

IoT 도시빅데이터를 활용한 도로교통특성과 유해환경요인 간 영향관계 분석

Impact of Road Traffic Characteristics on Environmental Factors Using IoT Urban Big Data

박 병 훈* · 유 다 영** · 박 동 주*** · 홍 정 열****

* 주저자 : 서울시립대학교 교통공학과 학사과정
 ** 공저자 : 서울시립대학교 교통공학과 석사과정
 *** 공저자 : 서울시립대학교 교통공학과 교수
 **** 교신저자 : 계명대학교 교통공학과 조교수

Byeong hun Park* · Dayoung Yoo* · Dongjoo Park* · Jungyeol Hong**

* Dept. of Transportation Eng., University of Seoul
 ** Dept. of Transportation Eng., Keimyung University

† Corresponding author : Jungyeol Hong, jyhong9868@kmu.ac.kr

Vol.20 No.5(2021)

October, 2021
 pp.130~145

pISSN 1738-0774
 eISSN 2384-1729
<https://doi.org/10.12815/kits.2021.20.5.130>

Received 20 August 2021
 Revised 14 September 2021
 Accepted 21 September 2021

© 2021. The Korea Institute of
 Intelligent Transport Systems. All
 rights reserved.

요 약

스마트 서울 정책의 일환으로 도시 빅데이터 활용의 중요성이 부각되고 있으며, 미세먼지, 소음과 같이 교통과 관련된 도시환경 요소가 시민들의 삶의 질에 미치는 영향에 대한 사회적 관심이 증가하고 있다. 본 연구에서는 IoT 도시 빅데이터와 교통 빅데이터를 매칭하여 통합 DB를 구축하고, 이를 활용하여 특정 공간이 도로 영향권 내에 포함되는지 여부에 따라 미세먼지, 소음 피해에 유의한 차이가 있는지 분석하였다. 또한 시계열 클러스터링을 통하여 도로교통 특성 및 환경요인들이 유사한 특성을 가지는 공간 단위들을 군집화하였으며, 이 결과를 통하여 미세먼지 또는 초미세먼지 hot-spot, 소음 hot-spot 등 도시공간 단위의 환경위험 판리를 체계적으로 구축하는 기반을 마련하고자 하였다.

핵심어: IoT 도시 빅데이터, 도로 영향권, 미세먼지, 소음, 시계열 클러스터링

ABSTRACT

As part of the Smart Seoul policy, the importance of using big urban data is being highlighted. Furthermore interest in the impact of transportation-related urban environmental factors such as PM10 and noise on citizen's quality of life is steadily increasing. This study established the integrated DB by matching IoT big data with transportation data, including traffic volume and speed in the microscopic Spatio-temporal scope. This data analyzed the impact of a spatial unit in the road-effect zone on environmental risk level. In addition, spatial units with similar characteristics of road traffic and environmental factors were clustered. The results of this study can provide the basis for systematically establishing environmental risk management of urban spatial units such as PM10 or PM2.5 hot-spot and noise hot-spot.

Key words: IoT bigdata, Road-effect zone, PM10, Noise, Time-series clustering

I. 서 론

점차 고도화되는 도시 기능과 도로상의 차량들의 통행증가로 교통이 도시환경 및 시민들의 건강, 안전에 미치는 영향은 꾸준히 확대되고 있다. 특히 미세먼지, 소음과 같이 교통과 관련된 도시환경 요소들은 시민들의 건강 및 삶의 질과 긴밀한 관계에 있어 이러한 위험 요소들의 감소를 위한 방안 모색에 사회적 관심이 매우 높은 실정이다. 도시 내 환경오염을 유발하는 대표적인 요인 중 하나인 도로 재비산먼지는 자동차 배기가스, 타이어 마모, 브레이크 패드 마모로 인해 도로 위에 침적된 먼지들이 차량의 이동에 의해 대기 상으로 확산되면서 발생하는 것이며, Al, K, Ca 등의 지각 물질에 기인하는 자연적 성분 외에도 Cd, Pb, Cr 등 유해한 인위 성분을 포함하기 때문에 일반 먼지에 비하여 인체에 더욱 해롭다. 전국의 미세먼지(PM10) 배출량 중에서 비산먼지가 차지하는 비율은 약 50.2%이며, 이 중 도로 재비산먼지 배출량은 약 36.7%에 달한다(CAPSS, 2017). 특히 교통, 산업, 건설 활동이 빈번한 대도시일수록 인체 건강에 큰 피해를 주고 있다.¹⁾

미세먼지와 마찬가지로 도시 내 소음은 시민들의 건강에 악영향을 준다. 따라서 세계보건기구에서는 소음을 삶의 질에 영향을 미치는 주요 요인으로 지정하여 관리하고 있다. 국내 환경부에 따르면 40dB(A) 이상의 소음은 수면에 영향을 미치기 시작하고, 50dB(A) 이상부터는 불쾌감을 호소하는 사람이 발생하며 60dB(A) 이상이 되면 신체적 건강에 직접적인 영향을 미치기 시작한다.²⁾ 그러므로 건강한 도시를 조성하기 위해서는 미세먼지, 소음의 건강 위해 요인들과 도시교통 간의 영향 관계 및 위험 수준을 정확히 파악할 필요가 있으며, 이러한 분석 결과를 바탕으로 시민들에게 건강한 교통환경서비스를 제공할 수 있는 정책을 마련해야 한다.

스마트 서울 정책의 일환으로 구축된 서울시 도시 빅데이터는 2020년 4월부터 서울시 전역에 걸쳐 설치된 약 800개 이상의 IoT 센서를 통하여 매 2분 단위의 환경요인들에 대한 데이터를 수집·저장되고 있다. 따라서 본 연구에서는 스마트 서울 도시 데이터 센서(이하 S-DoT)로부터 수집된 초미세먼지(PM2.5), 미세먼지(PM10), 소음 등의 도시환경 데이터와 도로 교통량, 속도 등 교통 빅데이터를 통합하여 특정 도시 공간 단위가 교통 영향권 내 포함되었는지에 따라 미세먼지, 소음 등 유해 환경요인들이 직접적인 영향 관계가 있는지 분석하고, 도출된 분석결과를 바탕으로 미시적 범위의 시공간 측면의 도시교통관리전략을 제시하고자 하였다.

서울시 전역에 설치된 IoT 센서는 도시 현상을 매우 미시적으로 확인할 수 있는 최적의 시스템으로 빅데이터 수집을 위한 핵심 임무를 수행할 수 있다. 그러나 기존의 선행 연구들은 IoT 센서로부터 수집된 빅데이터 활용의 한계로 교통과 환경요인 간의 영향 관계를 미시적·복합적으로 평가·분석할 수 없었다. 또한, 실시간 IoT 도시환경 데이터와 도로교통데이터의 통합데이터베이스를 구축하여 시간 단위 및 공간 단위의 도시 건강 향상을 위한 교통 분야의 정책을 도출하는 연구는 본 연구가 처음이라는데 의의가 있다. 본 연구에서 교통-도시환경 통합데이터를 구축하고 미세먼지, 초미세먼지, 소음의 도시 공간별 변화패턴을 도로교통특성과 연계하여 도출하는 것은 향후 도로교통관리를 통하여 미세먼지 또는 초미세먼지 hot-spot, 소음 hot-spot 등 도시 공간 단위의 환경위험 관리를 체계적으로 할 수 있는 가이드라인이 될 수 있다.

