

도로 주변 지역 미세먼지 농도 추정 모형 개발 -교통량, 기상, 배경농도 이력자료를 기반으로-

정도영

한국건설기술연구원 도로교통연구본부

A Model for Estimating Near-road Particle Matter Concentrations -Based on Historical Data for Traffic, Meteorology, Background Concentration-

Do-young JUNG

Department of Highway and Transportation Research,
Korea Institute of Civil Engineering and Building Technology

요약 본 연구는 교통량, 기상조건, 배경농도의 이력자료를 기반으로 도로 주변 지역의 미세먼지 농도를 추정하는 연구이다. 본 연구는 수집된 데이터셋을 기반으로 두 가지 다중 선형회귀 모델을 개발하는 것으로 하나는 도로주변 지역의 PM 2.5 농도를 예측하기 위한 것이고, 다른 하나는 도로주변 지역의 PM 10 농도를 예측하기 위한 것이다. 교통량, 배경농도 등 다양한 설명 변수와 풍속, 온도 등 여러 기상 변수를 고려하여 모델을 개발하고자 하였다. 그런 다음, 개발된 모형이 다른 연구 영역에 적용될 수 있는지 확인하기 위해 모델 이전 가능성을 테스트하였다. 모형 결과는 고려된 변수가 도로주변 지역 PM 농도를 결정하는데 영향력이 있는 매개 변수라는 것을 확인하였다. 특히, 시간당 1,000대의 차량이 증가하면 PM 2.5 농도는 $2.421\mu\text{g}$, PM 10 농도는 $6.769\mu\text{g}$ 이 증가하는 것으로 나타나 교통량에 영향을 미치는 것을 확인하였다. 모형 이전성 테스트 결과는 개발된 모형을 다른 지역에 적용할 수 있음을 시사하고 있으며, 다른 지역에 대한 도로주변 PM 농도의 약 40~60%의 차이를 설명하는 것으로 나타났다. 개발된 모형은 정확도 측면에서 신뢰할 수 있는 동시에 최소한의 데이터 요구 사항과 모델링 시간으로 PM 농도를 예측할 수 있는 애플리케이션을 제공한다.

Abstract This study was done to develop a model for predicting the particulate matter (PM) concentration near roads based on observation data for traffic counts, previous near-road PM concentrations, background PM concentrations, and meteorology. Based on a collected dataset, two multiple linear regressions models were developed: one for PM2.5 and one for PM10. The models were developed by considering various explanatory variables, including traffic, background PM concentrations, and several meteorology variables, such as wind speed, temperature, etc. Then, the models' transferability was tested to confirm that they can be applied to other regions. The model results indicate that the considered variables are influential parameters that determine near-road PM concentrations. Notably, the marginal effects of traffic were identified, suggesting that an increase of 1,000 vehicles per hour increases the PM2.5 concentration by $2.421\mu\text{g}$ and the PM10 concentration by $6.769\mu\text{g}$. The model transferability test results suggest that the developed models can be applied to other regions, explaining around 40~60% of the variance in near-road PM concentrations in such regions. The developed models are reliable in terms of accuracy while also providing easy-to-use applications for predicting PM concentrations with minimal data requirements and modeling time.

Keywords : Particulate Matter, Traffic-Related Air Quality, Traffic, Meteorology, Model Transferability

본 논문은 국토교통부의 연구비지원(과제명: 2022년 도로교통량 조사(TMS))에 의해 수행되었음.

*Corresponding Author : Do-young JUNG(Korea Institute of Civil Engineering and Building Technology)

email: jdy@kict.re.kr

Received September 26, 2022

Revised October 27, 2022

Accepted November 4, 2022

Published November 30, 2022

1. 서론

최근 미세먼지를 포함한 각종 오염물질과 온실가스가 인간의 건강을 해치고, 지구환경을 위협하기 시작하면서, 오염물질 및 온실가스를 줄이기 위한 전 세계적인 노력이 이루어지고 있다. 특히 교통부문은 미세먼지(PM), 일산화탄소(CO), 오존(O₃) 등을 포함하는 각종 오염물질과 온실가스(CO₂)의 주 발생원 중 하나로 알려져 있으며, 이에 따라 관련 정부부처와 연구기관에서는 교통부문으로부터 발생하는 오염물질과 온실가스 발생량을 줄이기 위한 각종 노력을 기울여 왔다. 이러한 노력의 일환으로 여러 연구기관에서는 교통부문의 오염물질과 온실가스 발생량을 결정하는 영향 요인을 분석하는 것에 초점을 맞추어 다양한 연구를 수행해왔다. 그 결과 여러 국내의 연구기관에서는 도로주행 시험결과 등을 바탕으로 차종/연료/주행 특성 등에 따른 오염물질 및 온실가스 발생량을 추정하는 다양한 모형을 개발하였다. 그 대표적인 모형으로 국내에서는 국립환경과학원에서 개발된 차종별/연료별 배출계수식[1-3]이 있으며, 국외에서는 미국환경청(USEPA)에서 개발된 MOVES 모형[4], 캘리포니아 대기자원위원회(California Air Resource Board, CARB)에서 개발된 EMFAC 모형[5], 기후 변화에 관한 정부간 협의체(IPCC)에서 개발된 EFDB 모형[6] 등이 있다. 이처럼 차량 오염물질 및 온실가스 발생량 추정에 관한 연구는 상당히 오랜기간에 걸쳐 개발이 이루어져 왔으며, 차종별, 연료별, 주행 특성에 따른 오염물질 발생량 추정을 가능케하여, 추정결과와 정확성이 어느 정도 확보되어 왔다.

교통부문에서의 오염물질에 관한 연구는 도로교통오염원, 즉 차량에서 발생하는 오염물질이 공간에 어떻게 확산되는 지에 관한 연구로 확장되어왔다. 이는 차량에서 발생하는 오염물질이 대기 중에 확산되어, 실제 인간의 거주공간에서 인간에게 어느 정도 수준으로 부정적 영향을 미치는지 파악하기 위함이며, 이러한 연구는 주로 환경공학분야에서 이루어져 왔다. 특히 도로 교통오염원에 따른 피해는 도로 인접지역(약 도로로부터 200m 이내)에서 크게 발생하며[7], 또한 대부분의 도시 인구가 이러한 영향권내에서 거주 또는 활동하기 때문에, 많은 선행 연구들은 주로 도로주변지역의 오염물질 농도를 추정하는 모형을 개발하는 데 초점을 맞추어 왔다(이를 미시적 대기확산모델링 또는 microscopic dispersion modeling 이라 함). 이러한 연구는 오염물질이 변화무쌍한 기상 조건(예: 기온, 습도, 풍향, 풍속 등)에서 어떻

