

# 소음측정망 데이터를 활용한 통계적 환경소음 예측

## Statistical Environmental Noise Prediction Using Data from a Noise Monitoring Network

김 경 민\* · 김 동 범\*\* · 김 근 한\*\* · 장 서 일†  
Kyoungmin Kim\*, Dongbeom Kim\*\*, Geunhan Kim\*\* and Seo Il Chang†

(Received December 17, 2025 ; Revised February 1, 2026 ; Accepted February 3, 2026)

**Key Words :** Noise Monitoring Network(소음측정망), Statistical Noise Map(통계적 소음지도), Extra-Trees(엑스트라 트리), Shapley Additive Explanations(SHAP, 샤플리 가산 설명)

### ABSTRACT

In this study, we developed a statistical method to predict environmental noise using machine learning models trained on measured data from a noise monitoring network in Seoul, Korea. Daytime and nighttime annual equivalent noise levels were used as dependent variables, and traffic, climate, topographic, landscape, and land-use characteristics were used as explanatory variables. Feature variables were aggregated within buffer distances of 20 to 80 m around monitoring sites to identify the optimal range of influence. Among several models that we evaluated, an extremely randomized trees (Extra-Trees) model showed the highest predictive performance with a coefficient of determination of 0.729 and a root mean square error of 3.4 dB(A) for daytime noise at a buffer radius of 30 m. We then applied Shapley additive explanations (SHAP) to analyze the contribution of each variable, and the results showed that factors related to traffic were the most influential, followed by land-use characteristics. The trained model was applied to a 10 m × 10 m grid to generate a statistical noise map. This study highlights the potential of explainable machine learning-based statistical noise mapping for urban noise management and land-use planning.

### 1. 서 론

우리나라는 「소음·진동관리법」 제4조의 2에 준하여 교통기관 등으로부터 발생하는 소음을 적정하게 관리하기 위하여 필요한 경우에는 일정 지역의 소음 분포 등을 표시한 소음지도를 작성할 수 있다<sup>(1)</sup>. 또한, 소음지도는 조사 목적 및 활용계획에 따라 소음

환경기준 또는 교통소음관리기준 초과지역 현황 파악 등에 요구되는 정밀한 공학적 소음지도와 도시 내 전반적인 소음 노출분포 파악에 용이한 통계적 소음지도로 구분할 수 있다. 여기서, ‘통계적 소음지도’는 소음지도 작성 과정에서 많은 비용과 시간 및 자원을 필요로 하는 공학적 소음지도를 대체할 수 있는 하나의 방안으로 제시된 개념으로<sup>(2)</sup>, 도시 소음과 관련 영향인자인 도시 구성요소와의 관계를 통계

† Corresponding Author ; Member, Dept. of Environmental Engineering, University of Seoul, Professor  
E-mail : schang@uos.ac.kr

\* Member, Dept. of Environmental Engineering, University of Seoul, Ph.D. Candidate

\*\* Korea Environment Institute, Researcher

‡ Recommended by Editor Jiyoung Hong  
© The Korean Society for Noise and Vibration Engineering

적 기법으로 모델링하여 통계적 소음지도를 작성하는 연구가 수행되었으며, 도시 내 주요 공간에서의 소음 노출분포 특성을 빠른 시간에 효율적으로 파악하기에 적합하다<sup>(2-6)</sup>. 이는 기존 물리적 소음 전파 모델링을 기반으로 SoundPLAN 등 소음 예측 소프트웨어를 사용하여 작성하는 공학적 소음지도와는 개념적 차이가 있다.

한편, 소음지도 작성 과정에 활용되는 도로, 건물 및 지형 등의 자료는 소음 계산 모형의 입력자료임과 동시에 도시를 구성하고, 그 형태와 특징을 나타내는 요소로 의미가 있다. 여기에 소음 영향을 받는 지역 인구, 토지이용도 등도 마찬가지로 도시를 구성하고, 형태와 특징을 나타내는 요소이다. 따라서 도시의 형태 및 특징을 나타내는 요소와 소음의 관계를 통계적으로 분석하고, 각 요소가 소음에 미치는 영향을 분석하여 통계적으로 소음지도를 작성할 수 있을 것이다<sup>(2)</sup>.

이에 대한 선행연구로 2011년에 Xie et al.이 중국의 Dalian 도시를 대상으로, 도시 대기질 예측에 활용되어 온 토지이용회귀(land use regression; LUR) 방법을 도시 소음예측모델 개발에 최초로 적용하여 소음지도를 작성하였다<sup>(7)</sup>. 즉, 특정 공간 주변 주거지역 등 토지이용 유형을 기반으로 관련 회귀함수를 통해 소음의 공간적 변동성을 설명하는 연구로, 도시계획 차원의 소음 예측 및 합리적 의사결정을 지원할 수 있는 효과적인 도구로 LUR 모델의 활용 가능성을 보여주었다. 유럽에서 Salomons et al.은 교통소음의 공간적 분포와 도시 형태 등과의 연관성을 규명하였고<sup>(8)</sup>, 2017년 Ryu의 연구에서는, 청주시를 대상으로 도시

형태 지표(지형, 도로, 건물, 교통량 등 8개)가 야간 시간대 도로교통소음 노출에 미치는 영향을 예측소음도 기반의 공간통계모델로 분석하여 유의미한 결과를 도출하는 등 최종적으로 도시계획 단계에서 건물의 배치나 도로망 설계, 토지이용 조정 등을 통해 소음 노출 저감 가능성을 시사하였다<sup>(3)</sup>. 여기서, ‘예측소음도 기반의 공간통계모델’이란 소음도와 도시계획요소 간의 상관관계를 공간통계모형으로 추정하여 도로교통소음을 예측하고 통계적 소음지도를 작성할 수 있는 모델을 의미한다. 또한, 2021년 Kim의 연구에서는 도시를 구성하는 기본 도시형태지표와 예측소음도를 기반으로, 인공신경망과 일반선형회귀 모델인 OLS (ordinary least squares)를 이용하여 통계적으로 소음지도를 작성하였으며, 두 모델 모두 기존 소음지도와 비교해 유사한 결과를 확인하였다<sup>(2)</sup>.