II. 문학 검토

2000년대부터 국내외 연구들을 통하여 교통 요인은 미세먼지의 주요 원인이라는 것이 증명되어 오고 있

1) 도로 재비산먼지 관리시스템 (환경부, 한국환경공단)

2) <http://stat.me.go.kr/nesis/mesp/info/statPolicyHealth4.do> (환경부, 한국환경산업기술원)

으며, 미세먼지가 도시민의 건강에 미치는 영향을 정량화하는 방법론에 관한 연구가 꾸준히 시도되어 오고 있다(Penard-Morand et al., 2006; Vardoulakis et al., 2005; Ponce et al., 2005; Gulliver and Briggs, 2005; Hochadel et al., 2006; Kuehni et al., 2006; Lipfert et al., 2006; Gastaldi et al., 2014). 교통 현황과 미세먼지 농도와의 관계를 고찰한 Kim et al.(2018)의 연구에서는 11개의 대기 환경 감지 센서로부터 수집된 미세먼지 데이터와 교통 현황의 상관관계 분석을 위하여 대전 권역 전체 중 센서와 매칭을 할 수 있는 일부 구간(대덕대로와 계룡로 2개의 도로)을 공간적 범위로 정하여 분석을 수행하였으며, 그 결과 차량의 통행속도가 높아질수록 미세먼지 농도가 증가하는 추세를 보이고, 혼잡도가 높을수록 미세먼지 농도가 감소한다는 사실을 선형 상관관계 분석 방법론을 통해 제시하였다. Zhang et al.(2018)의 연구에서는 485개의 RFID 및 카메라 센서로 수집된 시간 대별 교통량 및 속도 데이터, 가솔린 차량의 배출 계수(CO₂ 등 주요 대기오염원)의 시공간별 변화 추이를 시각화하여 분석한 결과 가솔린 차량의 배출 계수는 교통변화와 유사한 패턴을 보이고 있으며, 도심 지역에 주로 집중적으로 분포하고 있는 것을 도출하였다. 또한 Jeong et al.(2019)의 연구에서는 주요 간선도로와 고속도로의 15m 이내에 위치한 도심 및 고속도로 현장에서 PM_{2.5}를 측정하여 비교·분석한 결과 고속도로 인근의 미세먼지 배출 기여도는 도심 인근보다 2~3배가량 높은 수준을 보이는 것을 발견하였다. Wu et al.(2020)은 중심 상업 지구 내 2개의 모니터링 지점에서 대기 환경 지표를 분석하였으며, 그 결과 도로 인근 지점에서 대기 환경이 더욱 심각한 수준인 것을 발견하였다. 이처럼 기존 연구들은 미세먼지 발생량이 교통량, 차량 통행속도 등 도시 내 교통 요인들의 특징과 높은 상관관계를 가진다는 것을 도출하였다.

교통시설로 인해 발생하는 소음 역시 사람들의 신체적, 정신적 건강에 어떠한 영향을 가지는지 도출하기 위하여 국내의 연구자들은 교통 소음에 의한 건강 위해성을 평가할 수 있는 방법들을 제시해 왔다. 특히 교통 소음은 시민들의 수면 방해, 심장질환 발병과 연관성이 있는 것으로 밝혀져 정부 차원에서 교통 소음 저감 대책을 위한 다양한 정책 및 관리방안을 제시하고 있다(Korea Environment Institute, 2007). Zuo et al.(2014)의 연구에서는 캐나다 토론토 지역 내 554개 지점을 대상으로 30분 단위의 교통 소음을 측정하였으며 교통량과 소음이 높은 상관관계(피어슨 상관계수가 0.79 이상)를 보이는 지점 중 무작위로 62개 지점을 분석하였다. 그 결과 교통량과 소음은 시간적(Temporal) 영향보다 지리적 특성과 같은 공간적(Spatial) 영향이 더 큰 것으로 나타났으며, 교통량이 많고, 간선도로이며, 공업시설이 많은 지역은 소음과 연관성이 매우 높다는 결과를 도출하였다. Welch et al.(2013)의 연구에서는 교통 소음이 직접 또는 간접적으로 건강에 미치는 영향을 분석하기 위하여 고속도로에서 50m 이내 거주하는 사람들과 고속도로 영향권에서 먼 곳에 거주하는 사람들에게 소음 민감도에 대한 설문을 실시하였다. 그 결과 고속도로로부터 50m 이내에 거주하는 사람들은 물리적, 심리적, 사회적, 환경적 측면의 건강 관련 삶의 질이 낮은 것으로 나타났으며 소음으로 인한 건강 문제를 야기하고 있는 것을 발견하였다. Can and Aumond(2018)의 연구에서는 차량 속도에 따라 소음 수준을 추정하였으며, Makarewicz and Gałuszka(2011)는 교통체증은 차량의 속도를 감소시킴으로써 소음 수준을 낮춘다는 사실을 speed-flow 다이어그램을 통해 제시하였다. 이외에도 도로 교통량과 소음 간 연관 관계를 도출하고, 효과적으로 소음 정도를 예측할 수 있는 방법론 연구 또한 활발하게 진행되어 왔다(Ali et al., 2019; Nourani et al., 2020). 이처럼 교통시설과 소음, 소음과 건강 간의 유의한 연관성은 기존 연구들을 통해 알 수 있으며, 교통 소음의 수준을 미시적 도시 공간 단위별로 추정 가능할 경우 시민들의 건강에 위협한 고위험 수준의 도시 공간들에 대하여 사전에 교통관리 전략 및 정책 마련이 가능할 수 있다는 것을 알 수 있다.

<Table 1> Contents of Literature Review

Literature	Author	Title	Summary	
			Objective	Methodology
Traffic & Atmospheric Environment	Seo and Kim(2013)	Estimation of greenhouse gas emissions from road traffic: A case study in Korea	Finding the principal cause of the main vehicle emissions of highways	Simple Comparison Analysis by Road type
	Kim et al.(2018)	Correlation Analysis between Vehicular Traffic and PM Using Sensor Big Data	Analyzing the correlation between PM and vehicular traffic in the Daejeon area	Derived the correlation between PM and vehicular traffic(speed, occupancy, congestion) by using Linear Regression
	Zhang et al.(2018)	Fine-grained vehicle emission management using intelligent transportation system data	Quantifying the traffic and emission by vehicle type by region and used in vehicle operation restriction policies	Visualization the heterogeneity of temporal and spatial emission distribution by region
	Jeong et al.(2019)	Temporal and spatial variability of traffic-related PM _{2.5} sources: Comparison of exhaust and non-exhaust emissions	Comparison between tailpipe and Non-tailpipe emissions factor by region(downtown, highway)	Derived the Spearman correlation coefficients
	Wu et al.(2020)	Development of the Real-time On-road Emission (ROE v1.0) model for street-scale air quality modeling based on dynamic traffic big data	Analyzing the real-time street level traffic emissions and performing air quality assessment for traffic policy suggestion	Build the integrated database about Traffic&Air Quality for comparison analysis with regional type
Traffic & Noise Environment	Makarewicz and Gałuszka (2011)	Road traffic noise prediction based on speed-flow diagram	Calculation of the annual average sound level of the road traffic noise with traffic flow	Derived the formula between speed-flow diagram and annual average sound level
	Welch et al.(2013)	Road traffic noise and health-related quality of life: A cross-sectional study	Comparison between Motorway and Non-motorway dweller's health-related factors	Capitalize on ANOVA and Pearson's r for exploring the association between Health index and road effective
	Zuo et al.(2014)	Temporal and spatial variability of traffic-related noise in the City of Toronto, Canada	Analyzing the most effective variability of traffic-related noise	Derived the Pearson correlation coefficient of noise level between the variability
	Can and Aumond(2018)	Estimation of road traffic noise emissions: The influence of speed and acceleration	Comparing four noise emission models with road traffic representation (speed & acceleration)	Derived RMSE for each models by Road segment length and speed
	Ali et al.(2019)	Developing machine learning models to predict roadway traffic noise:An opportunity to escape conventional techniques	Developing and comparing the performance of four different ML models to predict roadway traffic noise models	Regression Tree(Complex, Medium, Simple), SVM(Linear, Quadratic, Gaussian), Ensembles(Boosted, Bagged), ANN
	Nourani et al.(2020)	Artificial intelligence based ensemble model for prediction of vehicular traffic noise	Performance Comparison between four kinds of Ensemble method	FFNN(feed-forward neural network), ANFIS, SVR(support vector regression), MLR(multi linear regression)

