

논문

그래프 신경망 기반 고해상도 $\text{NO}_2\text{-O}_3$ 추론장의 화학 정합성 평가: 수도권-강원권 지역의 지표 산화제 분배를 중심으로

Chemical Consistency of GNN-Inferred High-Resolution $\text{NO}_2\text{-O}_3$ Fields: Surface Oxidant Partitioning in the Seoul Metropolitan Area-Gangwon Region

정필수, 손승민, 왕경희, 김현수, 최대련¹⁾, 윤희영^{1),*}

안양대학교 일반대학원 환경공학과, ¹⁾ 안양대학교 환경에너지공학과

Peel-Soo Jeong, Seung-Min Sohn, Kyung-Hui Wang, Hyeun-Soo Kim,
Dae-Ryun Choi¹⁾, Hui-Young Yun^{1),*}

Department of Environmental Engineering, Graduate School of Anyang University,
Anyang, Republic of Korea

¹⁾Department of Environmental & Energy Engineering, Anyang University, Anyang,
Republic of Korea

접수일 2026년 4월 14일
수정일 2026년 5월 15일
채택일 2026년 6월 2일

Received 14 April 2026

Revised 15 May 2026

Accepted 2 June 2026

*Corresponding author

Tel : +82-(0)31-5183-2135

E-mail : huiyoung@anyang.ac.kr

Abstract This study investigates a 1 km $\text{NO}_2\text{-O}_3$ concentration field inferred by a deep learning (DL) graph neural network (GNN) trained solely on publicly available reanalysis, emission inventories, and ground monitoring data over the Seoul Metropolitan Area-Gangwon region. Temporal holdout validation over the full year 2023 confirmed high predictive performance, with R^2 of 0.892 for NO_2 and 0.926 for O_3 . However, existing evaluations of high-resolution air quality fields have largely relied on such statistical fit metrics, leaving the chemical consistency of off-station spatial patterns insufficiently examined. To address this gap, we applied supplementary diagnostics based on Ox conservation and partitioning, $\text{NO}_x\text{-O}_x$ regression, weekend ozone effect, urban-forest temporal structure, and satellite formaldehyde-to- NO_2 ratio (FNR). The data-driven model incorporating physics-based prior information captured a broad west-east transition of the surface oxidant partitioning ratio ($r = \text{NO}_2/\text{O}_x$), from strong NO_2 titration in central Seoul toward an O_3 -dominant background in Gangwon forests. The urban-forest contrast of O_3 diurnal amplitude matched station observations, and the spatial pattern of the weekend ozone effect was consistent with the expected chemical response. The 1 km partitioning map further showed consistent negative correlations with both GEMS and TROPOMI FNR (Spearman $\rho \approx -0.5$), providing supplementary evidence in observation-sparse areas. Interannual inference for 2020~2023 showed spatial changes in line with regional NO_x emission trends. However, since the oxidant partitioning ratio r is not a direct measure of ozone sensitivity, these results should be viewed as a chemically informed consistency check, not as a validation of off-station absolute accuracy or forecast skill. Overall, these multi-source diagnostics support DL-inferred concentration fields as a chemically coherent basis for interpreting air quality in observation-sparse regions.

Key words: High-resolution air quality inference, Graph neural network, Chemical consistency, Surface oxidant partitioning, Formaldehyde-to- NO_2 ratio

1. 서론

대기오염의 장기적 분석은 측정소의 수와 공간 배

치에 따라 그 신뢰성이 크게 달라질 수 있다(Lee and Shin, 2025). 국내 지방자치단체 대기오염측정망(2024년 12월 말 기준) 674개소 중 도시대기측정소가 531개

소(약 80%)를 차지하며(NIER, 2025) 강원·경북 등 비도시 지역은 측정소 밀도가 극히 낮아 지역 대기환경 평가의 불확실성이 크게 증가한다. 이러한 관측 공백을 해소하기 위해서는 시공간적으로 연속된 고해상도 농도장 산출이 필요하다. 특히 1 km 이하 고해상도 농도장은 환경보건 코호트 연구의 개인 노출 추정, 대기정책의 고농도 우심지역 식별, 환경영향평가의 사업 단위 영향 분석 등 다양한 분야에서 수요가 꾸준히 증가하고 있으며 이러한 요구에 대응하여 화학수송모델(Chemical Transport Model, CTM)·토지이용회귀(Land Use Regression, LUR)·국지 확산모델 하이브리드(Oh *et al.*, 2021), 국지 확산모델과 이중 지상 관측 동화를 결합한 도시 규모 고해상도 매핑(Mijling, 2020) 등 다각적 접근이 활발히 시도되어 왔다. 다만 기존 접근은 LUR의 비선형 상호작용 포착 한계(Hoek *et al.*, 2008), CTM의 배경농도 편향과 분해능 제약, 국지 확산모델의 높은 계산 비용 등이 구조적 제약으로 남아 있다.

최근에는 딥러닝(deep learning, DL) 기반 고해상도 대기질 농도장 추론이 유력한 대안으로 부상하고 있다(Choi *et al.*, 2025). Shen *et al.* (2024)은 CTM 시뮬레이션을 물리 기반 사전정보로 딥러닝에 통합하여 관측 희소 지역을 포함한 전지구 PM_{2.5} 농도장을 추정하는 바 있으며 정규 격자 대기질 자료의 딥러닝 다운스케일링도 다각적으로 시도되어 왔다(Gomez Gonzalez, 2023). 다만 정규 격자 입력을 가정하는 전형적 접근인 합성곱 신경망(Convolutional Neural Network, CNN)은 측정소가 불규칙하게 분포하고 관측 자료(점)와 격자 자료(면)의 해상도·형태가 이질적인 자료 환경의 대기질 추론에는 한계가 있다. 이에 비균질한 측정소 네트워크의 구조를 직접 활용하고 이중·다해상도 자료를 단일 체계에서 융합할 수 있는 그래프 신경망(Graph Neural Network, GNN) 기반 접근이 주목받고 있다(Terroso-Saenz *et al.*, 2024; Li *et al.*, 2023). GNN은 측정소, 격자, 광역 노드를 이중 그래프(heterogeneous graph) 위에서 계층적 메시지 패싱으로 연결하여 관측 정보를 미관측 격자로 일관되게 전파할 수 있다.