게 대기에 확산되며, 피해지역에 어느 정도 수준으로 오염물질이 농축되는 지에 관하여 예측하는 것을 목표로 하고 있다. 관련 연구분야를 선도하고 있는 국가는 미국이며, 관련 연구를 주도하고 있는 기관으로는 USEPA와 CARB가 있다. 특히 USEPA는 AERMOD, R-LINE이라는 대기확산모형을 개발하였으며[8], CARB는 CALINE 모형을 개발하여[9], 차량에서 발생하는 오염물질이 대기에 확산하는 형태를 모델링하고, 이를 통해 관측하고자 하는 지점의 오염물질 농도를 추정한다. 다만, 국내에서는 차량 오염물질의 대기확산 모형 개발에 관련된 연구는 부재한 실정이며, 몇몇 국내 연구진에서는 미국의 모형을 도입하며 도로교통오염원으로 인한 주변지역의 오염물질 농도 추정을 시도해왔으나[10,11], AERMOD와 같은 대기확산모형은 정밀한 기상입력자료를 요구하고 있어, 국내의 경우 이러한 입력자료를 구축하고 관리하는 기관이 부재하여, 실무자 또는 연구자가 이처럼 복잡한 모델링을 수행하는 데 현실적으로 어려움이 있다. 또한 미국 모형을 국내에 적용함에 있어 모형의 정산이 이루어지지 않아 추정값의 정확성에 대한 의문이 제기되어 왔다.

대기확산 모델링을 활용한 추정방식과 더불어 도로교통 오염원의 오염물질 농도를 추정하는 방식으로 주로 활용되는 방법론은 도로주변의 오염물질 농도에 대한 관측값과 해당도로의 교통현상(예: 교통량, 속도, 차종 등)과 해당 지역의 기상 관측값을 바탕으로 통계모형을 구축하는 방식이다. 이러한 접근법은 실제 관측값을 바탕으로 모형이 구축되기 때문에 경우에 따라서 모형의 추정력이 높게 나타날 수 있다는 장점이 있으나, 관측지점의 조건과 다른 지역에 적용될 때 오염물질 농도 추정값의 신뢰성이 낮아질 수 있는 단점이 있다. 예를 들어, 모형구축의 기반이 되는 관측자료(교통량, 기상변수 등)의 변동성이 충분히 확보되지 못할 경우 모형의 설명력이 다른 지역을 설명하기에 충분하지 않을 수 있다. 또한 일반적으로 교통량 또는 미세먼지농도 관측을 위해서는 매우 높은 비용과 시간이 요구되기 때문에, 관측은 특정 기간/지점에 한정되며, 이에 따라 기존 연구에서는 모형 구축을 위한 자료의 변동성이 충분히 확보되어 오지 못했다는 단점이 지적되어 왔다. 이처럼 자료의 변동성이 충분히 확보되지 못할 경우, 이를 바탕으로 구축된 모형은 조사된 자료의 범위내에서만 유의하며 그 범위를 넘어설 경우에는 추정치의 신뢰성이 낮아진다는 단점이 있다. 예를 들어, 여름철 자료를 바탕으로 미세먼지농도를 추정하는 모형이 구축되었다면, 구축된 모형을 통해 겨울

철 미세먼지농도를 추정할 때, 추정치의 신뢰성이 낮아질 수 있다.

이처럼 신뢰성있는 도로주변지역의 오염물질농도를 추정하는 것은 모델링의 복잡성, 자료구축의 한계 등의 이유로 매우 어렵다. 또한 국내에는 무수히 많은 도로가 있으며, 이처럼 무수히 많은 도로에 대해서 기존의 대기 확산모델링 분석방법 등을 활용하여 오염물질의 농도를 예측하는 것은 자료구축과 분석시간 등을 고려해 보았을 때, 현실적으로 매우 어렵다. 이러한 상황에서 상대적으로 적용이 용이한 기존 통계모형을 활용한 오염물질농도 추정을 고려해 볼 수 있으나, 선행 연구에서 제시된 모형의 낮은 신뢰성, 국내 여건을 고려한 모형의 부재 등의 이유로 선행연구에서 제시된 통계모형을 직접적으로 적용하기는 어려운 실정이다.

본 연구는 이러한 기존연구의 한계를 극복하고자 국내 여건에 부합하고, 분석이 용이하며, 신뢰성있는 오염물질농도 추정을 가능하도록 모형을 개발하고자 한다. 구체적으로 본 연구는 교통 및 기상 현상과 도로인접지역의 미세먼지농도의 상관성을 분석하고, 도로 인접지역 미세먼지농도를 추정할 수 있는 모형을 개발하는 것을 목표로 한다. 또한 그 결과를 바탕으로 궁극적으로 교통 부문에서 발생하는 오염물질의 공간적 확산형태를 이해하고, 이를 바탕으로 도로에서 발생하는 오염물질이 실질적으로 인간에게 미치는 영향을 파악할 수 있는 기틀을 제공하고자 한다. 이를 위해, 본 연구는 교통량, 기상 조건(예: 온도, 습도, 강수 등), 배경농도(Background concentration)와 도로주변 미세먼지농도(PM2.5, PM10) 자료를 장기간에 걸쳐 수집하고, 구축된 자료를 바탕으로 변수들간의 상관관계를 분석하며, 통계모형을 구축하여 도로인접지역의 미세먼지농도를 추정하는 모형을 개발하고자 한다. 또한 통계모형 결과를 바탕으로 궁극적으로 교통량이 도로인접지역 미세먼지농도에 미치는 한계효과(예: $\mu\text{g}/\text{대}/\text{시}$)를 추정하여, 실무자 또는 관련분야 연구자들에게 유용한 정보를 제공하고자 한다. 마지막으로 연구지역의 자료를 기반으로 개발된 모형이 과연 다른 연구지역에 적용될 수 있는지 확인하고자, 모형의 이전성(Model Transferability)을 검증한다.

2. 선행 연구 고찰

본 연구에서는 도로주변지역의 미세먼지 농도를 추정하는 모형을 개발하였다. 모형 개발을 위해, 본 연구는

도로 주변지역의 미세먼지 농도에 영향을 미치는 영향인자에 대한 문헌조사를 수행하였으며, 문헌조사 결과를 요약하면 Table 1과 같다. Table 1에서 보는 바와 같이 다양한 요인들이 도로인접지역의 오염물질 농도에 중요한 영향을 미치는 것으로 나타났다. Liu and Kim (2019)와 Zhang et al. (2018)은 도로오염원으로부터 발생하는 오염물질의 총량이 증가함에 따라 도로인접지역의 오염물질 농도가 증가할 수 있음을 제시하였다 [12,13]. 여기서 도로오염물질의 발생량은 교통량/통행 속도/차량구성 등 여러 교통류 특성에 영향을 받으므로, 결국 도로인접지역의 오염물질 농도를 추정하기 위해서는 위에 제시된 다양한 교통류 변수들을 고려해야 함을 제시하고 있다. Liu et al. (2017), Wu and Niemeier (2016), Igri et al. (2011) 등 여러 연구에서는 도로로부터의 거리(Distance from road segment)가 멀어질수록 오염물질의 농도가 낮아짐을 제시하고 있다[14-16]. 특히 HEI(2010)은 일반적으로 도로오염물질의 영향권을 200m 이내로 보고 있으며[7], 특히 도로에 인접할수록 오염물질 농도가 매우 높게 나타나며, 거리가 멀어질수록 농도가 점차적으로 낮아지며, 거리가 200m 넘어가게 되면 배경농도(background concentration)에 다르게 된다고 제시하고 있다.