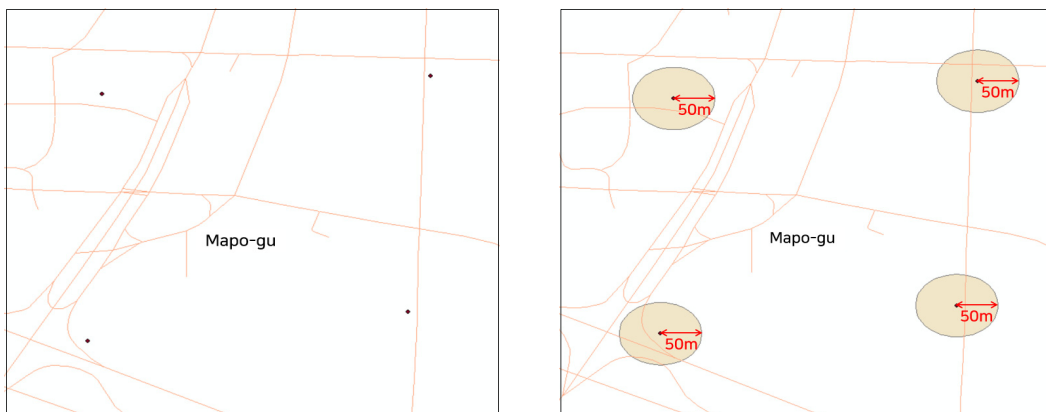
Ⅲ. 데이터 수집

1. 연구의 시공간 범위

S-DoT 센서는 서울시 행정동(2019년 기준 424개)별로 약 한 개 이상 지정·설치되었으며, 주거공간, 상업공간, 개방공간, 정책수요공간으로 설치 용도를 구분하여 서울시 전역에 걸쳐 약 800개 이상이 설치되어 있다. 본 연구에서는 특정 도시공간이 교통 영향권 내에 포함되는지 여부에 따라 미세먼지, 소음 등 도시환경요인들의 위험수준 변화 패턴을 분석하기 위하여 서울시 전역에 분포하고 있는 S-DoT 센서로부터 매 2분 단위로 수집된 도시환경요인들과 공간범위 내 교통량, 통행속도 간 관계를 분석하였다.

본 연구에서 분석 공간범위의 유형을 구분하는 주요 기준은 도로 영향권 여부이다. Britter&Hanna의 연구 등 기존 연구들에서는 도시 내 대기환경모델링을 위한 공간적 범위를 Street Scale(~100m), Neighborhood Scale(~1km), City Scale(~10km), Regional Scale(~100km)로 구분하였다(Britter and Hanna, 2003; Wu et al., 2020). 그러나 Hagler et al.(2009)는 도로변 비산먼지의 발생 지점으로부터 거리가 멀어질수록 10m당 평균 8.5%씩 영향력이 감소하며, Barzyk et al.(2009)는 도로에서 발생하는 VOCs(Benzene, Toluene, Ethylbenzene) 등 독성 물질들의 영향 범위는 최대 반경 250m, 직선 거리 100m까지 유효하다고 주장하였다. 이 외에 대부분의 연구들에서도 도로변을 기준으로 분석 반경을 작게 할수록 반경 안에 있는 대기환경 요인들로부터 일관성 있는 분석 결과를 보여주는 것으로 나타났다. 또한 Kang et al.(2008)은 서울시 주거지 유형 및 유형별 특성 분석을 위한 기본 단위로 12m 이상의 도로 등으로 분리되는 구간을 하나의 블록으로 지정하였고, 서울시의 블록 사이즈 중 소형 블록은 35~40m * 80m의 가구표준도를 기준으로 하였다(Sohn and Shin, 2003).

이처럼 도로 영향권과 비영향권을 구분하기 위한 적절한 공간적 범위를 설정하기 위해서는 S-DoT 센서를 기준으로 수집된 주요 도시환경 요인들의 관측 결과가 도로로부터 어느 정도까지 이격되어 있을 때 도로의 영향을 받는다고 판단할 수 있는지에 대한 적절한 판단 근거가 필요하다. 따라서 앞서 살펴본 기존 문헌들의 고찰 결과를 바탕으로 분석 반경을 최소화하면서 일반적인 서울시 도로 블록별 사이즈를 함께 반영한 반경 범위를 비교·분석한 결과, 각각의 S-DoT 센서로부터 50m 반경을 기준으로 하여 도로 영향권과 비영향권을 구분했을 때 분석 범위가 중복되지 않고, 보다 정확한 대기환경 요인들의 영향을 반영할 수 있는 데이터를 수집할 수 있었으며, 각 권역별 S-DoT 센서의 분포 비율(360:387)도 가장 이상적으로 나타났다.



<Fig. 1> Example of Radius Setting for each sensor

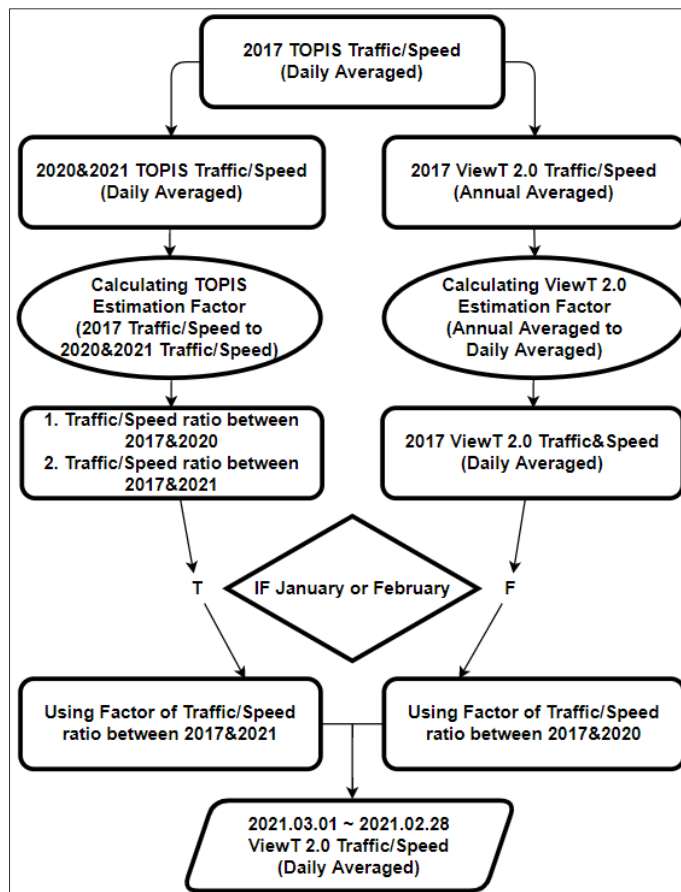
2. 데이터 수집 및 전처리

1) 도로 교통량 및 속도 데이터

국가교통DB에서 제공하는 2017년 서울시 링크별 교통량 및 속도 데이터를 서울시에서 제공하는 TOPIS 교통량과 속도 데이터를 활용하여 2020년 3월 1일부터 2021년 2월 28일까지의 서울시 링크별 추정 교통량 및 속도 데이터를 보정·구축하였다. 따라서 3월부터 12월까지는 2017년과 2020년 TOPIS 자료를 활용하고, 1월과 2월의 경우 2017년과 2021년 TOPIS 자료를 활용하여 2017년 ViewT2.0 자료를 2020년으로 보정하였다. 또한 일평균으로 구축되어 있는 TOPIS 자료를 활용하여 연평균으로 구축되어 있는 ViewT2.0 자료를 보정하기 위해서 2017년 일평균 TOPIS 교통량/속도의 월별/일별 분포 비율로 2017년 연평균 ViewT2.0 서울시 링크별 교통량/속도를 일평균으로 보정하는 과정이 필요하다.