한편, 대류권 오존(O₃)은 질소산화물(NO_x)과 휘발성유기화합물(VOCs)의 광화학 반응으로 생성되는 대표적 2차 오염물질이다. 국내에서는 NO₂, SO₂ 등 1차 오염물질 농도가 장기적으로 감소해 온 반면 오존 농도는 지속적으로 증가하는 추세를 보여왔다(Park *et al.*, 2024; Jeong *et al.*, 2022; Yeo and Kim, 2021). 이는 O₃ 생성이 전구물질 농도에 대해 강한 비선형성을 갖기 때문에 특히 VOC 제한(VOC-limited) 환경에서는 NO_x 배출 저감이 NO에 의한 O₃ 적정(titration) 효과를 약화시켜 오히려 O₃ 농도를 상승시킬 수 있다(Colombi *et al.*, 2023). 효과적인 O₃ 저감 정책 수립을 위해서는 대상 지역의 오존 생성 민감도 체제에 대한 공간적 이해가 선행되어야 한다(Lee *et al.*, 2020; Kim *et al.*, 2019). 수도권은 도시의 NO_x 적정 우세에서 교외의 광화학 O₃ 생성 기여로 이어지는 도심-교외 간 민감도 차이가 보고되어 있으며(Baek *et al.*, 2023; Jeong *et al.*, 2022) 최근 연구에서는 한국의 대부분 지역이 NO_x 제한에 가깝고 VOC 제한은 수도권 핵심부에 잔존하는 것으로 보고되고 있다(Oak *et al.*, 2025; Lee *et al.*, 2024b). 반면 강원 산림·산악권은 인위적 NO_x 배출이 희소하여 배경 O₃가 우세한 환경이나 관측망이 극히 희소하여 공간적 민감도 진단이 어려운 실정이다(Lee and Shin, 2025; Jeong *et al.*, 2022).

이러한 민감도 체제의 공간 진단에는 위성 기반 HCHO/NO₂ 컬럼 비(formaldehyde-to-NO₂ ratio, FNR)(Duncan *et al.*, 2010; Martin *et al.*, 2004)와 CTM 기반 배출 민감도 분석이 주로 활용되어 왔으나, FNR은 임계값의 지역적 불확실성(Tian *et al.*, 2025; Jin *et al.*, 2020)과 위성 픽셀 크기의 분해능 제약이 내재하며 CTM은 방대한 계산 비용과 배출 인벤토리 불확실성이 수반된다. 이중·다해상도 자료를 융합하는 DL 기반 고해상도 농도장 추론은 이러한 분해능·계산 비용의 한계를 보완할 잠재력을 가지나, 기존 DL 대기질 연구의 다수는 측정소 위치에서의 점 단위 통계 지표(R², RMSE 등)에 의존하여 학습·검증을 수행해 왔다. 명시적 물리·화학 메커니즘이 부재한 자료 기반 접근은 관측망 밖 격자에서 물리적으로 일관되지 않은 결과를

산출할 수 있어(Shen *et al.*, 2024; Li *et al.*, 2023) 측정소가 희소한 지역에서는 점 단위 지표만으로 산출장의 공간 분포 타당성을 담보하기 어렵다. 이러한 구조적 제약은 점 단위 오차 지표를 보완할 대기화학적 경향 기반 보조 검증 전략의 필요성을 시사한다.

본 연구의 목적은 GNN 기반 1 km $\text{NO}_2 \cdot \text{O}_3$ DL 추론 농도장이 측정소가 없는 지역에서도 대기화학적으로 타당한 공간 패턴을 산출하는지를 점검하는 데 있다. 본 연구는 오존 생성 민감도를 직접 판정하기보다 산출장이 알려진 대기화학적 경향과 부합하는지를 검토하며 주요 내용은 다음과 같다. 첫째, 지표 산화제 농도 구조의 화학적 정합성을 점검한다. 토지피복별 Ox 보존성, NO_x - Ox 회귀 기울기, 주말-주중 차이 및 도시-산림의 O_3 일변화 구조를 통해 DL 농도장의 O_3 공간·시간 분포가 배출 강도와 시간 변동이 다른 시공간 맥락에서 대기화학적으로 기대되는 패턴과 일치하는지 살펴본다. 둘째, 1 km 해상도 지표 산화제 분배 구조의 외부 정합성을 평가한다. 분배비 $r(=\text{NO}_2/\text{Ox})$ 로 산출한 1 km 분배 지도를 위성 FNR과 교차 비교하고 수도권 도시에서 강원 산림으로 이어지는 동서 횡단

면에서 분배-민감도 공간 구조가 독립 자료와 방향적으로 일관되는지 확인한다. 마지막으로, 2020~2023년 4개년 추론을 통해 수도권-강원권 산화제 분배 구조의 다년도 경향을 분석한다.

2. 연구 방법

2.1 연구 지역 및 자료

2.1.1 지상 관측 자료

딥러닝 학습에는 대기오염측정망(AirKorea)의 6종 대기오염물질(PM_{10} , $\text{PM}_{2.5}$, NO_2 , O_3 , CO , SO_2)과 국립환경과학원(National Institute of Environmental Research, NIER)으로부터 제공받은 NO 등 총 7종 물질의 시간별 지상 관측 자료를 활용하였다.

분석 대상 도메인은 수도권-강원권을 포함하는 약 300 km × 180 km 영역(126.0~129.4°E, 36.5~38.2°N, UTM-K 좌표계(EPSG:5179) 기준 54,000개 1 km 격자)으로 도메인 내 309개소(도시대기 247, 도로변 38, 교외 11, 항만 13개소)의 시간별 농도를 활용하였다.

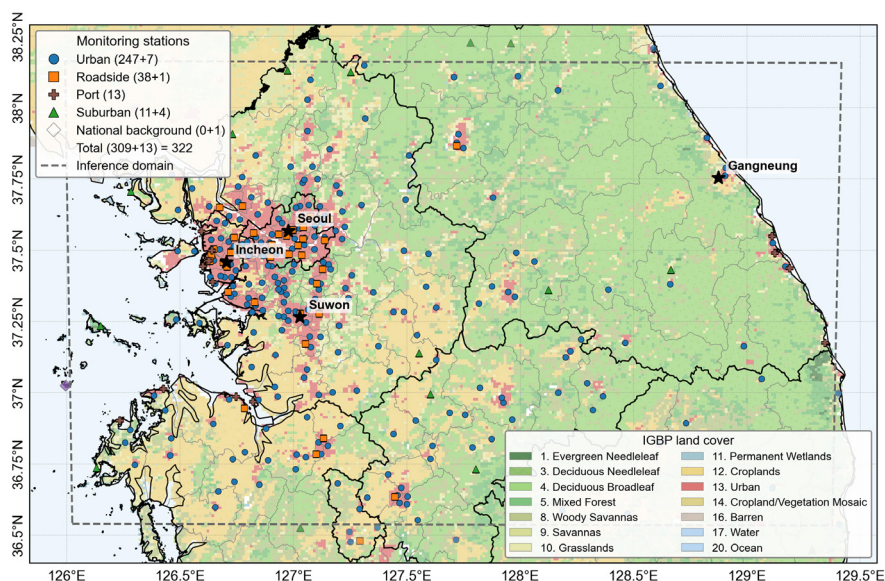


Fig. 1. Study area, monitoring stations by type, and MODIS IGBP 1 km land cover classification. Numbers in parentheses denote domain + boundary auxiliary stations within 10 km outside the domain boundary.