또한 문헌조사 결과는 오염물질 농도는 다양한 기상 조건(예: 온도, 습도, 풍향, 풍속 등)에 큰 영향을 받음을 제시하고 있다. 특히 풍속은 가장 중요한 기상요인 중 하나로 일반적으로 풍속이 높아짐에 따라 도로주변 지역의 오염물질 농도는 낮게 나타난다[13,17,18]. 이는 풍속이 높아짐에 따라 오염물질이 대기 중에 더욱 빨리 확산되기 때문이다. 이 밖에도 표면온도(Surface temperature)와 강수(Precipitation)는 도로주변지역의 오염물질 농도를 낮추는 기상 요인이다[13,17,19,20]. 반대로 습도(humidity), 기압(pressure), 운량(cloud cover)은 오염물질 농도와 음의 상관성이 있는 것으로 제시된다 [13,17-19]. 즉, 예를 들어 습도가 증가할수록 도로 주변 지역의 오염물질 농도는 증가할 수 있다.

선행 연구 고찰 결과는 도로 주변지역의 오염물질 농도에 영향을 미치는 요인은 매우 다양함을 보여준다. 즉, 도로 주변지역의 오염물질 농도를 예측하는 모형을 구축하기 위해서는 이처럼 다양한 영향인자를 모형에 고려할 필요가 있음을 시사하고 있다. 본 연구는 선행 연구 고찰 결과에서 제시된 다양한 영향인자를 분석 변수에 고려하였으며, 실제 자료의 수집 가능여부 등을 고려하여 최종 모형을 구축하였다.

Table 1. Factors Affecting Near-road Pollutant Concentrations.

Variables	Relation to concentration	References
Link emission rate	Positive	Liu and Kim (2019); Zhang et al. (2018)
Distance from road segment	Negative	Liu et al. (2017); Wu and Niemeier (2016); Igri et al. (2011)
Wind speed	Negative	Zhang et al. (2015); Tecer et al. (2008); Akpinar et al. (2008); Zhang et al. (2018)
Surface temperature	Negative	Zhang et al. (2015); Tecer et al. (2008); Akpinar et al. (2008); Zhang et al. (2018); Lin et al. (2015)
Precipitation	Negative	Tecer et al. (2008); Lin et al. (2015)
Humidity	Positive	Zhang et al. (2015); Tecer et al. (2008); Akpinar et al. (2008)
Pressure	Positive	Akpinar et al. (2008); Zhang et al. (2018)
Cloud cover	Positive	Tecer et al. (2008)

Note: The main contents of this table was retrieved and adjusted from the Table 2 illustrated in Kim et al. (2020).

3. 분석자료

본 연구는 도로인접지역의 초미세먼지(PM_{2.5})와 미세먼지(PM₁₀) 농도를 추정하는 연구를 수행하였다. 구체적으로 (초)미세먼지의 농도 추정을 위해 선행 연구에서 제시된 다양한 영향요인을 설명변수로 활용하였으며, 이를 위해 광범위한 분석자료를 수집하였다. 본 연구에서 수집한 자료는 연구대상도로의 교통량과 도로인접지역의 미세먼지 농도, 주변 지역의 기상 조건(온도, 습도, 풍속, 강수량 등)과 배경농도를 포함한다.

특히 관측자료를 기반으로 모형을 구축한 기존 선행 연구들에서는 자료 수집에 따른 시간/비용의 한계로 인해 일반적으로 자료의 수집기간이 짧아 자료 변동성이 매우 작으며, 이에 따라 모형의 신뢰성이 다소 떨어질 수 있다는 단점이 제기되어 왔다. 이러한 상황에서, 본 연구에서는 오랜 기간에 걸친(2019년 7월~2020년 2월, 244일) 광범위한 자료 수집을 통해 자료의 변동성을 확보하였으며, 이에 따라 모형의 신뢰성을 높일 수 있다는 점에서 기존 연구와는 차별성이 있다.

3.1 연구 대상 지역

본 연구에서는 두 개의 도로구간을 대상으로 자료를

수집하였다(① 지점ID 10347: 경기도 평택시 포승읍 만호리 489, ② 지점ID 40011: 충청남도 아산시 영인면신봉리511-7) (Fig. 1).

이는 한국건설기술연구원에서 운영 중이 도로교통량 조사장비를 활용하여 정확한 교통량을 수집하고 이에 도로교통량 조사장비에 미세먼지 센서를 추가설치하여 미세먼지와 교통량의 관계를 정확히 파악하고자 한다. 또한, 지방부의 다른 환경의 여건의 변화를 최소화하고자 한다.

도로 구간의 기상 조건과 관련된 자료는 기상자료개방포털을 통해 수집하였으며(Korea Meteorological Administration, 2021), 연구 도로와 인접한 두 곳의 기상청 측정 지점을 선정하였다(Fig. 1). 마지막으로 배경농도는 에어코리아(AirKorea, 2021)의 시간대별 미세먼지 농도 측정자료를 통해 수집되었으며, 본 연구를 위해 연구 도로지역을 둘러싼 5곳의 배경농도 관측 지점을 수집 대상으로 하였다(Fig. 1).

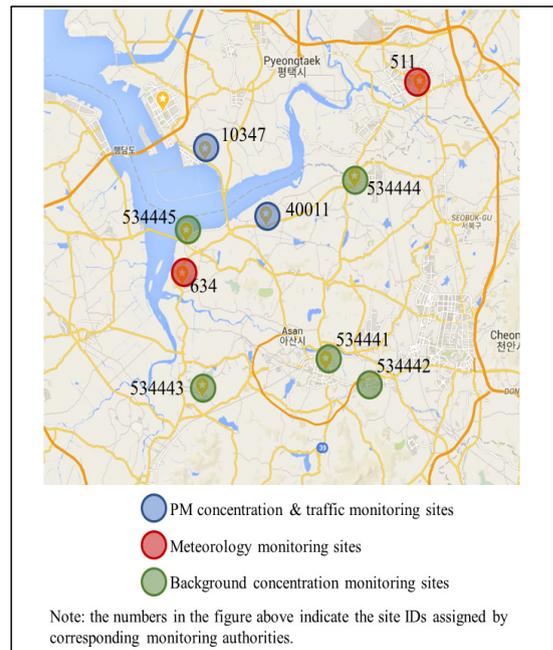


Fig. 1. Summary of data collection sites

3.2 미세먼지 농도-교통량 자료 수집

앞서 언급되었듯이, 본 연구는 두 곳의 도로구간을 대상으로 자료를 수집하였다. 자료 수집을 위해 해당 도로구간의 교통량과 미세먼지농도를 상시 수집하는 간이 측정기를 설치하였다(Fig. 2). 자료 수집기간은 2019년 7

월~2020년 2월(총 244일)이며, 5분 단위로 자료가 집계 및 수집되었다. Fig. 2에서 보는 바와 같이, (초)미세먼지 관측지점은 도로와 매우 인접해 있으며, 따라서 본 연구의 목표에 부합하는 도로 인접지역 (초)미세먼지 추정을 위한 자료로 활용되기에 적합해 보인다.