추정된 도로교통현황과 도시환경 데이터의 통합DB 구축을 위해 ArcGIS 상에서 S-DoT 센서 좌표를 기준으로 교통망 GIS DB(국가교통DB)에서 추출한 서울시 도로 링크와 매칭(One-to-

Many)한 후 도로 영향권과 비영향권으로 분류하여 통합하였다. <Fig. 2>은 2020년 서울시 링크별 교통량 및 속도 추정 과정을 보여준다.



<Fig. 2> Flowchart of Traffic & Speed Estimation Process

2) 도시환경요인 데이터

매 2분 단위로 수집되는 초미세먼지(PM2.5), 미세먼지(PM10), 소음(Noise)의 S-DoT 도시환경 데이터는 시간별 도로교통 현황데이터와 통합하기 위하여 1시간 단위로 집계하였다. 분석의 정확성을 높이기 위하여 주요 도시환경요인별 데이터를 통계적으로 검증하였으며, 극단적 이상값으로 판단되는 이상치는 전처리하였다. 따라서 주요 도시환경요인별 이상치 전처리 기준은 데이터 분포의 통계치를 기준으로 초미세먼지(PM2.5)와 미세먼지(PM10) $0\mu g/m^3 \sim 500\mu g/m^3$, 소음(Noise) $0dB$ 이상, 기온(Temperature) $50^\circ C$ 이하로 설정하였으며 이 외의 값들은 이상치로 간주하고 분석 대상에서 제거하였다.

각각의 도시환경 요인들은 서울시에서 제공하는 통합대기환경지수와 환경부 환경통계포털에서 제공하는 소음지수를 기준으로 <Table 2>에서 보여주는 바와 같이 Class1 (Very Unhealthy) ~ Class4 (Good)까지 카테고리화 변수로 변환한 뒤 시각화하였다. 이는 도시환경 요인별로 상이한 측정 범위와 관측 단위를 단일 기준으로 통일하여 이후 영향 관계 및 Hot-Spot 비교·분석을 용이하게 하기 위함이다.

<Table 2> Atmospheric Environment & Noise Pollution Index

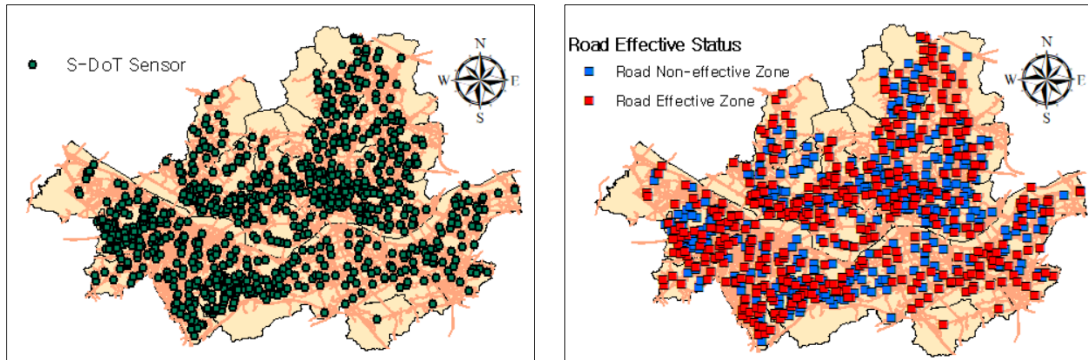
Legend	Range			Class
	PM2.5($\mu g/m^3$)	PM10($\mu g/m^3$)	Noise(dB)	
Good	0~15	0~30	0~39	4
Moderate	16~35	31~80	40~49	3
Unhealthy	36~75	81~150	50~59	2
Very Unhealthy	76~	151~	60~	1

3) 토지이용현황 데이터

도로환경요인들은 도로 교통량 및 속도 뿐 만 아니라 공간적 특성에 의해서도 변화패턴이 상이할 수 있으므로 토지이용현황 정보를 수집하였다. 국토지리정보원에서 제공하는 토지이용현황도를 활용하여 38가지 토지이용현황 UCB Code를 분석에 용이하도록 다음과 같이 5가지로 재분류하였다. Industry Facility(이하 공업시설)는 처리장, 매립지 등 공업시설과 발전시설을 포함하며 Public Facility(이하 공공시설)는 교육 및 군사시설을 비롯한 공공용지를 포함하고 Transportation Facility(이하 교통시설)는 도로, 철로 등 주변 지역과 공항, 항만 등의 주요 교통 시설로 이용되는 토지를 포함한다. 그 외 Residential Facility(이하 주거시설)는 일반 및 고층 주택지와 상업 등 시민들이 머무르고 주로 활동하는 지역이며 Green Area(녹지구역)는 초지 및 임목지와 골프장, 유원지, 하천, 댐 등을 포함한다.

4) 기초 현황 분석

본 연구에서 분석대상으로 선정한 800개의 S-DoT 센서 중 2020년 3월부터 2021년 2월까지 전체 분석 기간 중 데이터 누락 없으며 KTDB 도로 Link ID 매칭이 가능한 747개 S-DoT 센서를 재추출하였으며 이 센서들을 대상으로 통합DB를 구축하였다. 통합DB 분석 결과 <Table 3>과 같이 초미세먼지(PM2.5)는 최소 $0\mu g/m^3$ 에서 최대 $454.53\mu g/m^3$ 의 범위 내에 분포하고 있으며, 소음은 $0dB$ 에서 최대 $70dB$ 까지 분포하는 것으로 나타났다. 분석범위 내 평균교통량 및 평균속도는 각각 1,897대/시간, $29.08km/h$ 이다. <Fig. 3>에서 나타나는 바와 같이 S-DoT 센서들을 기준으로 50m 반경 내에 도로교통현황(속도, 교통량)에 대한 속성값이 존재하는 도로들과 매칭되는 도로 영향권 S-DoT 센서는 360개, 그 외 도로 비영향권 S-DoT 센서는 387개이다.



<Fig. 3> Smart Seoul Data of Things(S-DoT) Sensor Distribution/Road Effective Status of Seoul

<Table 3> Summary of Integrated DataBase

Features	Summary Statistics		
	Min.	Mean	Max.
PM2.5 ($\mu\text{g}/\text{m}^3$)	0.00	19.02	454.53
PM10 ($\mu\text{g}/\text{m}^3$)	0.00	25.31	499.00
Noise (dB)	4.16	48.53	70.00
Traffic (veh/h)	0	1,897	35,870
Speed (km/h)	4.25	29.08	133.76
Time periods	2020-03-01 00:00:00 ~ 2021-02-28 23:00:00		
Land use	Industry Facility	5 S-DoT Sensor	
	Green Area	31 S-DoT Sensor	
	Public Facility	45 S-DoT Sensor	
	Transportation Facility	224 S-DoT Sensor	
	Residential Facility	442 S-DoT Sensor	
Road Effective Status	Road Effective Zone(1)	360 S-DoT Sensors	
	Road non-effective Zone(0)	387 S-DoT Sensors	

IV. 시계열 군집분석(Time-Series Clustering Analysis)

본 연구에서는 초미세먼지(PM2.5), 미세먼지(PM10), 소음(Noise) 수준을 분석할 때 시공간별로 유사하게 변화하는 패턴을 고려하기 위하여 시계열 군집분석 기법을 활용하였다. 각 IoT 센서가 설치된 위치를 기반으로 설정된 공간단위를 시계열적 특성을 기반으로 군집화하였고, 군집별로 초미세먼지(PM2.5), 미세먼지(PM10), 소음(Noise) 수준 패턴의 유사성을 도출하였다.