이에 더하여 도메인 경계 외부 10 km 이내의 13개소를 입력 자료에 추가하여(총 322개소), 경계 격자에서 관측 정보의 급격한 단절을 완화하였다. 분석 대상 도메인의 1 km 해상도 토지피복과 측정소 분포는 그림 1에 제시하였으며 관측망이 도시 지역에 집중되고 비도시 지역에는 희박하게 분포하여 뚜렷한 공간적 불균형을 보인다.

2.1.2 위성 자료

위성 자료는 기후에너지환경부 정지궤도 환경위성(Geostationary Environment Monitoring Spectrometer, GEMS) (Kim *et al.*, 2020)의 Level-2 NO₂ 및 HCHO 대류권 수직컬럼밀도(Vertical Column Density, VCD) 산출물을 사용하였다. GEMS는 천리안위성 2B호(GEO-KOMPSAT-2B, GK-2B)에 탑재된 분광계로 아시아 지역을 약 1시간 간격으로 관측하여 일 8~9회 스캔을 수행하며 공간 해상도는 약 3.5 × 8 km이다(Bae *et al.*, 2025; Lange *et al.*, 2024). NO₂ VCD는 산출물에서 직접 제공되는 대류권 컬럼을 사용하였으며 GEMS HCHO VCD의 산출 품질은 독립 위성 및 지상 분광 관측과의 검증을 통해 동북아 지역에서 확인된 바 있다(Lee *et al.*, 2024a). 데이터 품질 스크리닝은 Lange *et al.* (2024)의 GEMS 대류권 NO₂ 검증과 동일하게 QF=0 (good only) 필터를 적용하였으며 이 필터는 구름 분율(cloud fraction) ≥ 0.3인 구름 영향 픽셀과 태양천정각(SZA) ≥ 70°의 고천정각 픽셀 등 불량 관측을 자동 배제한다.

독립 교차 검증을 위해 Sentinel-5P 위성에 탑재된 대류권 관측 분광기(TROPOspheric Monitoring Instrument, TROPOMI) (Veefkind *et al.*, 2012)의 Level-2 NO₂ 및 HCHO 대류권 VCD를 사용하였다. TROPOMI는 극궤도 위성으로 하루 1회 오후(약 13:30 KST)에 한반도 상공을 통과하며 공간 해상도는 약 3.5 × 5.5 km이다. 품질 스크리닝은 각 제품의 공식 권고 기준에 따라 NO₂에 대해 qa_value ≥ 0.75 (van Geffen *et al.*, 2022), HCHO에 대해 qa_value ≥ 0.5 (De Smedt *et al.*, 2018)를 적용하였다. 본 연구에서는 관측 궤도, 시간대, 산출 알고리즘이 서로 다른 두 위성을 독립 교차

검증 자료로 활용하여 딥러닝 추론 농도장의 공간 패턴을 검증하였다.

두 위성의 FNR은 식 (1)로 산출한 뒤 GEMS·TROPOMI 비교 선행 연구(Lee *et al.*, 2024a)와 동일한 0.1° 격자로 평균하였다.

$$\text{FNR}(p, t) = \frac{\text{VCD}_{\text{HCHO}}(p, t)}{\text{VCD}_{\text{NO}_2}(p, t)} \quad (1)$$

2.1.3 배출량 자료

인위적 배출량은 국가미세먼지정보센터 대기정책 지원시스템(Clean Air Policy Support System, CAPSS)의 1 km 격자 배출 인벤토리(2021년 기준)를 사용하였다. 2022~2023년에 대해서는 동 센터의 시군구별 배출량 통계(NAEIRC, 2025)에서 2021년 대비 변화율을 산출하여 격자 배출량에 적용하였다. 배출원은 이동, 면, 산업, 에너지, 선박, 유기용제의 6개 부문으로 분류하여 NO_x, VOC, PM_{2.5}, PM₁₀, SO_x, CO, NH₃의 7종을 격자별로 산출하였으며 연간 배출량의 시간별 배분은 Ryu *et al.* (2021)의 3단계 가중치 체계(월별 × 요일별 × 시간별)를 적용하였다. 생물기원 VOC는 CAMS-GLOB-BIO v3.1 (Granier *et al.*, 2019) 월별 기후학적 배출량(0.25°)에서 이소프렌, 포름알데히드, 모노테르펜, 세스퀴테르펜 4종을, 산불 배출량은 CAMS GFAS (Kaiser *et al.*, 2012) 일별 자료(0.1°)에서 CO 화재 플럭스 1종을 각각 5 km 격자로 보간하여 사용하였다. 다만 CAPSS는 국내 행정구역 기반 인벤토리이므로 도메인 내 북서측 북한 지역에는 인위적 배출량이 반영되지 않았다.

2.1.4 입력 자료 종합 및 검증 체계

본 연구의 1 km 격자 농도장 추론에 사용된 입력 자료는 표 1에 정리하였으며 수치형 53개와 범주형 1개 총 54개 입력 변수를 포함한다. 산화제 분배 분석의 대상인 NO₂, O₃, NO 3종에 대한 검증 체계는 표 2에 제시하였다.

표 1의 입력 자료는 해상도에 따라 두 계층으로 구분된다. 기상(ERA5-Land, ERA5), CTM 배경농도(GEOS-CF, CAMS FC), 광역 배출 인벤토리(CAMS GFAS,

Table 1. Input data sources, variables, and spatial resolutions.

Category	Source	Variables (n)	Native res.	Model res.
Ground obs.	AirKorea (NIER)	PM _{2.5} , PM ₁₀ , NO ₂ , O ₃ , CO, SO ₂ , NO (7)	322 stations (309 domain + 13 buffer)	Point
Meteorology	ERA5-Land/ERA5 (ECMWF) (Hersbach et al., 2020)	T _{2m} , BLH, u/v wind, RH, precip., SSRD, SSHF, inv. strength, u/v ₈₅₀ , wind shear (12)	0.1°/0.25°	5 km bilinear
CTM (prior)	GEOS-CF (NASA) (Keller et al., 2021)	PM _{2.5} , NO ₂ , O ₃ , CO, SO ₂ , NO, HCHO (7)	0.25°	5 km bilinear
CTM (ensemble)	CAMS FC (ECMWF) (Inness et al., 2019)	NO ₂ , O ₃ , CO, SO ₂ , NO, HCHO (6)	0.4°	5 km bilinear
Fire emission	CAMS GFAS (ECMWF)	CO fire flux (1)	0.1°	5 km NN
BVOC emission	CAMS-GLOB-BIO v3.1 (Granier et al., 2019)	Isoprene, formaldehyde, monoterpenes, sesquiterpenes (4)	0.25°	5 km bilinear
Anth. emission	CAPSS	6 sectors × 7 species (21: static 14 + temporal 7)	1 km	1 km
Topography	NASADEM (NASA) (NASA JPL, 2020)	Elevation, slope (2)	~30 m	1 km
Land cover	MODIS IGBP MCD12Q1 (Friedl and Sulla-Menashe, 2022)	21 classes (embedding)	500 m	1 km

Table 2. Validation hierarchy and interpretation scope for the 2023 target year.