미세먼지농도 수집에 사용된 간이 측정기는 미세먼지 ($0\sim 500\mu\text{g}/\text{m}^3$), 온도($-40\sim 125^\circ\text{C}$), 습도($0\sim 100\%$) 인자 측정이 가능하며, 공인기관 신뢰성 시험을 통과하였다. 현재 국내 공공기관(서울시, 한국도로공사 등)에서 사용중인 기기로 측정기를 통해 수집되는 데이터정보는 <Fig. 3>과 같다.



Fig. 2. Illustrations of Data Collection Sites for Traffic and PM Concentrations.

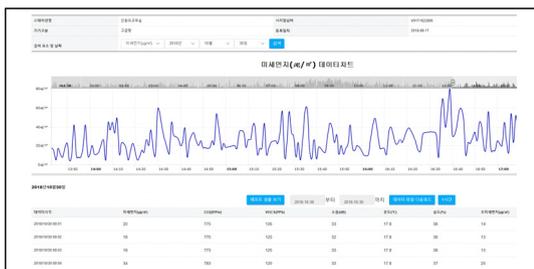


Fig. 3. Sample data of measuring devices

3.3 기상청 측량 자료 수집

본 연구는 도로주변지역 미세먼지농도를 설명하기 위해 다양한 기상 관련 변수를 고려하였다. 본 연구에서 활용된 기상 관련 자료는 기상청의 기상자료개방포털을 통해 수집되었으며, 본 연구에서 수집된 자료로는 풍향, 풍속, 강수량, 기압 등이 있다. 특히 기상자료개방포털에서는 시간대별로 자료를 집계하여 이력자료를 제공하고 있다. 본 연구에서 고려된 기상청 측량 자료는 연구 대상 도로와 인접한 두 곳의 관측지점(① 지점ID 551: 위도 36.98769, 경도 127.10879, ② 지점ID 634: 위도 36.84578, 경도 126.86536)을 대상으로 하였다(Fig. 1).

3.4 배경농도 수집

배경농도란 교통량 또는 차량주행의 영향이 미미한 지점(즉, 도로로부터 상당히 떨어진 지점)의 미세먼지농도를 의미한다. 배경농도를 형성하는 요인은 일반 자연환경에서 형성되는 오염원으로부터 형성되는 농도를 의미하며, 즉 황사와 같은 일반 자연환경오염원, 발전소, 난방 등과 같은 일반 도시오염원으로부터 형성되는 오염물질 농도를 의미한다. 본 연구에서는 에어코리아(airkorea.or.kr)에서 수집되는 시간대별 미세먼지농도 측정데이터를 활용하였다. 이를 위해, 연구 대상 도로와 인접한 에어코리아 관측지점에서 수집된 PM10과 PM2.5 배경농도 이력자료를 수집하였다. 본 연구에 고려된 배경농도 관측지점은 총 5지점이며, 관측지점에 정보는 다음과 같다(Fig. 1).

- 지점번호 534441: 충남 아산시 모종동 573-2
- 지점번호 534442: 충남 아산시 배방읍 배방로 38
- 지점번호 534443: 충남 아산시 도고면 기곡리 296-4
- 지점번호 534444: 충남 아산시 둔포면 석곡리 1481
- 지점번호 534445: 충남 아산시 인주면 인주산단로 23-28

4. 분석결과

4.1 분석방법론

본 연구는 수집된 자료를 바탕으로 도로주변지역의 미세먼지농도를 설명하기 위해 통계모형을 구축하였다. 분석을 위해 다양한 출처로부터 수집된 자료를 병합하고, 통계분석 수행을 위해 자료의 무결성(data integrity)를 확인하고, 자료정리(data cleaning) 프로세스를 수행하

Table 2. Descriptive Statistics of Collected Data

Data source	Variables		Collection rates (%)	Minimum	Maximum	Mean	Standard deviation	
Traffic & PM concentration monitoring device	Site ID: 10349	Flow rate (vehs/5mins)	97.6	1.0	409.0	141.4	94.3	
		PM2.5 concentration ($\mu\text{g}/\text{m}^3$)	99.1	0.0	361.0	26.9	21.1	
		PM10 concentration ($\mu\text{g}/\text{m}^3$)	99.1	0.0	895.0	45.7	40.9	
	Site ID: 40011	Flow rate (vehs/5mins)	97.5	0.0	322.0	53.2	62.0	
		PM2.5 concentration ($\mu\text{g}/\text{m}^3$)	98.5	0.0	159.0	21.4	18.1	
		PM10 concentration ($\mu\text{g}/\text{m}^3$)	98.5	0.0	461.0	28.7	28.1	
Korea Meteorological Administration (meteorology monitoring)	Site ID: 551	Temperature ($^{\circ}\text{C}$)	99.8	-11.2	35.2	12.9	11.0	
		Humidity (%)	99.8	20.6	99.9	73.6	19.1	
		Wind speed (m/s)	99.8	0.0	9.5	1.6	1.2	
		Precipitation (mm)	99.8	0.0	33.5	0.2	1.2	
	Site ID: 634	Temperature ($^{\circ}\text{C}$)	100.0	-10.7	35.2	12.9	10.9	
		Humidity (%)	100.0	22.4	99.9	70.1	19.4	
		Wind speed (m/s)	100.0	0.0	11.1	1.9	1.4	
		Precipitation (mm)	100.0	0.0	23.5	0.1	1.0	
airkorea.or.kr (background concentration monitoring agency)	PM2.5 concentration ($\mu\text{g}/\text{m}^3$)	Total	94.9	0.0	161.0	26.9	19.6	
		Site ID: 534441	92.0	0.0	151.0	27.0	19.8	
		Site ID: 534442	95.5	0.0	147.0	23.5	17.2	
		Site ID: 534443	95.5	0.0	122.0	25.7	19.0	
		Site ID: 534444	94.0	0.0	159.0	31.5	21.9	
		Site ID: 534445	97.5	0.0	161.0	27.0	19.0	
	PM10 concentration ($\mu\text{g}/\text{m}^3$)	Total	97.0	0.0	235.0	43.5	25.7	
		Site ID: 534441	96.4	0.0	235.0	41.2	26.4	
		Site ID: 534442	96.4	3.0	202.0	41.4	20.5	
		Site ID: 534443	98.0	0.0	230.0	41.0	26.1	
		Site ID: 534444	96.4	0.0	227.0	47.2	27.8	
		Site ID: 534445	97.8	0.0	233.0	46.6	26.0	

였다. 구축된 자료를 바탕으로 상관분석을 수행하였으며, 이를 통해 변수들간의 상관성을 파악하였다. 또한 통계 모형(즉, 다중선형회귀모형)을 구축하여, 미세먼지농도와 설명변수들간의 상관성과 한계효과를 측정하였다. 마지막으로 구축된 모형의 전이성(model transferability)을 평가하여, 본 연구에서 개발된 모형이 다른 지역에 적용이 가능한지에 대한 평가를 수행하였다.