군집분석에서 유사성은 거리(distance)기반으로 측정할 수 있다. 군집분석 연구에서 많이 사용되고 있는 유클리디안 거리(Euclidean distance)는 시계열 변수 값들을 같은 시점을 기준(time-rigid)으로 측정하므로 시간 유연성(time-flexible)이 반영되지 않는다. 따라서 변수의 관측치가 시간의 흐름에 따라 변화하는 시계열 데이

터를 속성이 비슷한 동적 개체들끼리 군집화하기 위해서는 유클리디안 거리보다 시간 지연으로 인한 맵핑(mapping) 오류와 이상치에 강한 동적타임워핑(DTW: Dynamic Time Warping) 거리가 더욱 적합하다. 특히 도시환경데이터와 같이 계절성, 추세성을 가지고 있는 데이터일수록, 패턴에 따라 관측치들을 묶을 수 있는 DTW거리 알고리즘이 적합하므로 시계열 클러스터링의 효과가 더욱 유의미하게 나타날 수 있다. 길이가 m 인 시계열 데이터 $\vec{x} = (x_1, \dots, x_m)$ 와 $\vec{y} = (y_1, \dots, y_m)$ 에서 벡터 간 원소들이 정렬되어 있을 때 유클리디안 거리는 식(1)과 같이 정의된다.

$$Euclidean Distance(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \dots\dots\dots (1)$$

시계열 데이터는 대체적으로 발생 및 측정 시점에 따라 나타나는 시간 지연과 시중점의 차이로 인해 벡터 공간 상 정렬되지 않는 경우가 많기 때문에 이러한 한계점을 보완한 DTW는 식(2)와 같이 정의된다.

$$DWT: d(x_i, y_j) = \sqrt{(x_i - y_j)^2} \dots\dots\dots (2)$$

$\vec{x} = (x_1, \dots, x_m)$ 와 $\vec{y} = (y_1, \dots, y_m)$ 를 바탕으로 $m \times m$ 개의 행렬집합 M 을 생성하고 각 행렬의 원소 (i, j) 는 x_i 와 y_j 의 거리 $d(x_i, y_j)$ 를 계산한다. 생성된 행렬 집합 M 에서 \vec{x} 부터 \vec{y} 까지 mapping 되는 경로의 집합이 $W = \{w_1, w_2, \dots, w_r\}$ with $r \geq m$ 을 만족하는 W 라고 정의할 때, W 의 r 번째 값 $r(i, j)$ 는 식(3)과 같이 정의된다.

$$r(i, j) = ED(i, j) + \min\{r(i-1, j-1), r(i-1, j), r(i, j-1)\} \dots\dots\dots (3)$$

따라서 DWT 거리 알고리즘은 시간차에 제약받지 않고 두 개의 시계열 데이터에 존재하는 서로 다른 원소들 간 가장 유사도가 높은(minimum distance) 쌍을 찾아 매칭한다. DWT 거리는 식(4)와 같다.

$$DTW(\vec{x}, \vec{y}) = \min \sqrt{\sum_{i=1}^r w_i} \dots\dots\dots (4)$$

시계열 군집분석의 성능을 최적화하기 위하여 초미세먼지(PM2.5), 미세먼지(PM10), 소음(Noise)의 도시환경요인 변수들을 정규화하여 0~1 사이의 값으로 변환한 뒤 비교·분석하였으며 분석의 성능평가를 위하여 한 군집 안의 데이터들이 다른 군집과 비교하여 얼마나 유사한지를 판단하는 SilhouetteI Index를 적용하였다.

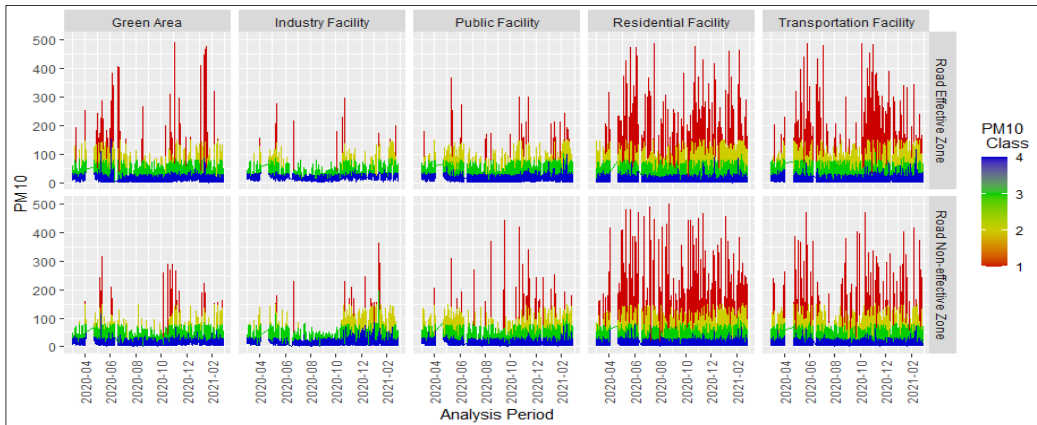
V. 결 과

1. 도시환경지표와 교통 및 토지이용 간 영향 관계

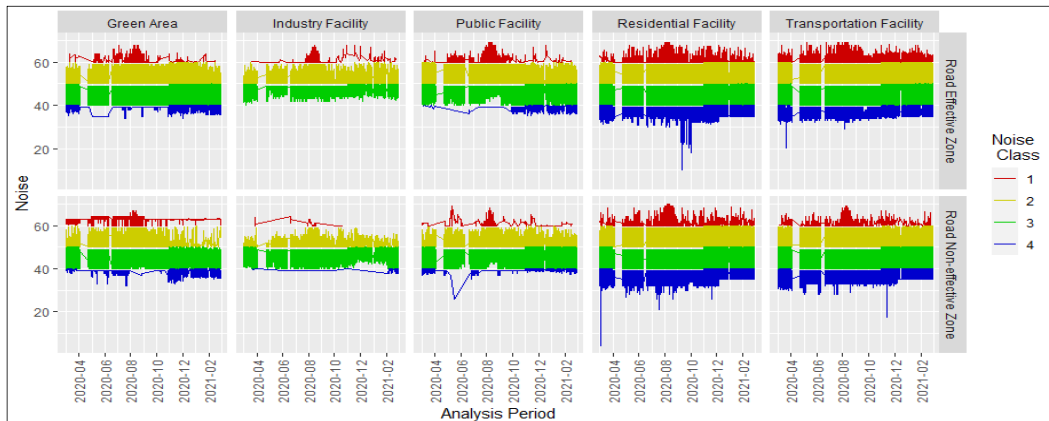
본 연구에서 구축한 IoT 도시·교통 통합데이터베이스를 기반으로 시간대별 토지이용현황과 도로 영향권 유무를 구분하여 동 시간대 주요 도시환경 요인 변화 추이를 비교·분석하였다. 그 결과 초미세먼지(PM2.5),

미세먼지(PM10), 소음(Noise) 3가지의 도시환경요인들이 공통적으로 주거시설과 교통시설 인근에서 Class1(Very Unhealthy) 수준의 시간대가 가장 빈번하게 나타났고, 교통시설 인근 지역의 Class1(Very Unhealthy) 수준의 시간대는 도로 비영향권보다 도로 영향권에서 더욱 빈번하게 발생하는 것으로 분석되었다. 특히 미세먼지와 초미세먼지는 녹지구역에서 도로 영향권 내 위치하는지 여부에 따라 차이가 가장 극명하게 드러나는 것으로 분석되었다. 녹지구역 내 도로영향권 지역은 봄과 가을의 초미세먼지 및 미세먼지 수치가 도로 비영향권 지역보다 높은 수준으로 증가하는 것을 발견하였다.