이 연구는 정부의 자동소음측정망 보급 확대 정책과 향후 소음측정망 데이터를 활용한 실시간 소음 모니터링을 위한 선제적 연구로, 서울시의 소음측정망 자료를 활용한 기계학습 기반 통계적 환경소음 예측 모델을 개발하고, 이 모델을 서울시 전역에 적용 및 통계적 소음지도를 제작하여, 도시소음 관리도구로 활용가능성을 검증하고자 한다.

## 2. 연구 방법

연구의 주요 절차는 Fig. 1과 같이 기계학습에 필요한 데이터셋을 구축하고, 기계학습 기반 통계적 환경소음 예측모델 개발 및 검증으로 수행하였다.

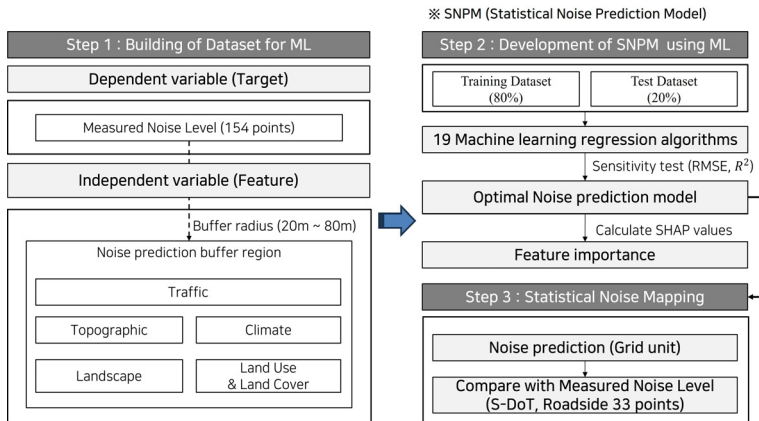


Fig. 1 Research process

### 2.1 데이터셋 구축

연구대상지역은 Fig. 2와 같이 서울시로 설정하고, 기계학습에 필요한 데이터셋을 Table 1과 같이 목표

변수와 특징변수로 구분하여 구축하였다.

#### (1) 목표변수

한국환경공단에서 운영하는 국가소음정보시스템의 환경소음측정망 실측소음도 데이터를 수집하여, 주간(06:00 ~ 22:00) 및 야간(22:00 ~ 06:00) 시간대에 따라 각각 연평균 등가소음도(dB(A))로 가공한 데이터를 사용하였으며, 최종적으로 Fig. 2와 같이 154개 지점(수동 145개, 자동 9개)의 실측소음도를 데이터셋으로 활용하였다. 여기서, 수동측정망은 총 150개 측정지점을 운영하고 있었지만, 일부 데이터가 누락된 5개 지점의 측정자료는 제외하였다.

#### (2) 특징변수

특징변수는 직접적 영향인자와 간접적 영향인자로 구분하여, 다음과 같이 구축하였다.

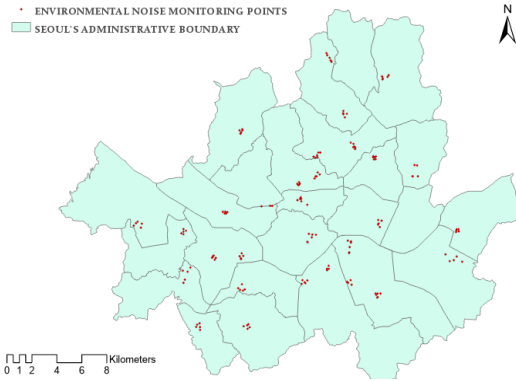


Fig. 2 Study area

Table 1 List of main variables used in the study

Variable	Data		Range	Spatial resolution	Source (year, ref., spatial resolution)
Target	Measured noise levels [Leq, dB(A)]		42 ~ 76	-	Environmental noise measuring network (2021)
Feature	Traffic features	Traffic volume [vehicles/h]	162 ~ 105 001	10 m × 10 m	Korea transport data base (2021)
		Traffic speed [km/h]	2 ~ 96		
		Heavy vehicle ratio [%]	1 ~ 70		
	Climate features	Average Temperature [°C]	9 ~ 14		MK-PRISIM (2017 ~ 2021, 1 km)
		Average Wind speed [m/s]	0.7 ~ 2.2		
	Topographic features	Elevation [m]	0.1 ~ 801		Digital elevation map (2021)
		Slope [°]	0 ~ 65		
	Landscape features	NDVI	-0.1 ~ 0.5		Sentinel-2 (2021, 30 m)
		GNDVI	-0.2 ~ 0.5		
		NDWI	-0.5 ~ 0.2		
LULC features	Urbanized area [%]	52.1	Land cover map (2021)		
	Agricultural area [%]	1.6			
	Forest area [%]	23.5			
	Grassland area [%]	12.1			
	Wetland area [%]	0.3			
	Bareland area [%]	5.1			
	Water area [%]	5.2			

\* NDVI, GNDVI, and NDWI denote the normalized difference vegetation index, green normalized difference vegetation index, and normalized difference water index, respectively.

직접적 영향인자는 교통 특징을 대표할 수 있는 데이터를 의미하며, 「소음지도의 작성방법」에 따라 교통량과 속도, 대형차량 혼입비율로 설정하였다. 관련 교통 데이터는 한국교통연구원에서 관리하는 국가교통데이터베이스를 통하여 Level 5.5 기준 교통량 및 속도데이터와 GIS 네트워크 데이터를 수집하였다. 여기서, 교통량은 연평균 주·야간 시간대를 기준으로 관측교통량의 공간적 범위를 보완할 수 있는 추정교통량을 활용하였으며, 속력은 주·야간 연평균 속도로 설정하였고, 대형차량 혼입비율은 CRTN 소음 예측식 기준 대형과 소형으로 구분하였다.