Level	Source /Method	Diagnostic indicator	Interpretation scope	Res.
(a) Direct fit	Temporal holdout (2023) + 5-fold spatial holdout	R ² , RMSE: NO ₂ , O ₃ , NO (309 domain stn, 2023)	Temporal generalization + spatial generalization at masked locations	Point
(b) Chemical consistency	Ox conservation	CV ratio (Ox/O ₃ by land cover)	Photostationary state reproduction	1 km
(b)	NOx-Ox regression	Slope, intercept by land cover	Oxidant production/ titration structure	1 km
(b)	Weekend ozone effect	ΔO ₃ , ΔNO ₂	Natural-experiment support for high-NOx partitioning	1 km
(b)	Urban-forest temporal structure	Diurnal O ₃ phase	Time-structure consistency across environments	1 km
(c) Cross-source external constraint	GEMS L2 FNR	Spearman ρ, directional consistency	Column-surface proxy alignment (not absolute match)	~3.5 × 8 km
(c)	TROPOMI L2 FNR	Spearman ρ, directional consistency	Independent polar-orbit cross-validation	~3.5 × 5.5 km

CAMS-GLOB-BIO) 등 저해상도 (0.1~0.4°) 격자 자료는 5 km 격자로 보간되어 광역 배경장을 제공하며 CAPSS 1 km 배출, NASADEM (~30 m) 지형, MODIS

(500 m) 토지피복 등 고해상도 자료와 측정소 관측은 1 km 격자에서 활용되어 국지 배출원과 지표 특성을 반영한다. 표 2의 검증 체계는 (a) 시간 독립 및 5-fold

공간 교차 검증을 통한 측정소 통계 일치도 평가, (b) O_x 보존, NO_x-O_x 회귀, 주말효과, 도시-산림 일변화를 이용한 화학 정합성 점검, (c) GEMS 및 TROPOMI 위성 FNR과의 독립 자료 교차 비교의 3단계로 구성된다. 본 연구의 모든 입력 자료는 별도의 자료 생산 과정(토지이용회귀 모델링, 고해상도 자료동화 농도장, 국지 확산 모델 산출물 등) 없이 기 배포된 재분석·CTM·배출 인벤토리·지형·토지피복 자료로 구성되며 이러한 외부 배포 자료와 측정망 관측 자료만으로 산출한 1 km 농도장이 화학적 정합성을 달성할 수 있는지를 실험하는 것이 본 연구의 목표 중 하나이다.

2.2 딥러닝 모델 설계

2.2.1 모델 아키텍처

대기오염물질의 농도 분포는 수백 km 규모의 종관 수송부터 1 km 이하의 국지 배출원 영향까지 다중 스케일이 중첩되어 형성된다. 본 연구는 이러한 특성을 반영하여 적응형 다중 해상도 이중 그래프 신경망(Adaptive Multi-Resolution Heterogeneous Graph Network, AMR-HGN; 이하 본 모델의 추론 산출물을 DL로 약칭)을 설계하였으며 다해상도 이중 그래프 구조, CTM 배경장 기반 적응형 편향 보정 예측, 관측 공백 지역의 신뢰성 확보 전략을 핵심으로 한다(그림 2). 모

든 입력 변수는 공유 인코더를 통해 잠재 공간으로 사상된 뒤 관측-농도 관계의 비선형 표현이 입력에서 농도 산출까지 단일 모델 내에서 일괄학습된다.

본 연구 도메인의 지상 농도 관측은 322개 측정소 위치(도메인 내 309개 + 경계 완충 측정소 13개)에만 존재하므로 1 km 격자 54,000개의 약 0.6%에 한정되는 공간적 희소성을 보인다. 정답 위치인 측정소(점 자료), 추론 대상인 1 km 격자(면 자료), 광역 종관 영향권의 5 km 격자(면 자료)는 공간 단위와 해상도가 서로 다른 이중 비균질 자료원으로 공존한다. 모든 격자에 정답이 조밀하게 존재한다고 가정하는 정규 격자 기반 CNN(Bronstein *et al.*, 2017)과 균질한 패치 토권을 가정하는 Vision Transformer (ViT) (Dosovitskiy *et al.*, 2021)는 이러한 희소·비균질 환경에 직접 적용하기 어렵다. 이에 본 연구는 정답 신호인 측정망의 비균질·희소 분포와 입력 자료의 이중·다해상도 특성을 단일 체계에서 동시에 다룰 수 있는 GNN을 채택하여 AMR-HGN을 설계하였다.

AMR-HGN은 기존 GNN 기반 대기질 모델과 다음 세 가지 점에서 구별된다. 첫째, 메시지 패싱 연산자로는 GATv2 (Brody *et al.*, 2022) 류의 동적 attention 아이디어를 참고하되 송신·수신(source·target) 노드 표현과 엣지 속성을 함께 입력으로 활용하는 엣지 속성 인

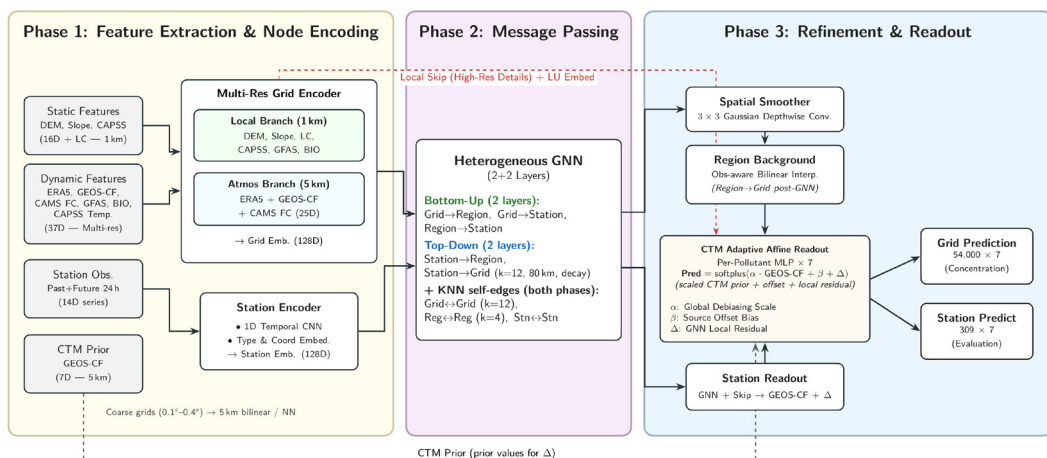


Fig. 2. Architecture of the Adaptive Multi-Resolution Heterogeneous Graph Network (AMR-HGN): Phase 1 (feature extraction and node encoding), Phase 2 (hierarchical message passing), and Phase 3 (refinement and readout).