시계열 소음 패턴은 5가지 토지이용 유형에서 공통적으로 도로 영향권이 도로 비영향권보다 Class1(Very Unhealthy)이 훨씬 빈번하게 발생하였고, 특히 공업시설 인근 지역은 Class3 (Moderate) 수준 또는 그 이상의 높은 소음이 지속적으로 발생하고 있는 것으로 분석되었다. 공업시설 이외의 지역들은 시계열 흐름에 따라 Class4 (Good)에서 Class1(Very Unhealthy) 사이를 폭넓게 변동하는 것을 알 수 있다. <Fig. 4>와 <Fig. 5>는 미세먼지와 소음의 시계열 변화패턴 분류한 수준별로 시각화한 것을 보여준다.



<Fig. 4> Comparison of PM10 Trend during Analysis Period by Landuse and Road Effect



<Fig. 5> Comparison of Noise Trend during Analysis Period by Landuse and Road Effect

2. 시공간별 도시환경요인의 위험수준 Hot-Spot 분석

S-DoT 센서별로 유사한 도시환경-도로교통현황 간 특성을 가진 그룹별로 비교·분석하기 위해 Silhouette Index를 기반으로 우수한 성능을 보이는 최적 군집을 도출하였다. 도로 영향권 내의 월평균·일평균 클러스터링을 도출한 결과 <Fig. 6>과 같이 Centroid가 3개일 때 뛰어난 군집성능을 보였으며 도로 비영향권 관측치들 또한 월평균·일평균 클러스터가 <Fig. 7>과 같이 Centroid가 3개일 때 뛰어난 Silhouette Index를 보여주었다.

Silhouette Index $S_{x(i)}$ 는 클러스터링의 유효성을 검증하기 위해 사용되는 지표 중 하나로 본 연구에서는 타 지표 대비 클러스터링 성능을 가장 직관적으로 평가하기에 적합하다고 판단되며 아래와 같이 정의된다.

$$S_{x(i)} = \frac{b_{x(i)} - a_{x(i)}}{\max(a_{x(i)}, b_{x(i)})} \dots\dots\dots (5)$$

where $x(i)$ is data point in the cluster($i=1,2,3, \dots, n$),

$a_{x(i)}$ is average distance between $x(i)$ and every data point in the same cluster,

$b_{x(i)}$ is minimum average distance between $x(i)$ and every data point in other cluster

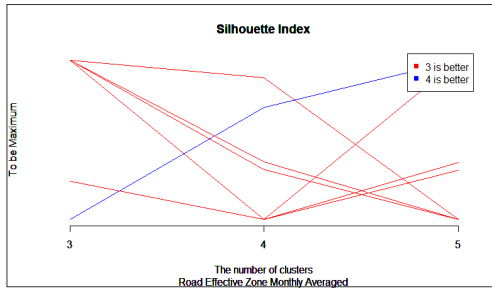
$$S_k = \frac{1}{n} \sum_{i=1}^n S_{x(i)} \dots\dots\dots (6)$$

where k is number of clusters,

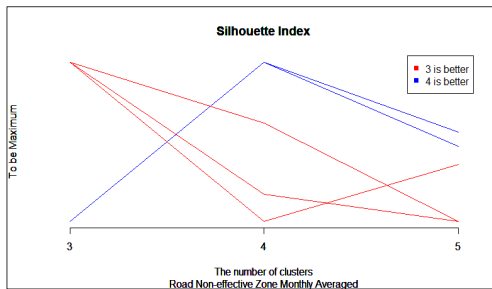
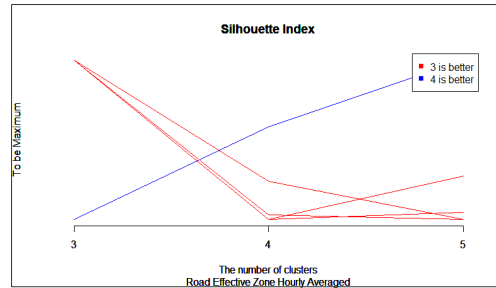
n is number of data points in the same cluster

$$S_{average} = \frac{1}{m} \sum_{k=1}^m S_k \dots\dots\dots (7)$$

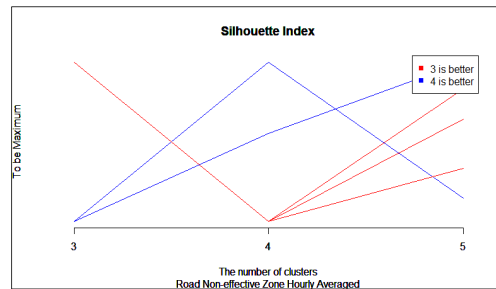
where m is number of all clusters



<Fig. 6> Best number of clusters in Road Effective Zone



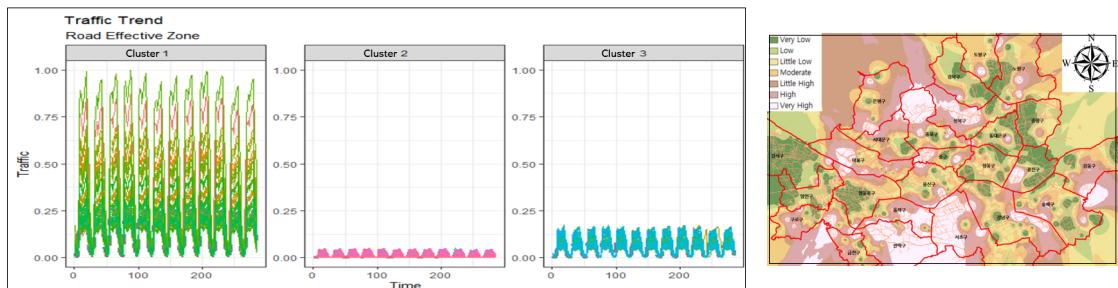
<Fig. 7> Best number of clusters in Road Non-effective Zone



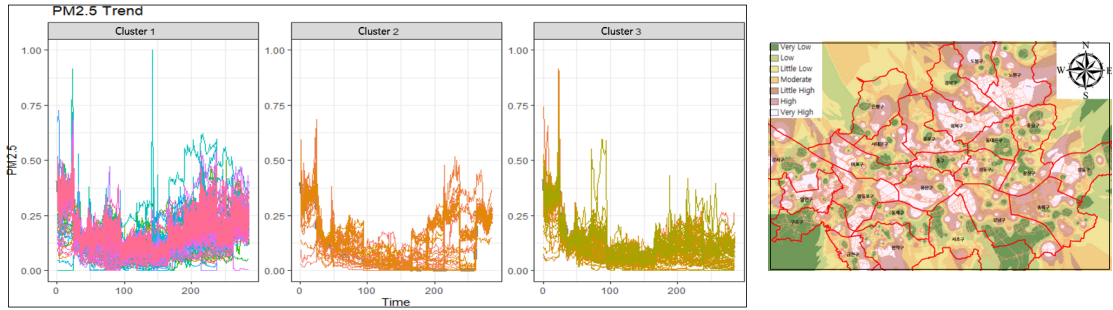
교통량 시계열 군집분석 결과(<Fig. 8>) 성북구, 서초구, 관악구 등이 Cluster 1로 군집되었다. Cluster 1은 1년 동안 꾸준히 교통량이 많은 지역으로 경부고속도로와 내분순환로가 통과되는 지역들을 포함한다. 이 지역들은 다른 Cluster들과 비교하였을 때 많은 통행량이 관측되는 지역으로 소음 측면의 교통관리가 중요한 지역이라고 할 수 있다. 서대문구, 광진구, 강북구 등이 포함되는 Cluster 2는 연간 교통량이 가장 적게 나타난 지역으로 주로 관측지점 인근에 산림 지역이 분포하고 있다. 주거시설 비율이 25개 자치구 중 각각 2위, 3위인 서대문구와 광진구는 Cluster 2로 분류되었다. Cluster 3에는 금천구, 강동구, 마포구 등이 포함되며 Cluster 1과 Cluster 2 사이의 교통량 수준이 나타나는 지역으로 주로 외곽의 간선도로를 경유하는 지역들이 분포하고 있다.