간접적 영향인자는 「소음지도의 작성방법」에 규정되어 있는 기상 및 지형 조건을 반영할 수 있는 데이터와 이외 경관 및 토지피복 데이터로 설정하여 다음과 같이 구축하였다.

첫째, 기상 데이터는 최근 5년간 연평균 자료를 적용하였으며, MK-PRISM을 이용하여 기상청이 제공하는 공간 해상도 1 km의 격자형 관측자료를 수집한 후, 타 데이터와의 공간적 해상도 일치를 위해 선형 내삽 방식의 통계적 다운 스케일링을 적용하여 10 m 해상도로 격자 재구성을 수행하였다.

둘째, 지형 데이터는 소음 전파에 영향을 미치는 특징변수를 의미하며, 소음지도 작성 시 지형 조건을 생성할 때 필요한 등고선 정보(표고, 경사도)를 사용하였다. 데이터셋으로는 국토지리정보원에서 제공하는 수치표고모델인 DEM을 활용하였다.

셋째, 경관 데이터는 경관 특성이 소음에 미치는 영향을 반영하기 위한 특징변수를 의미하며, 대표적으로 식생지수(NDVI, GNDVI)와 수역지수(NDWI) 등을

사용하였다. 상기 지수는 원격탐사 시 유럽 우주국에서 제공하는 Sentinel-2 위성 이미지로부터 도출하였으며, 데이터의 신뢰도와 재현성을 확보하기 위해 지오빅데이터 오픈플랫폼에서 제공하는 30 m 해상도의 전처리 완료된 시계열 위성 영상을 활용 및 선형 내삽 방식으로 공간 해상도를 10 m 해상도로 변형시켜 적용하였다.

마지막으로, 토지이용 및 토지피복 특성이 소음에 미치는 영향을 분석하기 위하여, 기후에너지환경부의 환경공간정보서비스에서 제공하는 2021년 기준 해상도 10 m의 토지피복지도를 활용하였다.

### 2.2 분석 방법

#### (1) 버퍼 기반 특징변수의 집계

특징변수는 Fig. 3과 같이 목표변수인 환경소음측정망 설치지점을 중심으로 버퍼를 생성하여, 대상 버퍼 반경 내부의 데이터를 집계 및 대표값으로 사용하였다. 여기서, 버퍼의 거리는 다양한 공간 규모에서 노출되는 환경소음에 대한 영향인자의 잠재적 영향을 평가하고 소음 예측 모델의 성능을 최적화할 수 있는 적정 영향범위를 의미하며, 최소 20 m부터 최대 80 m 까지 10 m 단위로 증가시켜 설정하였다.

#### (2) 환경소음 예측 회귀모형 탐색

통계적 환경소음 예측 모델 개발을 위하여, 환경소음측정망 실측소음도와 영향인자로 구성된 학습 데이터를 사용하여 모형을 학습시키고, 특징 변수들과의 관계를 바탕으로 연속적인 소음도를 예측하는 지도학습 회귀모형을 사용하였으며, 소음 예측에 적합한 최

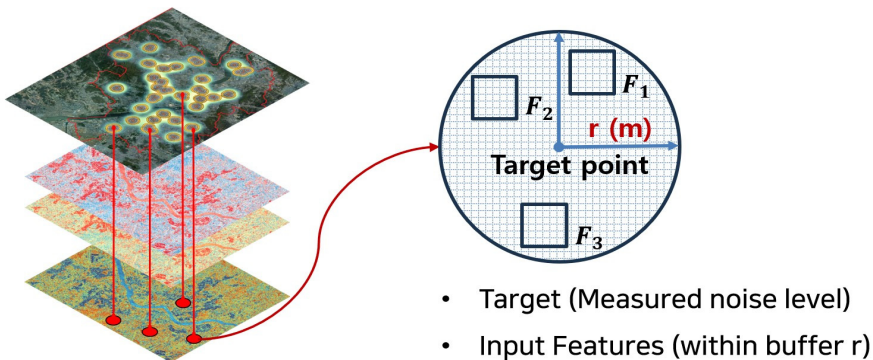


Fig. 3 Estimated features within buffer

적의 모델을 탐색하기 위하여 Table 2와 같이 총 19종의 회귀모형을 선정 및 모형 간 소음 예측성능을 비교·평가하고자 하였다. 여기서, 비교적 간단한 선형회귀모형(OLS)을 기초 모델로 설정하고, 소음 예측 등 비선형 관계의 회귀분석에 적합한 트리 구조 계열의 모형(Extra-Trees, Random Forest, XGBoost)을 후보 모델로 우선 검토하였다. 이외, 신경망 계열의 지도학습 회귀분석을 위한 클래스로 MLPRegressor(multi-layer perceptron)를 포함하였으며, Scikit-learn과 LightGBM, CatBoost 라이브러리 등에서 제공하는 다양한 회귀모형을 활용하였다.

(3) 특징변수 기여도 분석

통계적 환경소음 예측결과에 영향을 미치는 특징변수의 기여도를 분석하기 위하여, SHAP(Shapley additive explanations) 기법을 적용하였다<sup>9)</sup>. SHAP은 미국의 수학자인 Lloyd Stowell Shapley가 제안한 게임 이론의 Shapley value에 기반하여, 특징변수별로 모델 예측에 기여하는 정도(증가 또는 감소)를 Shapley value로 계산하고 이를 합산하여, 최종 예측 결과인 목표변수에 대한 특징변수 간의 상대적 영향도(+/- 방향)를 시각적으로 설명하는 기법을 의미한다.

3. 연구 결과

3.1 Extra-Trees 기반 환경소음 예측모형 선정

환경소음측정망 설치지점으로부터 버퍼 반경 20 m에서 80 m까지 특징변수를 집계하는 등 관련 데이터셋으로부터 hold-out 방식으로 모형별 훈련(80%) 및 테스트(20%)를 수행하였다.

주간 시간대 기준 모델별 성능평가 결과, Table 3 및 Fig. 4와 같이 버퍼 반경 20 m ~ 30 m 범위에서 트리 구조 계열의 모델(Extra-Trees, Random Forest, XGBoost)이 가장 우수한 성능을 보였으며, 버퍼 반경 40 m 이후부터는 성능이 저하되는 추세를 보인다.

