:2018.11.29, Accepted:2018.11.29)