<Fig. 9>에서 보여지는 바와 같이 초미세먼지 시계열 군집분석 결과 노원구, 강북구 등이 Cluster 1에 포함되며 겨울철(12월~2월) 초미세먼지 수치가 가장 높게 올라가는 지역으로 해당 계절에 초미세먼지 저감 조치가 우선적으로 관리되어야 하는 지역이다. 노원구는 교통시설 및 공공시설 비율이 25개 자치구 중 2위로 주로 교통시설과 공공시설 인근 지역에서 Cluster1과 같은 특성을 보이는 것으로 분석되었다. Cluster2는 중구와 서초구 등이 포함되며 봄철(3월~5월) 초미세먼지 수치는 가장 양호하지만 겨울철(12월~2월) 초미세먼지 수치는 Cluster1 보다는 양호하고 Cluster3보다는 심각한 지역이다. 중구는 공공시설의 비율이 25개 자치구 중 3위로 주로 공공시설 인근 지역에서 Cluster2와 같은 특성이 나타난다. Cluster3에는 동대문구와 강남구가 대표적이고 겨울철(12월~2월) 초미세먼지 수치가 가장 양호한 지역이다. 주거시설 비율이 높은 강남구와 동대문구에서 주로 발생하며 주거시설 인근 지역일수록 Cluster3과 같은 특성을 보인다. 미세먼지 시계열 군집분석 및 연평균 지역별 분포 비교·분석 결과 <Fig. 10>의 Cluster1은 초미세먼지 비교·분석 결과와 유사한 시공간상 특성을 보이나 Cluster2와 3의 경우 초미세먼지보다 한달 가량 지연되어 봄철(3월~5월) 중 4월부터 미세먼지 수치가 급등하는 경향이 있는 것을 도출하였다. 반면 토지이용현황에 따른 Cluster별 초미세먼지와 미세먼지 분포는 유사한 것으로 나타났다.

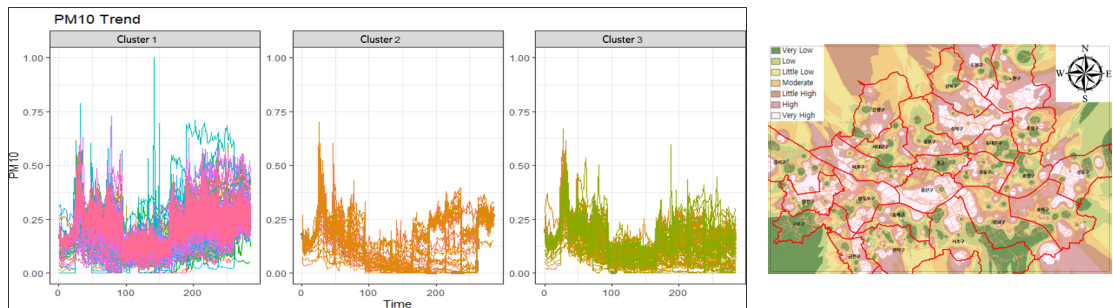
소음 시계열 군집분석 및 연평균 지역별 분포 비교·분석 결과 모든 Cluster에서 공통적으로 여름철(6월~8월)에 소음 발생이 급격하게 증가하는 것으로 분석되었다. 아울러 Cluster1과 같은 유형은 서대문구와 도봉구에서 주로 나타나며 가장 낮은 수준의 소음 발생이 관측된 지역으로 서대문구와 같이 주로 주거시설 인근 지역에서 Cluster1과 같은 특성을 보이는 것으로 나타났다. Cluster2 유형은 주로 영등포구와 강북구에서 나타나며 두 번째로 높은 수준의 소음 발생이 관측된 지역에 해당한다. Cluster3 유형의 경우 가장 높은 수준의 소음 발생이 지속적으로 나타나는 지역으로 금천구나 동작구와 같이 공업시설과 교통시설이 많은 지역에서 주로 관측되어 교통량과 토지이용은 소음과 매우 밀접한 관련이 있다.



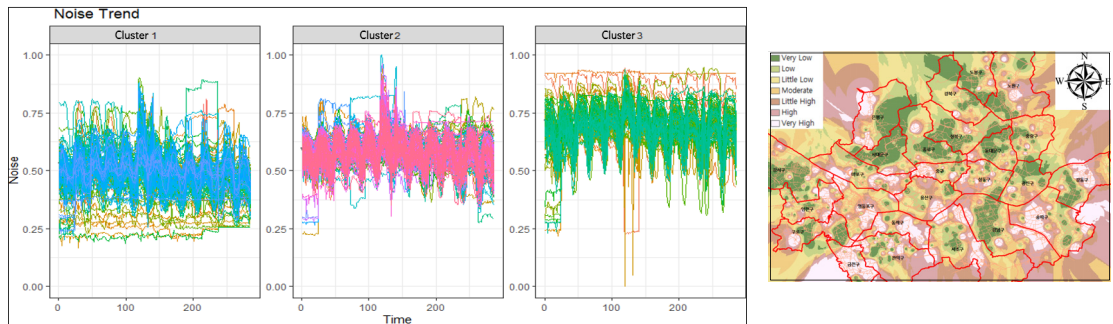
<Fig. 8> Monthly Averaged Traffic by S-DoT sensor/Annual Averaged Traffic by District in Seoul



<Fig. 9> Monthly Averaged PM2.5 by S-DoT sensor/Annual Averaged PM2.5 by District in Seoul



<Fig. 10> Monthly Averaged PM10 by S-DoT sensor/Annual Averaged PM10 by District in Seoul



<Fig. 11> Monthly Averaged Noise by S-DoT sensor/Annual Averaged Noise by District in Seoul

VI. 결 론

본 연구에서는 IoT 센서를 통해 수집한 스마트서울 도시빅데이터와 서울시 도로 링크별 교통데이터 및 토지이용 현황이 함께 반영된 교통-도시환경 통합데이터를 구축하여 도로교통특성과 유해 환경요인 간 영향관계를 비교·분석하였다.