특히, Extra-Trees(extremely randomized trees) 모델의 경우, 버퍼 반경 30 m 기준 RMSE 오차값과 R<sup>2</sup> 값이 각각 3.4 dB(A)와 0.729로 실제 환경소음 데이터의 변동 패턴을 가장 잘 설명하였으며, 실측소음 대비 예측오차도 평균적으로 가장 낮았다. 이에 이 연구에서는 환경소음의 공간적 변동성을 가장 잘 반영한 Extra-Trees 모형을 통계적 환경소음 예측 모델로 최종 선정하였다. 여기서, Extra-Trees 모형은 2006년 Geurts 등이 제안한 의사결정 트리(decision tree) 기

Table 2 Candidate supervised learning regression algorithms

Library	Model family		Class (regressor)
Scikit-learn	Linear models		Linear regression (ordinary least squares) Ridge, Lasso, Elastic-net, Bayesian ridge
	Kernel methods		Kernel ridge
	Support vector machines		SVR (support vector regression)
	Ensembles	Tree-based ensembles	ExtraTrees regressor, Random forest regressor, Gradient boosting regressor, HistGradient boosting regressor
		Meta-ensembles	Voting regressor, stacking regressor
	Nearest neighbors		KNeighbors regressor
	Gaussian processes		Gaussian process regressor
Neural networks		MLP regressor	
XGBoost	Gradient boosting (tree)		XGB regressor
LightGBM			LGBM regressor
Catboost			CatBoost regressor

반의 앙상블 학습 기법으로<sup>(10)</sup>, 각 트리에 사용하는 특징변수 데이터를 무작위로 선택하고 트리별 내부 분할 과정에서 무작위성을 극도로 크게 부여하여 트리 간의 상관성을 낮출 수 있으며, 다수 의사결정 트리를 구축 및 결합하여 예측성능을 안정적으로 향상시킬 수 있는 구조를 갖고 있어, 소음 예측 시 과적합을 완화하면서 일반화 성능을 높일 수 있다.

### 3.2 일반화 성능평가

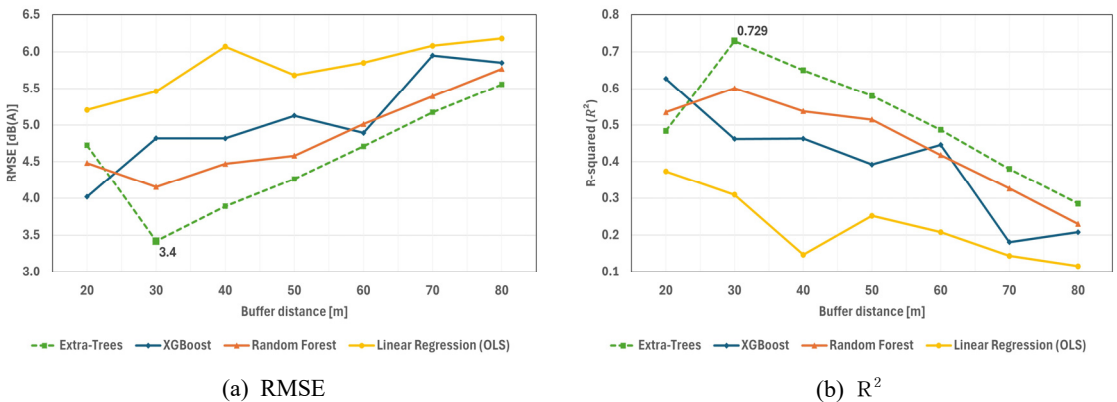
환경소음측정망과 같은 공간 데이터는 인접 지점끼리 소음도가 비슷한 공간적 자기상관(spatial autocorrelation)이 강한 특성을 갖고 있다. 이에 따라, 예측모델에 대한 성능평가 시, 훈련 데이터와 테스트 데이터 간의 공간적 이격거리에 따라 과대평가될 가능성이 존재한다. 따라서, Extra-Trees 모형 기반 통계적 환경소음 예측 모델에 대한 성능평가는 훈련 및 테스트

과정을 여러 번 다른 데이터셋을 활용하여 반복적으로 수행하는 교차검증 방식의 일반화 성능평가로 진행하였다. 여기서, 교차검증 횟수는 일반화 성능평가의 신뢰도와 계산 부담을 절충할 수 있는 변수로 의미가 있으며, 일반적으로 5회 또는 10회를 적용한다. 이 연구에서는 계산 효율 대비 일반화 성능평가에서 중요한 검증 데이터를 안정적으로 확보하기 위하여 교차검증 횟수를 5회(훈련 80%, 검증 20% 구조)로 설정하였다.

일반화 성능평가 결과, Table 4(하이퍼파라미터 튜닝 전)와 같이  $R^2$  값은 평균 약 0.4, RMSE 오차값은 약 4.8 dB(A)로 실측값 대비 5 dB(A) 이내의 오차를 확인할 수 있다. 다만, 초기 성능평가 결과에 비해 설명력이나 정확도가 낮아졌으나, 이는 실제 적용 환경에서 기대할 수 있는 예측성능을 보수적으로 반영한 결과로, 통계적 소음지도의 실무적 활용 가능성을 평가하는 데 의미가 있다.