지형(edge-aware) 이중 attention을 격자(1 km)·광역(5 km)·측정소 노드 사이의 다양한 엣지 유형으로 구성된 다해상도 이중 그래프(Terroso-Saenz *et al.*, 2024) 위에 적용한다는 점에서, 균질한 단일 해상도 그래프를 가정하는 표준 GCN(Kipf and Welling, 2017)·GAT(Veličković *et al.*, 2018) 기반 대기질 모델과 구조적으로 다르다. 둘째, 322개 측정소의 희소 정답으로부터 54,000개 1 km 격자의 조밀 농도장을 추론하는 문제 설정은 반지도 계층 GNN(Han *et al.*, 2023) 등 선행 연구를 본 단일 도메인으로 확장 적용한 것이다. 셋째, CTM 배경장을 외부 사전정보로 주입하되 학습 가능한 편향 보정 $\alpha \cdot \beta$ 와 GNN 잔차 δ (식 (3))가 결합 학습되도록 설계한 점에서 GNN 단독 구조와 구별된다.

모델은 격자(Grid, 54,000개, 1 km), 광역(Region, 2,160개, 5 km), 측정소(Station, 322개)의 세 종류 노드와 다중 엣지로 구성된 이중 그래프 위에서 작동한다. 격자 노드는 1 km 국지 정보(지형, 토지피복, 배출량)를, 광역 노드는 5 km 중규모 공간 정보를, 측정소 노드는 시간별 관측 시계열을 표현하며 계층적 메시지 패싱을 통해 스케일 간 정보가 교환된다.

Phase 1(그림 2, 특징 추출 및 노드 인코딩)에서 격자 노드의 입력 변수(표 1)는 두 경로(branch)로 분리 인코딩된다. 국지 경로(Local Branch)는 1 km 지형·배출 변수를, 대기 경로(Atmos Branch)는 5 km로 보간된 기상·CTM 변수를 각각 처리하여 국지 배출과 광역 기상의 다중 스케일 기여를 분리한다. 두 경로는 서로 다른 입력 부분집합과 별도 인코더 가중치를 사용하여 1 km 배출·지형 신호가 5 km 광역 평활화에 흡수되는 주파수 혼합을 인코더 수준에서 차단한다. 국지 경로 출력은 Skip 연결을 통해 Phase 3의 예측 단계 직접 전달되어 GNN의 공간 평활화 과정에서 소실되기 쉬운 1 km 배출 분해능의 고해상도 농도 구매를 보존한다. 측정소 노드는 과거·미래 각 24시간의 양방향 관측 시계열을 1D 합성곱으로 요약하여 현재 시점 농도 상태 표현을 강화하고 측정소 유형 임베딩(5종)과 좌표를 결합하여 128차원 표현을 생성한다.

Phase 2(그림 2, 메시지 패싱)에서는 이중 GNN이

상향 2층(격자 → 광역, 격자 → 측정소, 광역 → 측정소)과 하향 2층(측정소 → 격자, 측정소 → 광역)으로 구성된 총 4층의 계층적 메시지 패싱을 수행하며 각 층에는 동일 노드 유형 내 KNN 인접 엣지가 함께 포함된다. 광역 → 격자 방향의 정보 전달은 GNN 백본에서 분리하고 Phase 3의 광역 배경 보정 모듈에서 별도 처리하여 광역 평균장이 1 km 격자의 국지 잔차 학습을 과도하게 평활화하지 않도록 하였다. 엣지 유형별 독립 가중치와 4-head attention을 적용하여 스케일 간 정보 전달과 동일 유형 내 공간 문맥 집계를 분리 학습하며(Hu *et al.*, 2020; Wang *et al.*, 2019), 이는 이중 그래프에서 관계 유형별 attention을 분리 학습하는 표준 접근에 해당한다. 이러한 분리는 격자 ↔ 격자 메시지(농도 평활화 역할)와 측정소 → 격자 메시지(관측 정보 전파 역할)가 서로 다른 변환과 attention을 학습하도록 보장한다. 4층 메시지 패싱을 통해 다중 hop 범위의 노드 간 정보 교환이 이루어지며(Hamilton *et al.*, 2017), 측정소 관측 정보는 격자별 최근접 최대 12개 측정소로 구성된 측정소 → 격자 직접 엣지를 통해 1 km 격자에 전파된다. 공유 가중치 인코더가 전 도메인의 모든 격자에 동일한 변환을 적용하므로 측정소 밀집 지역에서 학습된 변수-농도 비선형 관계가 관측 공백 지역에서도 전이될 수 있다.

$$w(d) = \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (2)$$

여기서 d 는 측정소-격자 거리(km), σ 는 학습 가능한 감쇠 폭이다. 측정소 기원 메시지는 거리 감쇠로 영향권이 조절된 뒤 엣지 유형별 4-head attention으로 거리·기상·국지 변동의 신호 채널이 병렬 집계되어 공간 사전 분포 위에서 attention 가중치가 학습된다.

Phase 3(그림 2, 정제 및 예측)에서 최종 농도 예측은 CTM 배경장의 적응형 편향 보정과 국지 잔차 학습으로 산출된다.

$$y = \text{softplus}(\alpha \cdot \text{CTM} + \beta + \delta) \quad (3)$$

여기서 α 와 β 는 물질별 스칼라로서 CTM(GEOS-CF)의 체계적 편향을 전역적으로 보정하는 항이고 δ 는

GNN이 격자별·시간별로 산출하는 국지 잔차항이다. 전지구 CTM은 한반도에서 유의한 체계적 편향이 보고되어 있으므로 (Keller *et al.*, 2021; Ryu and Min, 2021) 저해상도 CTM을 광역 배경장으로 활용하되 관측·배출 정보가 체계적 편향을 보정하여 1 km 농도장을 산출하는 구조로 설계하였다. 이 구조는 CTM이 제공하는 대규모 공간 패턴을 물리적 사전 정보로 활용하면서 모델이 1 km 스케일의 국지 보정에 집중하도록 역할을 분리한다. 여기서 ‘적응형’은 $\alpha \cdot \beta$ 가 측정소 관측으로 자동 조정됨을, ‘다중 해상도’는 5 km CTM 배경과 1 km δ 잔차가 단일 손실로 결합 학습됨을 의미하며 단순 저해상도-고해상도 보간과 구별된다 (Reichstein *et al.*, 2019). 다만 전지구 CTM과 1 km 농도장 간 큰 해상도 격차 및 GEOS-CF 자체 편향으로 $\alpha \cdot \beta$ 의 식별 가능성에는 제약이 따를 수 있다.