그 결과, 도로 영향권에서의 초미세먼지(PM2.5), 미세먼지(PM10), 소음(Noise)은 Class 1(Very Unhealthy) 수준이 도로 비영향권보다 빈번하게 발생하는 것으로 나타났다. 또한 공업시설, 공공시설, 교통시설, 주거시설, 녹지구역 등 5가지 공간적 특성에 따른 유해 환경요인의 변화패턴으로부터 도로 영향권 유무와 무관하게 주

거시설과, 교통시설 인근에서는 초미세먼지(PM_{2.5})와 미세먼지(PM₁₀)로 인한 지속적인 건강 피해가 예상되었으며, 도로 영향권에 해당하는 녹지구역의 경우 도로 비영향권 대비 봄·가을철 초미세먼지(PM_{2.5})와 미세먼지(PM₁₀) 수치의 급격한 증가에 대한 사전 관리 대책이 필요할 것으로 나타났다. 공업지역은 타 지역 대비 높은 수준의 소음(Noise)의 피해가 지속적으로 발생하는 것으로 분석되어 소음 저감 정책 등의 주기적인 관리가 필요하다. 아울러 서울시 시공간적 특성에 따른 지역별 도시환경요인 위험수준 Hot-Spot 분석 결과, 교통시설과 공공시설 비율이 타 지역 대비 높은 노원구는 초미세먼지(PM_{2.5})와 미세먼지(PM₁₀)로 인한 건강 피해가 가장 심할 것으로 예상되는 Hot-Spot 지역으로 분류되었다. 또한 공업시설과 교통시설의 비율이 타 지역 대비 높은 금천구와 동작구는 소음(Noise)으로 인한 건강 피해가 가장 심할 것으로 예상되는 Hot-Spot 지역으로 분류되었다. 이러한 분석 결과로 미루어 보아 교통시설 인근 지역은 초미세먼지(PM_{2.5}), 미세먼지(PM₁₀), 소음(Noise) 등 주요 3가지 유해 환경요인으로부터 가장 취약할 것으로 분석된다.

본 연구를 통하여 향후에는 초미세먼지(PM_{2.5}), 미세먼지(PM₁₀), 소음(Noise) 등 지역별 주요 도시건강 위험 요소를 색출하여 시민들이 합리적으로 수용할 수 있는 건강교통체계를 구축하고 나아가 현재 서울 중심부 4대문 기준 ‘녹색교통 진흥 지역’과 같은 건강교통 정책 결정 과정에 있어 우선순위 선정 기준으로 활용할 수 있는 연구를 모색할 필요가 있다. 또한 추가적인 IoT 센서의 설치와 성능 향상을 통해 습도, 풍속, 풍향과 같은 추가적인 기상환경변수의 반영과 더불어 더욱 미시적인 도로 영향권 설정 기준으로 고도화된 분석 모델 구축도 가능할 것으로 보인다. 아울러 도시 빅데이터 수집 및 활용의 필요성이 증가하고 있고 이러한 흐름에 기인한 스마트 서울 정책 마련의 일환으로 빅데이터 수집 및 활용이 용이해짐에 따라 IoT 도시 빅데이터를 활용한 건강교통 관점에서의 서비스 제공 방안도 기대할 수 있을 것이다.

ACKNOWLEDGEMENTS

본 연구는 2020년 서울시립대학교 시정연구 지원사업비에 의하여 수행되었습니다.

REFERENCES

- Ali Khalil M. et al.(2019), “Developing machine learning models to predict roadway traffic noise: An opportunity to escape conventional techniques,” *Transportation Research Record*, vol. 2673, no. 4, pp.158-172.
- Barzyk T. M. et al.(2009), “Development of a distance-to-roadway proximity metric to compare near-road pollutant levels to a central site monitor,” *Atmospheric Environment*, vol. 43, no. 4, pp.787-797.
- Britter R. E. and Hanna S. R.(2003), “Flow and dispersion in urban areas,” *Annual Review of Fluid Mechanics*, vol. 35, no. 1, pp.469-496.
- Can A. and Aumond P.(2018), “Estimation of road traffic noise emissions: The influence of speed and acceleration,” *Transportation Research Part D: Transport and Environment*, vol. 58, pp.155-171.
- Clean Air Policy Support System(2017), *National air pollutants emission*, p.11.
- Gastaldi M. et al.(2014), “Evaluation of air pollution impacts of a signal control to roundabout

- conversion using microsimulation,” *Transportation Research Procedia*, vol. 3, pp.1031-1040.
- Gulliver J. and Briggs D. J.(2005), “Time-space modeling of journey-time exposure to traffic-related air pollution using GIS,” *Environmental Research*, vol. 97, no. 1, pp.10-25.
- Hagler G. S. W. et al.(2009), “Ultrafine particles near a major roadway in Raleigh, North Carolina: Downwind attenuation and correlation with traffic-related pollutants,” *Atmospheric Environment*, vol. 43, no. 6, pp.1229-1234.
- Hochadel M. et al.(2006), “Predicting long-term average concentrations of traffic-related air pollutants using GIS-based information,” *Atmospheric Environment*, vol. 40, no. 3, pp.542-553.
- Jeong C. H. et al.(2019), “Temporal and spatial variability of traffic-related PM_{2.5} sources: Comparison of exhaust and non-exhaust emissions,” *Atmospheric Environment*, vol. 198, pp.55-69.
- Kang S. J. et al.(2008), “Types, Locations and Characteristics of Residential Areas in Seoul,” *Journal of Korea Planners Association*, vol. 43, no. 3, pp.129-143.
- Kim C. J. et al.(2018), “Correlation analysis between vehicular traffic and PM using sensor Big Data,” *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*, IEEE, pp.644-648.
- Korea Environment Institute(2007), *An Environmental Study on the Health Effects of Traffic Noise*, pp.4-45.
- Kuehni C. E. et al.(2006), “Association between reported exposure to road traffic and respiratory symptoms in children: Evidence of bias,” *International Journal of Epidemiology*, vol. 35, no. 3, pp.779-786.
- Lipfert F. W. et al.(2006), “Traffic density as a surrogate measure of environmental exposures in studies of air pollution health effects: Long-term mortality in a cohort of US veterans,” *Atmospheric Environment*, vol. 40, no. 1, pp.154-169.
- Makarewicz R. and Gałuszka M.(2011), “Road traffic noise prediction based on speed-flow diagram,” *Applied Acoustics*, vol. 72, no. 4, pp.190-195.
- Ministry of Environment, <http://stat.me.go.kr/nesis/mesp/info/statPolicyHealth4.do>, 2021.08.08.
- Nourani V. et al.(2020), “Artificial intelligence based ensemble model for prediction of vehicular traffic noise,” *Environmental Research*, vol. 180, pp.108852.
- Penard-Morand C. et al.(2006), “Assessment of schoolchildren’s exposure to traffic-related air pollution in the French Six Cities Study using a dispersion model,” *Atmospheric Environment*, vol. 40, no. 13, pp.2274-2287.
- Ponce N. A. et al.(2005), “Preterm birth: The interaction of traffic-related air pollution with economic hardship in Los Angeles neighborhoods,” *American Journal of Epidemiology*, vol. 162, no. 2, pp.140-148.
- Seo Y. and Kim S. M.(2013), “Estimation of greenhouse gas emissions from road traffic: A case study in Korea,” *Renewable and Sustainable Energy Reviews*, vol. 28, pp.777-787.
- Sohn S. and Shin J.(2003), “A Study on the Spatial Structure of Residential Block in Seoul,” *Journal of the Architectural Institute of Korea*, vol. 19, no. 4, pp.83-90.
- Vardoulakis S. et al.(2005), “Spatial variability of air pollution in the vicinity of a permanent monitoring station in central Paris,” *Atmospheric Environment*, vol. 39, no. 15, pp.2725-2736.

- Welch D. et al.(2013), "Road traffic noise and health-related quality of life: A cross-sectional study," *Noise and Health*, vol. 15, no. 65, pp.224-230.
- Wu L. et al.(2020), "Development of the Real-time On-road Emission (ROE v1. 0) model for street-scale air quality modeling based on dynamic traffic big data," *Geoscientific Model Development*, vol. 13, no. 1, pp.23-40.
- Zhang S. et al.(2018), "Fine-grained vehicle emission management using intelligent transportation system data," *Environmental Pollution*, vol. 241, pp.1027-1037.
- Zuo F. et al.(2014), "Temporal and spatial variability of traffic-related noise in the City of Toronto, Canada," *Science of the Total Environment*, vol. 472, pp.1100-1107.