**Table 3** Model performance metrics across buffer distances for daytime noise prediction

Buffer distance [m]	RMSE [dB(A)]				$R^2$			
	Extra-Trees	XGBoost	Random forest	Linear regression (OLS)	Extra-Trees	XGBoost	Random forest	Linear regression (OLS)
20	4.7	4.0	4.5	5.2	0.484	0.627	0.535	0.373
30	3.4	4.8	4.2	5.5	0.729	0.462	0.601	0.309
40	3.9	4.8	4.5	6.1	0.650	0.463	0.538	0.147
50	4.3	5.1	4.6	5.7	0.580	0.392	0.515	0.252
60	4.7	4.9	5.0	5.9	0.487	0.446	0.418	0.208
70	5.2	5.9	5.4	6.1	0.380	0.181	0.326	0.144
80	5.6	5.9	5.8	6.2	0.285	0.208	0.230	0.116



**Fig. 4** Variation in model performance with buffer distance for daytime noise prediction

### 3.3 예측성능 최적화

소음 예측성능이 높은 최적의 하이퍼파라미터 조합을 탐색 및 Extra-Trees 모델에 적용하여 예측성능을 최적화하였다. 여기서, 하이퍼파라미터는 모형의 학습 방식을 제어하는 역할을 하는 설정값으로, 기계학습을 하기 전에 수동으로 설정하거나 기본값으로 적용된다.

Table 4(하이퍼 파라미터 튜닝 후)와 Table 5는 Extra-Trees 모델에 대한 하이퍼파라미터 설정값을 조정 한 결과로, R<sup>2</sup> 값이 높고 RMSE 오차값이 낮은 하이퍼파라미터 설정값의 조합으로 결정하였다. 또한, 하이퍼파라미터 조정에 따른 과적합을 방지하기 위하여 5회 교차검증을 병행하였다. 이외, Table 5에 제시한 하이퍼파라미터 설정값은 특정값으로 변경했을 때와 기본값으로 설정했을 때의 예측성능 차이가 없는 경우는 기본값을 그대로 적용하였다.

### 3.4 예측소음도 주요 기여인자 분석

예측소음도에 대한 특징변수별 기여도 분석결과, Fig. 5와 같이 왼쪽의 막대 그래프는 소음 예측에 영

향을 미치는 특징변수들의 기여도를 절대값 크기로 표현하였으며, 오른쪽의 beeswarm plot은 영향의 방향을 가로축의 SHAP value 수치(+/증가, -/감소)와 세로축의 feature value 색상(빨강/변수값 증가, 파랑/변수값 감소)으로 표현하였다.

먼저, 도로교통 특징변수는 속도, 교통량, 대형차량 혼입률이 높을수록 소음 증가에 가장 영향을 많이 미치는 중요한 변수로 도출되었다. 즉, 도로의 분포가 많고 관련 교통 활동이 증가함에 따라 도로변 소음도 증가에 가장 영향을 많이 미치는 것을 확인할 수 있다. 이외, 야간에는 교통량이 줄어들고 화물차량 비중이 증가하여 대형차량의 혼입률 영향이 교통량보다 더 크게 기여하는 결과를 보인다.

토지피복 특징의 경우에는 상업(commercial), 초지(grassland), 주거지역(residential)의 분포가 도로 교통 다음으로 소음도에 영향을 미쳤다. 즉, 주거지역과 초지지역의 분포가 많은 지역일수록 주변 소음도를 감소시키는 반면에, 업무시설이 많은 상업지역 일수록 소음도를 증가시키는 요인으로 작용하였다.

이외, 지형/경관/기상학적 특징은 토지피복 다음으로 영향을 미치는 변수로 분석되었으나, 상대적으로 기여도가 낮고 영향의 방향도 명확하지 않았다.

Table 4 Cross-validation performance (L<sub>day</sub>)

Hyper-parameter tuning	K-fold	R <sup>2</sup>	RMSE [dB(A)]
Before	1	0.7335	3.3934
	2	0.1127	6.1003
	3	0.2436	5.5421
	4	0.5412	4.7801
	5	0.5423	4.2991
	Average CV	0.4346	4.8230
After	Average CV	0.4640	4.7313

Table 5 Hyper-parameters of Extra-Trees model

Hyper-parameter	Value
max_depth	None*
max_features	0.5
min_samples_leaf	2
min_samples_split	2*
n_estimators	200

\* Reference values

### 3.5 통계적 소음지도 작성 및 검증

Extra-Trees 모형 기반 최적의 소음 예측 모델을 서울특별시 전역에 적용하여, Fig. 6과 같이 주-야간 시간대 통계적 소음지도를 작성하였으며, 소음지도의 격자 크기는 「소음지도 작성방법」에 따라 평면 소음 지도 계산 시 10 m × 10 m 단위를 적용하였다.

또한, 통계적 소음지도 작성에 따라 계산된 소음도 예측값은 「소음지도 작성방법」에 따라 실측값과 비교하여 평균오차 및 표준편차를 검토하였으며, 소음도 실측값은 서울시에서 자체적으로 운영하고 있는 스마트서울 도시데이터 센터(smart seoul data of things; S-DoT) 자료를 활용하였다.

여기서, S-DoT은 서울시 전역(주거지역, 상업지역, 공업지역, 도로, 공원 등)을 대상으로, 소음을 포함한 도시의 다양한 환경정보를 수집하기 위한 복합 센서를 의미하며, 2019년부터 현재까지 약 1100개소에 설치되어 2분 간격으로 수집되고 있다. 이 연구에서는 주로 도로변에 설치되어 있는 환경소음측정망 기반 예측소음도와 비교하기 위하여, Fig. 7과 같이 2021년