4.2 분석자료 구축

앞서 언급되었듯이, 본 연구는 다양한 출처로부터 자료를 수집하였으며, 자료의 기술통계량은 Table 2와 같다. 특히 Table 2는 연구도로에 설치된 간이 측정기로부터 수집된 교통량과 미세먼지농도의 변동성이 매우 크게 나타남을 보여주고 있다. 예를 들어, 연구도로의 교통류율은 5분당 1대에서 409대(측정소ID: 10349), 0대에서 322대(측정소ID: 10349)로 시간대별로 교통량의 변동이 매우 크게 나타남을 보여준다. 이와 더불어, 도로주변의 미세먼지농도 또한 큰 변동을 보여주고 있으며, 시간대에 따라 PM2.5 농도의 경우 최소 0.0 $\mu\text{g}/\text{m}^3$ 에서 최

대 361.0 $\mu\text{g}/\text{m}^3$ 사이에서 변동하며, PM10 농도의 경우 최소 0.0 $\mu\text{g}/\text{m}^3$ 에서 최대 895.0 $\mu\text{g}/\text{m}^3$ 사이에서 변동하는 것으로 나타났다. 또한 본 연구에서는 도로주변지역의 미세먼지 농도를 설명하기 위하여, 다양한 기상자료를 기상청으로부터 수집하고, 에어코리아로부터 미세먼지 배경농도자료를 수집하였으며, 기상자료와 배경농도의 기술통계량을 정리하면 Table 2과 같다. Table 2에 제시된 바와 같이 연구도로 지역의 기상과 배경농도 또한 변동성이 크게 나타나는 것으로 보인다. 특히 본 연구에서는 자료를 여름, 가을, 겨울에 걸쳐 약 6개월이라는 장기간에 걸쳐 수집하였으므로, 분석기간동안의 기상자료의 변동성 또한 크게 나타남을 보여준다. 특히 일반적으로 기존 연구에서는 조사에 따른 비용 등의 한계로 인해 짧은 기간 동안 자료수집을 수행하게 되게 되며, 이에 따라 계절에 따른 기상현상의 변동성을 충분히 반영하지 못하고 있으며, 따라서 본 연구에서는 기존연구와는 다르게 충분한 기상현상의 변동성을 고려하고 있다는 점에 기존 연구와는 큰 차별성이 있다. 예를 들어, <Table 2>의 기상청 자료를 살펴보면 기온의 범위가 최

Table 3. Pearson's Correlation Test Results

Near-road PM2.5 concentration vs. Explanatory variables						
Variables	Background PM2.5 concentration	Flow rate	Temperature	Humidity	Wind speed	Precipitation
PM2.5 concentration	0.802 ***	-0.061 ***	-0.302 ***	0.062 ***	-0.216 ***	-0.074 ***
Background PM2.5 concentration		-0.170 ***	-0.281 ***	0.135 ***	-0.242 ***	-0.069 ***
Flow rate			0.483 ***	-0.324 ***	0.020 *	0.042 ***
Temperature				-0.013	-0.036 ***	0.042 ***
Humidity					-0.238 ***	0.122 ***
Wind speed						0.031 ***

Near-road PM10 concentration vs. Explanatory variables						
	Background PM10 concentration	Flow rate	Temperature	Humidity	Wind speed	Precipitation
PM10 concentration	0.585 ***	0.086 ***	-0.156 ***	0.011	-0.163 ***	-0.052 ***
Background PM10 concentration		-0.089 ***	-0.229 ***	0.005	-0.166 ***	-0.093 ***
Flow rate			0.483 ***	-0.324 ***	0.020 *	0.042 ***
Temperature				-0.013	-0.036 ***	0.042 ***
Humidity					-0.238 ***	0.122 ***
Wind speed						0.031 ***

Note: ***, **, * mean significance at the 1%, 5%, and 10% levels, respectively.

소 -11.5도에서 최대 37.7도, 습도는 최소 18.6%에서 최대 99.9%로 계절별 기상현상의 변동성이 충분히 확보하고 있음을 보여준다.

본 연구는 도로주변의 미세먼지농도를 설명하기 위해서 통계모형을 구축하였으며, 통계분석을 위한 자료구축을 위해 여러 출처로부터 수집된 자료를 병합하였다. 다만, 교통량과 미세먼지 자료는 5분 단위로 수집되며, 기상청 및 에어코리아 자료는 1시간 단위로 수집되는 등 자료의 출처 간 자료수집 단위가 상이하므로, 통계분석을 위해 교통량/미세먼지 자료를 1시간 단위로 집계하여, 동일 시간대를 기준으로 자료를 병합하였다. 단, 수집된 자료는 관측장비의 성능 등의 이유로 자료수집률이 100%에 이르지 못하는 경우가 발생하고 있다. 예를 들어, <Table 2>에서 보는 바와 같이 간이 측정기에서 수집되는 교통량-미세먼지자료의 수집률은 97.6~99.1%, 기상청 자료의 수집률은 99.8~100.0%, 에어코리아의 배경농도 자료의 수집률은 92.0~98.0% 수준으로 나타났다. 다만, 통계분석을 수행하기 위해서는 자료의 무결성(integrity)이 확보되어야 하며, 따라서 본 연구에서는 누락된 시간대의 자료를 통계분석에서 제외하였다. 분석 자료 구축 결과, 자료가 누락된 시간대의 자료를 제외한 총 9,537시간의 자료가 통계분석에 활용되었으며, 이는 전체 분석시간(11,712시간=244일×24시간×2개의 도로) 대비 약 81.4%의 자료가 통계분석에 활용되었음을

의미한다. 통계분석을 위해서 에어코리아에서 수집된 배경농도는 연구도로 주변지역 5곳의 관측소에서 측정된 배경농도자료를 동일 시간을 기준으로 평균하고, 그 평균값을 분석에 활용하였다.