총 손실 함수는 다음과 같다.

$$\mathcal{L} = \mathcal{L}_{\text{station}} + \lambda_{\text{grid}} \cdot \mathcal{L}_{\text{grid}} + w(e) \cdot \lambda_{\text{emis}} \cdot \mathcal{L}_{\text{emis}} \quad (4)$$

세 항 모두 평균 제곱 오차 (mean squared error, MSE) 기반이다. $\mathcal{L}_{\text{station}}$ 은 322개 측정소에서 모델 예측과 실측 농도 간의 유효 관측 MSE로, 주 지도 신호이다.

$$\mathcal{L}_{\text{station}} = \frac{1}{\sum m_{i,p}} \sum_{i,p} m_{i,p} \cdot (y_{i,p} - \hat{y}_{i,p})^2 \quad (5)$$

여기서 i 는 측정소, p 는 물질, $m_{i,p}$ 는 유효 관측 여부 지시자(유효=1, 결측=0)이다. $\mathcal{L}_{\text{grid}}$ ($\lambda_{\text{grid}}=0.5$)는 각 측정소에 대응하는 1 km 격자에서의 MSE로, 이중 그래프에서 측정소 노드와 격자 노드 간 예측 정합성을 확보한다. $\mathcal{L}_{\text{emis}}$ ($\lambda_{\text{emis}}=1.0$)는 격자 노드의 은닉 표현으로부터 CAPSS 배출량을 예측하는 보조 MSE 손실로, 배출 강도와 농도의 양적 관계를 은닉 표현에 내재화하여 관측 공백 지역의 추론 신뢰성을 높인다. 총 파라미터 수는 약 167만 개이며 주요 학습 하이퍼파라미터는 보충자료 표 S1에 수록하였다. 본 모델은 2022년 (8,760시간)을 학습하고 검증은 시간 독립 검증(2023년)과 5-fold 공간 교차 검증(도메인 내 309개소 중 각 fold에서 약 62개소 차단)의 이중 검증 방식으로 구성

된다. 경계 완충 측정소 13개소는 평가 대상이 아닌 GNN 입력 보조용으로 모든 검증 조건에서 활성 상태를 유지하였다.

2.2.2 관측 공백 지역의 추론 신뢰성 확보 방법

측정소가 희소한 지역에서도 물리적으로 일관된 농도장을 추론하기 위해 예측 구조와 학습 전략 양면에서 보완 장치를 두었다. 식(3)의 편향 보정 구조는 관측 정보가 부족한 원격 격자에서 국지 잔차 δ 의 학습 신호가 자연스럽게 약화되어 CTM 배경장에 가까운 예측이 산출되도록 설계되었다. 본 연구는 측정망 관측과 CAPSS 배출 인벤토리에 대한 과적합을 완화하고 측정망 희소 지역으로의 일반화 성능을 확보하기 위해 측정소 드롭아웃과 배출량 드롭아웃을 도입하였다. 두 기법은 학습 시 일부 배치에서 측정소 입력 또는 CAPSS 배출 입력에 대해 배치 단위 마스킹을 적용하여 모델이 특정 입력에 과도하게 의존하는 것을 억제하며 적용 비율은 보충자료 표 S1에 정리하였다. 한편 배출량 보조 손실 ($\mathcal{L}_{\text{emis}}$)은 은닉 표현에 배출 강도와 농도의 양적 관계를 내재화하여 측정소가 부재한 격자에서도 배출 정보가 농도 추론에 기여할 수 있도록 하였다.

2.3 화학 정합성 검증 체계

2.3.1 산화제(Ox) 보존성

본 연구는 DL 추론 1 km 농도장이 측정소 밖에서도 대기화학적으로 일관된 공간 구조를 산출하는지를 보조적으로 점검하는 데 초점을 둔다. NO₂와 O₃의 합으로 정의되는 총 산화제 Ox (식 (6))는 광화학 정상상태에서 적정-광분해 순환의 영향을 덜 받아 개별 성분보다 공간 변동이 억제되는 합성 지표이다 (Jenkin, 2004; Clapp and Jenkin, 2001).

$$\text{Ox}(x, t) = \text{NO}_2(x, t) + \text{O}_3(x, t) \quad (6)$$

Ox 보존성은 토지피복 유형 내 Ox와 O₃의 공간 변동계수(coefficient of variation, CV)를 비교하여 점검하였다.

$$CV(c, k) = \frac{\sigma_{x \in k}[\bar{c}(x)]}{\mu_{x \in k}[\bar{c}(x)]} \quad (7)$$

여기서 c 는 Ox 또는 O_3 , k 는 토지피복 유형, $\bar{c}(x) = (1/T) \sum_t c(x, t)$ 는 격자 x 의 시간 평균 농도, σ 와 μ 는 유형 k 에 속하는 1 km 격자 집합에 대한 공간 표준편차와 공간 평균이다. $CV(Ox)/CV(O_3) < 1$ 이면 적정·광분해에 의한 합성 지표로서 Ox 보존이 강함을 시사한다.

2.3.2 NOx-Ox 선형 회귀를 통한 광역·국지 오존 기여 분해

Clapp and Jenkin (2001)이 제안하고 García-Yee *et al.* (2018)이 도시·교외 공기와 분류에 적용한 Ox - NOx 선형 회귀를 통해 O_3 의 기원을 광역 수송 성분과 국지 광화학 성분으로 분해하였다.

$$Ox = a \cdot NOx + b \quad (8)$$

Ox 절편 b 는 장거리 수송 및 성층권 기여를 포함한 광역 배경 Ox 의 대리값을 나타낸다. 기울기 a 는 국지 NOx 변화에 대한 결보기 Ox 반응을 반영하며 양의 기울기는 NOx 증가에 따른 Ox 순증가를, 음의 기울기는 NOx 적정에 의한 Ox 감소 우세를 의미한다. 회귀 분석은 토지피복 유형(도시, 산림, 농경지)별로 그룹화하여 측정소가 희소한 지역을 포함한 도메인 전역에서 수행하였으며 동일 유형에 속하는 측정소 관측에 대해서도 동일 분석을 병렬 수행하여 DL 격자와 비교하였다.

2.3.3 산화제 분배비(r)와 민감도 체제 지표(FNR)

Ox 내 NO_2 의 비율인 분배비 r (식 (9))은 NOx 적정의 정도를 연속적으로 나타내며 r 이 높을수록 적정이 강한 환경을, 낮을수록 O_3 우세 배경을 나타낸다.

$$r(x, t) = \frac{NO_2(x, t)}{Ox(x, t)} \quad (9)$$

다만 r 은 산화제의 현재 분배 상태를 보여줄 뿐 배출 변동에 대한 O_3 생성 반응을 직접 측정하지 않으며 고 NOx 조건이 일반적으로 VOC 민감과 상관하더라도 혼합고·이류·광분해를 변동에 따라 r 과 민감도 방향이 분리될 수 있다. 따라서 본 연구에서 r 은 배출 민

감도의 판정 지표가 아니라 NO_2 - O_3 적정-분배 구조의 진단 도구로 한정하여 활용한다. r 기반 분류에는 문헌에서 확립된 임계값이 존재하지 않으므로 운영적 기준을 도입하였다. 상한 $r = 0.5$ 는 $NO_2 = O_3$ 의 분배 등점이며 하한 $r = 0.2$ 는 Ox 의 80%가 O_3 로 구성되는 수준으로 이 이하에서는 NOx 적정의 기여가 상대적으로 작아 광역 배경 O_3 가 우세한 분배 상태에 해당한다. 이에 따라 $r \geq 0.5$ (고적정, NO_2 우세), $0.2 \leq r < 0.5$ (전이), $r < 0.2$ (저적정, O_3 우세)로 구분하였다. 이 3분류는 배출 민감도 체제 분류가 아니라 분배 상태의 공간 유형화로 활용된다.