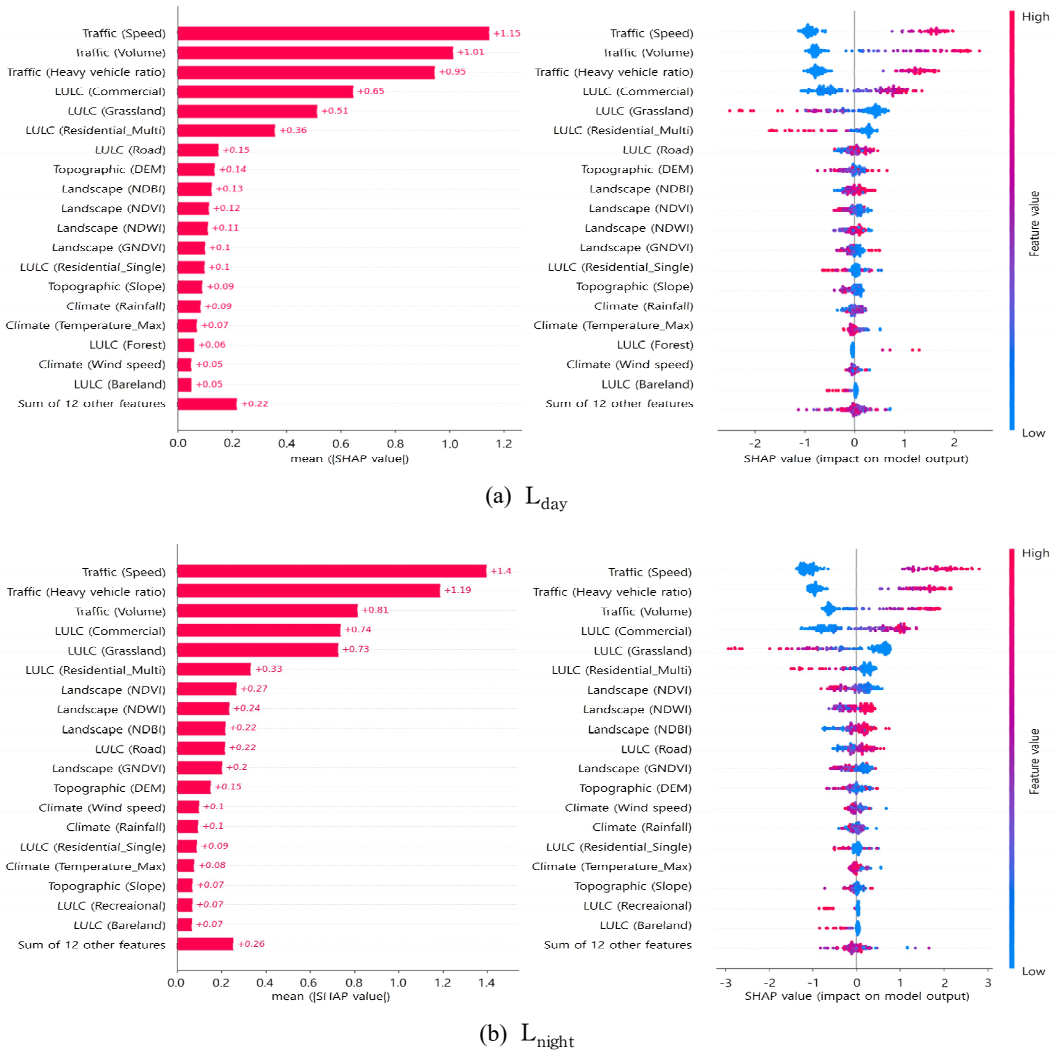


Fig. 5 Feature importance and direction based on SHAP values

기준 주요 도로에 설치되어 있는 고정식 S-DoT 33개 지점의 시간 평균 실측소음도를 수집하였으며, 주·야간 시간대(주간 06:00 ~ 22:00, 야간 22:00 ~ 06:00)를 구분하여 연평균 등가소음도(dB(A))로 가공한 데이터를 활용하였다.

Fig. 7과 같이 주요 도로변 총 33개 지점을 대상으로 통계적 소음지도에서 예측한 소음도 결과와 S-DoT 실측값을 비교한 결과, 「소음지도의 작성방법」 제5조 제2항 관련 [별표 2] 소음지도의 작성방법에서 고시하는 검증기준에 따라 Table 6과 같이 오차의 평균 (bias)이 +2 dB 이내로 실측값 대비 평균적으로 과대

추정하는 경향을 보이며 기준치인  $\pm 3$  dB을 만족하였지만, 표준편차는 기준치인 3을 초과하였다. 다만, 표준편차의 경우 기준치를 약간 상회하는 수준(주간 3.4, 야간 3.1)으로, 특정 지점별로는 Fig. 8과 같이 55 dB(A) 이내의 일부 저소음 영역에서 오차가 크게 발생하였지만, 전체 지점 측면에서는  $\pm 5$  dB(A) 오차 범위 내에 80% 이상 분포하고 있어 비교적 안정적인 예측 결과를 보인다.

시간대별로는 야간 시간대의 RMSE 오차와 표준편차가 주간에 비해 낮아 더 안정적이고 우수한 예측 성능이 확인되었다. 특히,  $\pm 5$  dB(A) 오차 범위 내에