4.3 피어슨 상관분석 결과

본 연구는 연구도로주변지역의 미세먼지농도와 영향 인자와의 상관성을 파악하기 위해서 피어슨상관분석(Pearson's correlation test)을 수행하였으며[21], <Table 3>은 분석결과를 보여준다. 분석결과, 도로주변 지역의 미세먼지농도(PM concentrations)와 배경농도(background PM concentrations)는 높은 양의 상관성을 보이는 것으로 나타났다(상관계수가 0.585~0.802 수준임). 이는 연구지역의 전반적인 미세먼지농도가 연구도로주변지역의 미세먼지농도와 매우 밀접한 상관성이 있음을 보여준다. 다만, 교통량과 도로주변의 미세먼지농도간의 상관성은 매우 약한 것으로 나타났다(계수의 절대값이 0에 가까움). 다만, 기존 연구에서는 교통량이 증가할수록 도로주변지역의 미세먼지농도가 동시에 증가할 것이라는 연구결과를 제시하고 있어, 교통량-미세먼지농도와의 관계를 파악하기 위해서는 고급통계분석(예: 회귀분석)을 수행해야 함을 시사하고 있다. 또한 <Table 3>은 도로주변지역의 미세먼지농도와 기상변수들간의 상관성이 있음을 보여주고 있으며, 특히 풍속/온도와는

Table 4. Multiple linear regression model results

Variables	Model for PM2.5 concentration		Model for PM10 concentration	
	Coefficients	Beta coefficients	Coefficients	Beta coefficients
Intercepts	3.670 ***	0.000	-8.257 ***	0.000
Background PM10 concentration ($\mu\text{g}/\text{m}^3$)	0.853 ***	0.771	0.807 ***	0.560
Flow rate (1000 vehicles/hour)	2.421 ***	0.148	6.769 ***	0.226
Temperature ($^{\circ}\text{C}$)	-0.260 ***	-0.157	-0.419 ***	-0.139
Humidity (%)	0.003	0.003	0.133 ***	0.069
Wind speed (m/s)	-0.519 ***	-0.035	-1.665 ***	-0.062
Precipitation (1 if precipitation is over 0.0mm, and 0 otherwise)	-2.247 ***	-0.029	-2.459 **	-0.017
Adjusted R2	0.668		0.381	

Note: ***, **, * mean significance at the 1%, 5%, and 10% levels, respectively.

약한 음의 상관성을 보이는 것으로 나타나, 본 연구에서 구축된 기상변수와 미세먼지농도의 상관성이 선행 연구에서 제시된 연구결과와 어느 정도 부합하고 있음을 보여준다.

4.4 회귀모형 구축 결과

구축된 자료를 바탕으로, 도로인접지역의 미세먼지농도를 추정하는 다중선형회귀모형을 구축하였다. 본 연구에 제시된 다중선형회귀모형을 수식으로 나타내면 다음과 같다.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon \quad (1)$$

여기서, Y 는 종속변수로서 본 연구에서는 도로주변지역의 PM2.5 또는 PM10의 농도이며, X_j 는 j 번째 독립변수(즉, 교통량, 배경농도 및 기상변수)이며, β_0 는 Y 절편, β_j 는 j 번째 모수, ε 는 오차항이다. 본 연구에서 회귀분석의 계수는 최소자승법(ordinary least squares, OLS)을 통해 추정되었으며, 통계소프트웨어 R을 통해 추정되었다[22].

〈Table 4〉에 나타난 것처럼, 본 연구는 두 개의 다중선형회귀모형을 구축하였다. 구체적으로 첫 번째 모형은 도로주변지역의 PM2.5 농도를 예측하며, 두 번째 모형은 도로주변지역의 PM10 농도를 예측하기 위해 구축되었다. 모형 구축 결과, 모형의 조정 R2이 0.381(PM10 모형)에서 0.668(PM2.5 모형) 수준이며, 회귀모형이 미시자료를 기반으로 구축되었음을 고려할 때, 상당히 높은 설명력을 갖고 있음을 보여준다. 추정된 계수를 살펴보면, 도로주변지역의 미세먼지농도를 설명하는 가장 유의한 변수는 배

경농도로 나타났다(베타계수의 절대값이 가장 크게 나타남). 주목할만한 점은 회귀모형 구축 결과, 피어슨상관분석에서는 나타나지 않았던 미세먼지농도와 설명변수들간의 상관성이 통계적으로 유의하게 나타났다는 점이다(대부분의 설명변수의 계수가 유의수준 0.01에 통계적으로 유의함을 보임). 특히 교통량과 도로인접지역 미세먼지농도간의 통계적 유의성이 확보되었으며, 추정된 계수값은 교통량의 미세먼지농도에 대한 한계효과를 제시하고 있다. 예를 들어, 시간당 교통량이 1000대 증가함에 따라 도로주변지역의 PM2.5 농도는 약 $2.421\mu\text{g}$ 증가함을 나타내며, PM10 농도는 약 $6.769\mu\text{g}$ 증가함을 제시하고 있다. 이와 더불어, 기상변수와 미세먼지농도와의 상관성 또한 유의하게 나타났으며, 대다수 계수의 부호는 선행연구에서 제시된 변수들간의 상관성과 일치하는 것으로 나타났다. 예를 들어, 기온/풍속이 증가할수록 또는 강수 여부에 따라 도로주변지역의 미세먼지농도는 감소하며, 습도가 증가할수록 미세먼지농도는 증가하는 것으로 나타났다. 이러한 결과는 도로주변지역의 미세먼지농도를 추정하기 위해서는 교통량 뿐만 아니라 다양한 기상변수를 고려해야 함을 제시하고 있다.

4.5 모형 이전성 평가

마지막으로 본 연구는 구축된 회귀모형이 과연 다른 도로지역의 미세먼지농도를 추정하는 데 적용이 가능한지에 대한 평가를 수행하였다. 본 연구에서는 두 곳의 사례지를 대상으로 자료를 구축하였으므로, 한 지역을 대상으로 구축된 모형을 다른 지역에 적용하여 관측치와 예측치의 비교를 통해 모형 이전성에 대한 평가를 수행하였다. 대상지역 외의 검증은 교통량 조사장비와 미세먼지 센서가 설치되어 있는곳이 없는 관계로 대상사이트