FNR (=HCHO VCD/ NO_2 VCD)은 오존 민감도 대리지표로 확립되어 있으나(Duncan *et al.*, 2010; Martin *et al.*, 2004) 임계값은 지역별 배출 구조에 따라 상당히 달라질 수 있다. Tian *et al.* (2025)은 1996~2022년 전지구 장기자료를 통해 FNR 전이 임계값을 전지구 평균 약 3.1, 동아시아 약 2.8로 제시하여 임계값의 지역적 불확실성을 재확인한 바 있다. 이에 더하여 위성 산출 오차(Souri *et al.*, 2023)와 컬럼-지표 대표성 불일치가 내재하므로 본 연구에서는 FNR을 절대 경계 확정보다는 공간적 방향성 일관성 수준에서 r 과 교차 검증하는 방식으로 활용한다. FNR 3분류는 Duncan *et al.* (2010)의 VOC 민감 하한($FNR < 1$)과 Jin *et al.* (2020)의 전이 상한($FNR \approx 3-4$)을 종합하여 $FNR < 1$ (VOC 민감 방향), $1 \sim 3.5$ (전이), > 3.5 (NOx 민감 방향)로 설정하였다.

2.4 분석 시간창과 화학 정합성 점검 요소

본 모델의 학습은 2022년 1개 연도 전 기간(8,760시간)의 자료로 수행하여 일변화·요일·계절 변동을 포함한 연중 시간 구조를 학습 분포에 반영하였다. 통계적 성능 검증(측정소 위치 R^2 ·RMSE 및 시공간 독립 검증)은 추론 대상 연도인 2023년 전체 기간을 대상으로 수행하였다. 한편 본 연구의 분석 대상이 지표 산화제 분배 구조의 공간 진단이므로 본 절에서 다루는 화학적 정합성 검증(Ox 보존성, NOx - Ox 회귀, 위성 FNR 비교)은 광화학 정상상태 가정이 타당하고 위성 FNR과의 교차 비교가 가능하며 도시·배경 분배 구배

의 대비가 가장 명료한 시간창으로 선행 연구(Lee *et al.*, 2024b; Baek *et al.*, 2023; Jeong *et al.*, 2022; Yeo and Kim, 2021)를 참고하여 광화학 활성기의 주간을 선정하여 수행하였다.

Ox 보존, r 기반 분배 분류와 FNR 교차 검증은 2023년 광화학 활성기(5~9월) 주간(10~16 KST)을 대상으로 하였다. 지표 산화제 분포와 위성 FNR 민감도 체제 비교를 위해 GEMS 관측이 가능한 주간 시간대 내로 한정하였다. NO_x-Ox 회귀와 주말효과 분석은 광화학 신호가 극대인 JJA(6~8월, 여름철) 주간(10~16 KST)에 한정하였다. NO_x-Ox 회귀는 DL 1 km 격자와 측정소 관측을 비교하여 DL 농도장의 NO_x-Ox 반응 구조가 관측과 부합하는지 점검하였다. 주말효과 분석은 주중-주말 간 NO_x 배출 차이에 따른 O₃ 농도 응답을 비교하는 실험으로서 도심 고NO_x 적정 영역에서 주말 O₃ 상승이 나타나는지를 통해 DL 농도장이 배출-적정 관계를 반영하는지 점검하였다. 도시-산림 O₃ 일변화 진폭 비교는 2023년 전체 기간을 대상으로 도시의 강한 적정-광분해 일주기와 산림 배경의 완만한 일주기 간 시간 구조 차이를 통해 DL 농도장의 환경 유형별 시간 패턴 재현성을 점검하였다. 마지막으로 2020~2023년의 산화제 분배 구조의 다년도 변동을 분석하였다.

3. 결과 및 고찰

3.1 딥러닝 모델의 통계적 성능 평가

표 3에 이중 검증 결과를 제시하였다. 시간 독립 검증은 학습에 사용되지 않은 2023년 전체에 대해 도메인 내 309개 측정소 위치에서 예측-관측 일치도를 평

가한 것이며 5-fold 공간 교차 검증은 동일 학습 모델로 309개 도메인 측정소 중 각 fold에서 62개소(20%)의 관측 입력을 추론 시점에 차단하고 해당 측정소 위치의 예측-관측 일치도를 fold별로 평가한 후 5-fold 평균을 산출한 것이다. 학습 단계에서는 309개 측정소 전부를 사용하므로 본 검증은 추론 시점의 관측 의존성을 차단한 공간 외삽 성능 평가에 해당한다. 전체 격자 추론 시에는 가용한 모든 관측 정보를 GNN 입력으로 활용하였다.

시간 독립 검증 결과 NO₂ ($R^2 = 0.892$)와 O₃ ($R^2 = 0.926$)는 학습에 사용되지 않은 2023년에서도 높은 재현력을 보였다. NO는 Ox 산출 및 NO_x-Ox 회귀 분석(3.2.2절)의 보조 물질로 함께 제시하였다. 5-fold 공간 교차 검증에서는 NO₂ $R^2 = 0.539$, O₃ $R^2 = 0.732$, NO $R^2 = 0.590$ 로 나타났다. O₃의 공간 교차 검증 성능이 가장 높은 것은 광화학 생성과 기상 수송이 농도장을 공간적으로 평활화하여 측정소 정보 없이도 외삽이 용이하기 때문으로 판단된다. 반면 NO₂와 NO는 도로변·산업단지 등 국지 배출원에 대한 의존도가 높고 NO의 경우 대기 수명이 매우 짧아 공간적 변동성이 크기 때문에 관측 정보가 차단된 조건에서 공간 외삽이 상대적으로 어렵다. 시간 독립 대비 공간 교차 R^2 의 하락은 GNN의 측정소 정보 의존도를 반영하며 관측 공백 지역에서의 추론 불확실성이 시간 외삽보다 크다는 것을 시사한다. 7종 전체 물질의 검증 결과는 보충자료 표 S2에 제시하였다.

3.2 지표 산화제 농도 구조의 화학적 정합성

3.2.1 Ox 공간 분포 및 보존 특성

본 절의 토지피복별 정량 통계는 CAPSS 배출 인벤토리

Table 3. Deep learning model dual validation performance for NO₂, O₃, and NO.