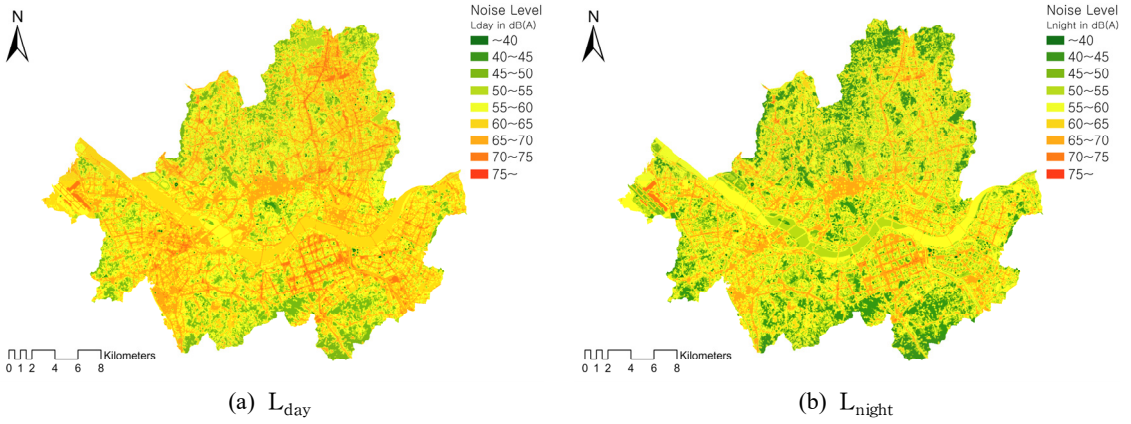


Fig. 6 Statistical noise map in Seoul

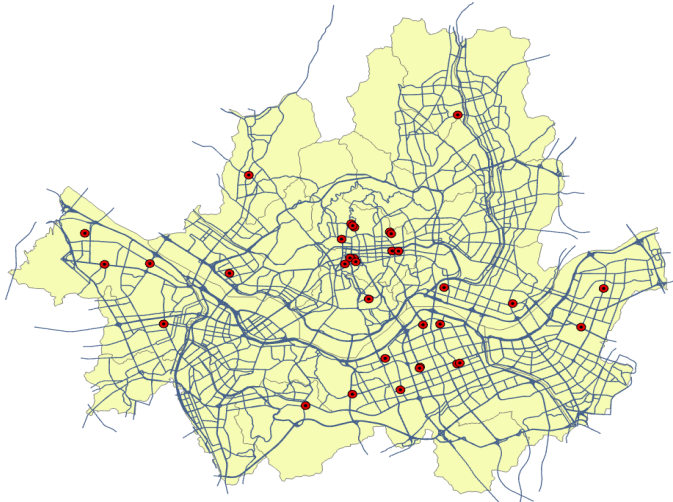


Fig. 7 Validation sites based on S-DoT noise measurements along the main road

Table 6 Results of cross analysis between measured and predicted noise level from Extra-Trees model [unit: dB(A)]

Time zone	Samples		Verification criteria		Probability for difference range [%]		RMSE	R <sup>2</sup>
	N	Range	Bias	Standard deviation	±3 dB(A)	±5 dB(A)		
Day	33	50.5 ~ 71.9	+2.0	3.4	63.6	84.8	3.9	0.733
Night	33	47.3 ~ 66.7	+1.6	3.1	57.6	90.9	3.4	0.727

분포할 확률이 야간(90.9%)에 주간(84.8%)보다 높게 나타나 야간 시간대에 실측값과 더 일관되게 예측하는 경향을 보였다. 이는 도로교통소음에 영향을 미치는 주요 소음원인 교통량이 야간에 낮게 발생하는

등 교통 활동의 변동성 감소에 따른 소음 발생 패턴이 단순화되고, 이에 따라 주변 공간 특징 기반의 예측 모델이 소음 노출 수준을 더 안정적으로 설명할 수 있기 때문으로 해석될 수 있다.

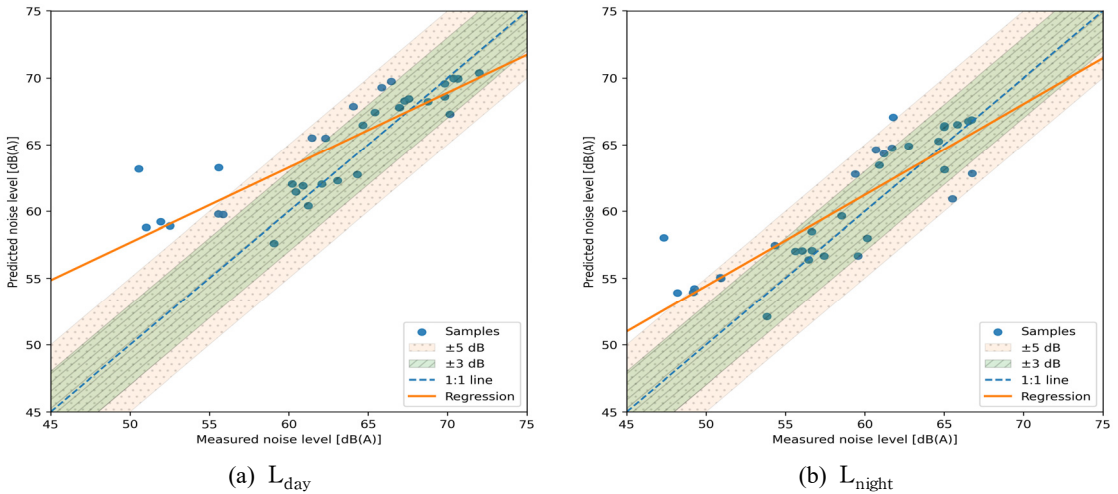


Fig. 8 Scatter plots between measured and predicted noise level from Extra-Trees model

4. 논의 및 결론

이 연구에서는 환경소음측정망을 이용하여 기계학습 기반 통계적 환경소음 예측모델을 개발하고, 대상 도시 전역에 대한 통계적 소음지도를 작성 및 도로변 실측소음도와 비교·검증하여, 소음 모니터링에 활용할 수 있는 수준의 예측성능을 확인하였다.

통계적 소음 예측 시 선행연구에서는 토지이용계획 정보를 입력자료로 활용하고 SoundPLAN 모델링 결과를 출력자료(라벨)로 활용하여(2,4), 계획과 예측소음도에 기반한 불확실한 정보를 함께 학습할 가능성이 존재한다. 반면에, 이 연구에서는 매년 갱신되는 토지피복지도를 활용하여 실제 지표면의 특성을 반영하고 소음측정망에서 관측된 실측소음도를 기계학습에 활용하여, 대상도시의 실제 토지피복 환경에서 노출되는 환경소음의 공간적 변동성을 현실적으로 예측하였다는 점에서 차별성을 가진다.

소음측정망 기반 통계적 환경소음 예측모델 설계 시 적합한 모형을 선정하기 위하여, 기계학습에서 활용할 수 있는 지도학습 회귀모형을 비교·평가하였다. 그 결과, 비선형성과 변수 간의 상호작용을 효과적으로 학습할 수 있는 트리 구조 계열의 모형이 선형회귀모형에 비해 우수한 예측성능을 보였으며, 특히 버퍼 반경 30 m 기준 Extra-Trees 모형이 설명력과 예측오차 측면에서 가장 안정적인 성능을 보였다. 여기서, 변수 간의 상호작용이란 교통량 등 특정 인자의

소음 영향이 독립적으로 작용하지 않고 주변 토지의 분포 유형과 같이 도시를 구성하는 다수의 영향인자에 따라 달라지는 특성을 의미하며, 이는 교통, 토지 이용 등의 소음 영향인자가 복합적으로 작용하는 도시 환경소음의 특성을 반영한다는 의미가 있다. 또한, 버퍼 반경 30 m는 서울시의 도로변 환경소음과 주변 소음 영향인자 간의 관계를 가장 효과적으로 설명할 수 있는 통계적으로 최적화된 공간 규모를 의미한다.