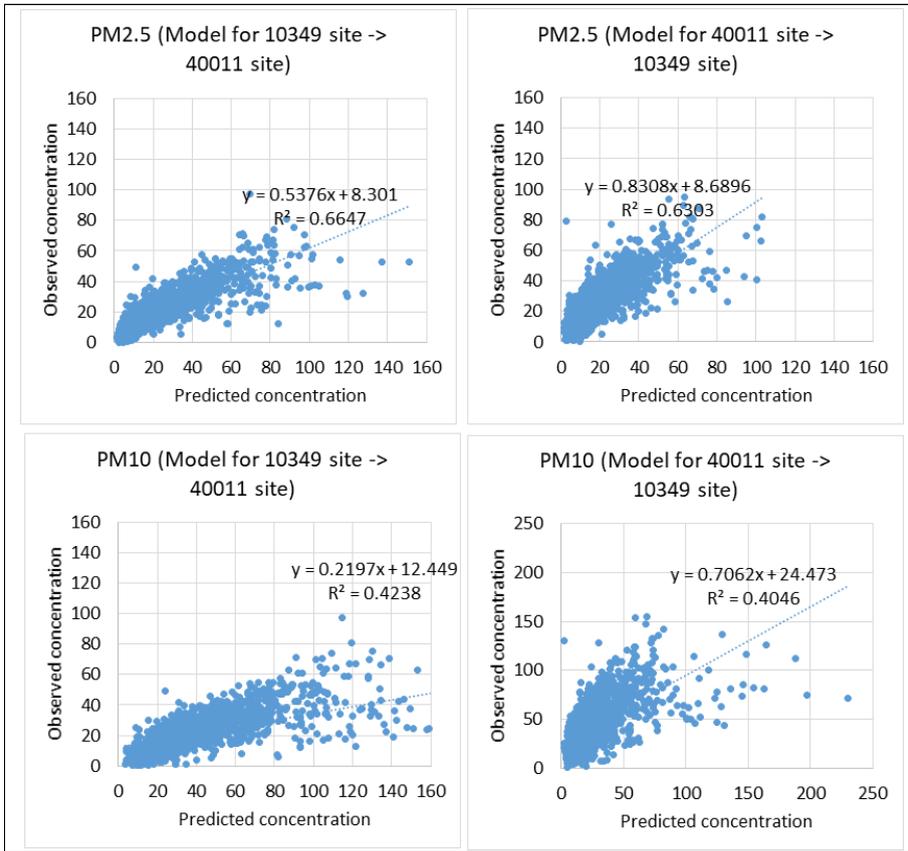


Fig. 4. Result of model transferability evaluation

로 이전성 평가를 수행하였다. 모형 이전성 평가 프로세스를 요약하면 다음과 같다.

- 지점ID 10349 지역의 간이 측정소 자료를 바탕으로 구축된 회귀모형을 지점ID 40011 도로지역의 미세먼지농도를 예측하는 데 활용 → 지점ID 40011 도로 지역의 미세먼지농도 추정치와 관측치의 차이를 비교
- 지점ID 40011 지역의 간이 측정소 자료를 바탕으로 구축된 회귀모형을 지점ID 10349 도로지역의 미세먼지농도를 예측하는 데 활용 → 지점ID 10349 도로 지역의 미세먼지농도 추정치와 관측치의 차이를 비교

(Fig. 4)는 모형 이전성 테스트 결과를 나타내며, 미세먼지농도 관측치와 추정치의 유사도(R^2)를 확인한 결과, R^2 는 0.40~0.66 수준으로 나타났다. 이는 한 지역을 대상으로 구축된 모형이 다른 지역의 미세먼지농도의 변동을 약 40~60% 정도 설명할 수 있음을 보여준다. 또한 (Fig. 4)에 나타난 바와 같이 미세먼지농도의 추정치와 관측치의 산점도가 추세선을 중심으로 집중하고 있어, 추정치가 관측치의 추세를 잘 반영하고 있음을 보여준다.

5. 결론

본 연구는 도로주변지역의 미세먼지농도를 추정하는 모형을 구축하였다. 구체적으로 연구대상도로에 설치된 간이 측정기를 통해 수집된 교통량/미세먼지 관측자료와 기상청의 기상자료, 에어코리아의 배경농도 자료를 바탕으로 도로주변지역의 미세먼지농도를 예측하는 회귀모형을 구축하였다. 회귀모형 구축 결과, 모형의 조정 R^2 이 0.381(PM10 모형)에서 0.668(PM2.5 모형) 수준이며, 미시자료를 기반으로 모형이 구축되었음을 고려할 때, 본 연구에서 추정된 회귀모형의 설명력 상당히 높음을 보여준다. 또한 모형구축 결과는 도로주변의 미세먼지농도에 영향을 미치는 가장 중요한 요인은 배경농도이며, 이와 더불어 교통량과 기상조건 또한 유의하게 영향을 미치는 것으로 나타났다. 특히 주목할만한 점은 본 연구에서는 선행 연구에서 제시되지 못했던, 교통량의 미세먼지농도에 대한 한계효과를 제시하였다는 점이다. 예를 들어 시간당 교통량이 1000대 증가함에 따라 도로주변

지역의 PM2.5 농도는 약 2.421 μ g 증가할 수 있으며, PM10 농도는 약 6.769 μ g 증가할 수 있음을 제시하고 있다. 다만, 모형에서 설명하지 못하고 있는 나머지 약 33~62%(즉, 1-조정 R²)는 본 연구에서 고려되지 못한 다른 변수들(예: 차량의 속도와 같은 운행상태, 혼합높이와 같은 기타 기상변수)의 영향으로 보이며, 향후 연구에서는 다음과 같은 기타 변수를 추가하여 모형의 설명력을 추가적으로 높일 필요가 있다.

본 연구는 구축된 회귀모형이 과연 다른 도로주변지역의 미세먼지농도를 추정하는 데 적용이 가능한지에 대한 모형이전성 평가를 수행하였다. 모형 이전성 평가 결과, 본 연구에서 개발된 모형이 다른 지역의 미세먼지농도를 약 40~60% 정도 설명할 수 있음을 보여주고 있다. 또한 회귀모형을 통한 미세먼지농도의 추정치가 실제 미세먼지농도의 추세를 잘 따르고 있음을 확인하였다. 다만, 모형 이전성 평가 결과 설명되지 못한 약 나머지 40~60%는 한 지역을 대상으로 구축된 자료의 변동성이 다른 지역의 변동성을 충분히 반영하지 못할 수 있음을 시사하고 있으며, 따라서 향후 연구에서는 자료 수집의 공간적/시간적 범위를 확장하여, 다양한 도로의 교통량-미세먼지 자료를 수집하고, 이를 통해 모형의 설명력을 더욱 높일 필요가 있다.

다만 본 연구의 강점은 구축된 모형의 활용이 매우 용이하며, 미세먼지농도 추정을 위한 자료 및 분석시간에 대한 요구를 기존 대기확산모델링과 비교하여 매우 크게 줄여줄 수 있다는 점이다. 또한 기존연구에서는 자료의 수집에 요구되는 비용/시간의 한계로 인해 자료의 변동성이 충분히 확보되지 못하여 통계모형의 추정력이 다소 떨어진다는 단점이 제기되었으나, 본 연구에서 개발된 모형은 장시간에 걸친 자료수집을 통해 자료의 변동성이 충분히 확보되어 모형의 추정력이 매우 신뢰할만하다는 점에서 기존 연구와의 차별성이 있다. 본 연구의 결과는 궁극적으로 교통부문에 발생하는 오염물질의 농도를 추정하는 매우 용이한 틀을 실무자/정책가에게 제공하여, 향후 교통부문의 미세먼지를 저감할 수 있는 정책 마련에 도움이 될 것으로 기대한다.

References

- [1] National Institute of Environmental Research (NIER) (2008). Establishment of Climate Change Responding System for Transportation Sector(I) - Development of Bottom-up Type GHGs Emission Factor for Vehicles -, NIER No. 2008-27-977 (in Korean).
- [2] National Institute of Environmental Research (NIER) (2012). Establishment of Climate Change Responding System for Transportation Sector(V) - Development of Bottom-up Type GHGs Emission Factor for Vehicles -, NIER-RP2012-204 (in Korean).
- [3] National Institute of Environmental Research (NIER) (2016). Development of Light Duty Vehicle Emission Factors and Inventories based on Experimental Results over Various Driving Cycles and Operating Conditions(III) - Advanced Study on the Traffic Activity and Regional Allocation -, NIER-SP2016-308 (in Korean).
- [4] U.S. Environmental Protection Agency (USEPA) (2019). MOVES2014b: Latest Version of Motor Vehicle Emission Simulator. <https://www.epa.gov/moves/latest-version-motor-vehicle-emission-simulator-moves> (Accessed: April 10, 2021).
- [5] California Air Resource Board (CARB) (2021). "EMFAC". <https://arb.ca.gov/emfac/> (Accessed: April 10, 2021).
- [6] Intergovernmental Panel on Climate Change (IPCC) (2021). Emission Factor Database (EFDB). <https://www.ipcc-nggip.iges.or.jp/EFDB/main.php> (Accessed: April 10, 2021).
- [7] Health Effects Institute (HEI) (2010) Traffic-Related Air Pollution: A Critical Review of the Literature on Emissions, Exposure, and Health Effects. Special Report 17.
- [8] U.S. Environmental Protection Agency (USEPA) (2021). "Air Quality Dispersion Modeling - Preferred and Recommended Models." <https://www.epa.gov/scram/air-quality-dispersion-modeling-preferred-and-recommended-models> (Accessed: April 10, 2021).
- [9] P. E. Benson(1984). "Caline 4-a dispersion model for predictiong air pollutant concentrations near roadways" (No. FHWA-CA-TL-84-15 Final Rpt). <https://trid.trb.org/view/215944> (Accessed: April 10, 2021).
- [10] G. Lee, J. Hahn(2017). "Analysis Method for Air Quality Improvement Effect of Transport and Environment Policy.", *J. Korean Soc. Transp.*, Vol.35, No.1, pp.37-49. DOI: <https://doi.org/10.7470/jkst.2017.35.1.037>
- [11] C. Yang, Y. S. Koo, I. S. Kim, J. G. Jung, (2013). "Studies on the Methodology of a Hybrid Model for Emission Dispersion Analysis.", *J. Korean Soc. Transp.*, Vol.31, No.2, pp.37-49. (in Korean) DOI: <https://doi.org/10.7470/jkst.2013.31.2.069>
- [12] H. Liu, D. Kim, (2019). "Simulating the Uncertain Environmental Impact of Freight Truck Shifting Programs" *Atmospheric Environment*, 214, 116847. DOI: <https://doi.org/10.1016/j.atmosenv.2019.116847>
- [13] N. Zhang, H. Huang, X. Duan, J. Zhao, B. Su, (2018).

"Quantitative Association Analysis between PM2.5 Concentration and Factors on Industry, Energy, Agriculture, and Transportation." *Scientific Reports*, 8(1), 9461.

DOI: <https://doi.org/10.1038/s41598-018-27771-w>

- [14] H. Liu, X. Xu, M. Rodgers, Y. Xu, R. Guensler (2017). "MOVES-Matrix and Distributed Computing for Microscale Line Source Dispersion Analysis." *Journal of the Air & Waste Management Association*, 67(7). DOI: <https://doi.org/10.1080/10962247.2017.1287788>
- [15] Y. Wu, D. Niemeier(2016). "Strategy of AERMOD Configuration for Transportation Conformity Hot-spot Analysis." Presented at 95th Annual Meeting of the Transportation Research Board, Washington, D.C.
- [16] P. Igri, D. Vondou, F. Kamga(2011). "Case Study of Pollutants Concentration Sensitivity to Meteorological Fields and Land Use Parameters over Douala (Cameroon) using AERMOD Dispersion Model." *Atmosphere*, vol.2, No.4, pp.715-741. DOI: <https://doi.org/10.3390/atmos2040715>
- [17] H. Zhang, Y. Wang, J. Hu, Q. Ying, X. M. Hu (2015). "Relationships between Meteorological Parameters and Criteria Air Pollutants in Three Megacities in China." *Environmental Research*, Vol.140, pp.242-254. DOI: <https://doi.org/10.1016/j.envres.2015.04.004>
- [18] L. H. Tecer, P. Süren, O. Alagha, Karaca, F. Tuncel(2008). "Effect of Meteorological Parameters on Fine and Coarse Particulate Matter Mass Concentration in a Coal-Mining Area in Zonguldak, Turkey." *Journal of the Air & Waste Management Association*, Vol.58, No.4, pp.543-552. DOI: <https://doi.org/10.3155/1047-3289.58.4.543>
- [19] S. Akpınar, H. F. Oztop, E. K. Akpınar(2008). "Evaluation of Relationship between Meteorological Parameters and Air Pollutant Concentrations during Winter Season in Elazığ, Turkey." *Environmental Monitoring and Assessment*, Vol.146, pp.211-224. DOI: <https://doi.org/10.1007/s10661-007-0073-9>
- [20] G. Lin, J. Fu, D. Jiang, J. Wang, Q. Wang, D. Dong(2015). "Spatial Variation of the Relationship between PM2.5 Concentrations and Meteorological Parameters in China." *BioMed Research International*. DOI: <https://doi.org/10.1155/2015/684618>
- [21] J. Wang(2013). "Pearson Correlation Coefficient." In: Dubitzky W., Wolkenhauer O., Cho KH., Yokota H. (eds) *Encyclopedia of Systems Biology*. Springer, New York, NY. DOI: https://doi.org/10.1007/978-1-4419-9863-7_372
- [22] R Core Team(2013). "R: A language and environment for statistical computing." *R Foundation for Statistical Computing, Vienna, Austria*. <http://www.R-project.org/>

정 도 영(Do-young Jung)

[정회원]



- 2008년 2월 : 서울시립대학교
교통공학과 (교통공학석사)
- 2008년 5월 ~ 2012년 1월 : 한국
교통연구원 연구원
- 2022년 8월 : 아주대학교 교통공
학과 (교통공학박사)
- 2014년 8월 ~ 현재 : 한국건설기
술연구원 전임연구원

<관심분야>

교통안전, 도로설계, 자율주행, ITS, 교통공학