통계적 소음지도의 검증을 위하여 서울시 주요 도로변에 설치된 S-DoT 실측소음도와 비교 결과, 통계적 소음지도에서 예측한 소음도 결과는 실측값과의 허용오차 기준치를 만족하는 수준이다. 특히 실측값 대비 평균적으로 과대추정되는 경향을 보였으며, 도로변 환경소음 기준치인 55 dB(A) 이상 영역에서는 오차가 비교적 작게 나타났다. 이는 잠재적인 소음 영향까지 보수적 평가를 가능하게 하고, 도로변 등 고소음 노출지역을 효과적으로 선별 관리할 수 있는 도시 소음 모니터링에 적합한 도구로서 소음저감대책 등 관련 의사결정 우선순위 설정에 있어 통계적 소음지도의 실용성을 강화한다.

소음 영향인자의 예측 기여도(SAHP) 분석결과, 토지 피복 특성은 교통 특징 다음으로 소음에 큰 영향을 미치는 주요 인자로 분석되었다. 이는 주변 토지이용 환경에 따라 실제 소음 노출 수준이 달라질 수 있다는 점을 시사한다. 즉, 도시계획 관점에서 향후 토지이용계획이나 개발계획 수립 시 교통 시나리오뿐만 아니라 주변 토지이

용의 차이에 따라 예상되는 소음 노출 수준을 비교·검토하는 등 적정 공간 구조 설계를 통한 사전 예방적 관리로 규모 통계적 소음지도 활용을 시사한다.

또한, 이 연구는 통계적 소음지도 작성을 통하여 기존 공학적 소음지도의 갱신 한계를 보완하고, 도시 소음 노출 분포 모니터링 및 소음측정망과 연계 운영할 수 있는 실용적 도구로 활용될 수 있다.

다만, 이 연구는 도로변 소음측정망 기반의 공간적으로 제한된 통계적 환경소음 예측모델로서, 도시 규모의 토지용도 지역별 환경소음을 충분히 반영하는 데에는 한계를 지니며, 기존 정밀한 공학적 소음지도와는 활용범위를 구분할 필요가 있다. 또한, 이 연구에서 도출된 최적의 버퍼 반경은 서울시의 도로변 환경소음측정망을 중심으로 주변 도시 구성요소에 의존한 결과로, 다른 도시로의 일반화에는 추가적인 검증이 필요하다. 향후 연구에서는 IoT 소음측정망과 연계한 상시 갱신형 통계적 소음지도로의 확장 및 도시 구성요소 특징이 다른 도시를 대상으로 모델을 일반화할 필요가 있다.

최종적으로 통계적 환경소음 예측은 기계학습에 활용하는 데이터셋의 공간적 분포와 신뢰할 수 있는 수준의 품질에 강하게 의존하는 데이터 기반의 접근방법이다. 따라서 통계적 소음 예측결과의 일관성과 정확도를 확보하기 위해서는 측정지점의 공간적 대표성을 갖춘 충분한 양의 소음도 자료와 예측 대상인 소음도의 공간적·시간적 변동 특성을 효과적으로 반영할 수 있는 특징변수 데이터의 안정적인 확보가 정책적으로 지속 가능한 통계적 소음지도 구축의 핵심 전제조건이라 할 수 있다.

## 후 기

이 논문은 정서 교란 소음의 성가심 영향분석 연구 사업의 지원을 받아 수행되었습니다(No. 202505142001).

## References

(1) Korea Ministry of Government Legislation, 2025,

Noise and Vibration Control Act, Korea Ministry of Government Legislation, Act No. 21065, Seoul, Korea.

(2) Kim, P. L., 2021, Artificial Neural Network Model for Statistical Road-traffic Noise Mapping Based on Elementary Urban Form Indicators, Doctoral Dissertation, University of Seoul, Seoul, Korea.

(3) Ryu, H. J., Park, I. K., Chun, B. S. and Chang, S. I., 2017, Spatial Statistical Analysis of the Effects of Urban Form Indicators on Road-traffic Noise Exposure of a City in South Korea, Applied Acoustics, Vol. 115, No. 1, pp. 93~100.

(4) Ryu, H. J., 2017, Spatial Statistical Analysis of Road-traffic Noise for Supporting Urban Planning in the Strategic Stage, Doctoral Dissertation, University of Seoul, Seoul, Korea.

(5) Kim, P. L., Ryu, H., Jeon, J. J. and Chang, S. I., 2019, Artificial Neural Network Model Development based on Road-traffic Noise and Urban Form Indicators, Transactions of the Korean Society for Noise and Vibration Engineering, Vol. 29, No. 5, pp. 577~583.

(6) Kim, P. L., Ryu, H., Jeon, J. J. and Chang, S. I., 2021, Statistical Road-traffic Noise Mapping Based on Elementary Urban Forms in Two Cities of South Korea, Sustainability, Vol. 13, No. 4, 2365.

(7) Xie, D., Liu, Y. and Chen, J., 2011, Mapping Urban Environmental Noise: A Land Use Regression Method, Environmental Science and Technology, Vol. 45, No. 17, pp. 7358~7364.

(8) Salomons, E. M. and Pont, M. B., 2012, Urban Traffic Noise and the Relation to Urban Density, Form, and Traffic Elasticity, Landscape and Urban Planning, Vol. 108, No. 1, pp. 2~16.

(9) Lundberg, S. M. and Lee, S. I., 2017, A Unified Approach to Interpreting Model Predictions, Proceedings of the 31st International Conference on Neural Information Processing Systems, pp. 4768~4777.

(10) Geurts, P., Ernst, D. and Wehenkel, L., 2006, Extremely Randomized Trees, Machine Learning, Vol. 63, No. 1, pp. 3~42.