Pollutant	In-sample R^2 (2022)	Temporal holdout R^2 (2023)	5-fold spatial holdout R^2 (2023)	Spatial holdout RMSE
NO ₂	0.910	0.892	0.539	8.84 ppb
O ₃	0.944	0.926	0.732	10.55 ppb
NO	0.883	0.842	0.590	9.93 ppb

R^2 , RMSE from grid readout at station-collocated 1 km grid cells
Train: 2022, Eval: 2023

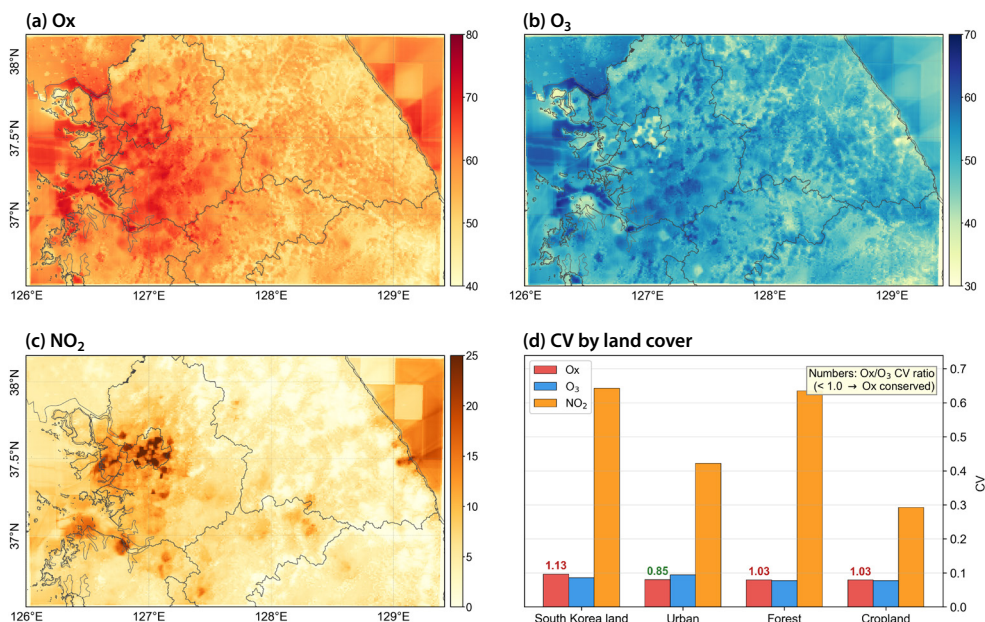


Fig. 3. Spatial distribution of (a) Ox, (b) O₃, (c) NO₂ (ppb; color scale) for May~Sep daytime (10~16 KST) 2023, and (d) coefficient of variation (CV) of Ox, O₃, and NO₂ for four South Korea land categories (South Korea land = sido boundary ∩ IGBP ≠ 20; Urban = IGBP 13; Forest = IGBP 1, 4, 5; Cropland = IGBP 12, 14). Ox/O₃ CV ratio annotated above each Ox bar (red = ratio ≥ 1, green = ratio < 1).

토리가 미적용된 북한 격자와 해양 격자를 제외하고 시도 행정경계 내 남한 육상 격자만을 대상으로 산출하였다. 그림 3d의 'South Korea land' 카테고리는 남한 육상 격자(IGBP 해양 코드 20 제외, 시도 행정경계 내 격자) 전체를 의미하며 도시·산림·농경지 카테고리는 남한 육상 내 각 토지피복 격자를 대상으로 한다. 해양 격자는 측정소가 부재하여 직접 검증과 관측 보정이 모두 제한적이며 도메인 경계부에 위치하여 격자 경계 불연속이 관찰되므로(그림 3a~c) 정량 분석에서 제외하였다.

그림 3은 2023년 광화학 활성기(5~9월) 주간(10~16 KST) 평균의 Ox, O₃, NO₂ 공간 분포를 나타낸다. 시도 행정경계 기준으로 서울 도시 격자 평균 O₃는 ~48 ppb로 가장 낮았으며 경기 도시 격자는 ~52 ppb로 가장 높게 나타났는데 그림 3b의 공간 분포에서 경기 남부 교외가 도내 최고 농도 영역을 형성하였다. 강원 산림 격자의 평균 O₃는 ~50 ppb로 경기 도시보다 소폭 낮았는데 이는 광화학 활성기에 국지 광화학 O₃

생성이 적고 광역 배경 수준에 가까운 농도를 보이는 것으로 해석된다. 300 km × 180 km 도메인에서 Ox (=NO₂ + O₃) 자체는 서울 도시(~66 ppb)에서 강원 산림(~53 ppb)으로의 체제적 경사를 보여 서로 다른 공기과 기원에 따른 광역 Ox 보존 수준의 차이를 반영하였다.

Ox 보존성은 토지피복 유형 내에서 변동계수 CV (식 (7))의 비율 CV(Ox)/CV(O₃)로 평가하였으며 CV 비율이 작을수록 해당 유형 내 Ox의 공간 변동이 O₃보다 작음을 의미한다. 육상 유형 중 도시의 Ox/O₃ CV 비율은 0.85로 가장 낮게 나타났는데, 이는 공간적으로 집중된 NO 배출이 국지 적정을 통해 O₃의 공간 변동을 확대하는 반면 O₃ + NO₂로 정의되는 Ox는 이러한 적정 과정에서 상대적으로 보존적으로 거동하기 때문으로 해석되며 2.3.1절에서 기술한 Ox의 보존적 거동이 1 km 격자 단위에서도 확인됨을 보여준다. 산림(1.03), 농경지(1.03), 남한 육상 전체(1.13)는 CV 비율이 1 부근 또는 그 이상으로 도시와 대비되었다

(그림 3d).

본 분석에서 사용한 토지피복별 CV 비율은 측정소 단위의 NO_x-O_x 회귀(Jenkin, 2004; Clapp and Jenkin, 2001)와 달리 1 km 격자 농도장의 공간 분산 구조를 평가하는 방법론적 보강이다. 산림·농경지·남한 육상 전체에서 CV 비율이 1 이상으로 산출된 결과는 본 평가 방식이 국지 적정·광분해 효과뿐 아니라 광역 배경 O_x 구매와 토지피복 영역의 공간 이질성을 함께 반영하기 때문으로 해석된다. 도메인 전역에 넓게 분포하는 유형에서는 서로 다른 공기괴 기원과 배경 농도 차이에 따른 광역 O_x 구매가 비율을 1 부근 또는 그 이상으로 끌어올릴 수 있으며 본 평가 방식만으로 적정 효과와 광역 배경 구매 효과를 정량적으로 분리하기에는 한계가 있다. 사전 공간 추세 제거 또는 하위 영역 단위 분석 등 토지피복별 O_x 보존성 평가의 방법론적 정제는 후속 과제로 남긴다.

이상의 토지피복별 CV 비율 대비는 모델이 토지피복별 O_x 보존 수준의 차이를 1 km 격자 단위에서 재현하고 있음을 시사한다.

3.2.2 NO_x-O_x 반응 특성과 배출 강도와의 관계

본 절에서는 García-Yee *et al.* (2018)의 접근을 따라 NO_x-O_x 선형 회귀의 기울기와 절편으로부터 공기괴의 신선·노화 정도와 국지 광화학 생성·광역 수송의 상대적 기여를 토지피복 유형별로 진단하였다. 그림 4는 토지피복 유형별 NO_x-O_x 선형 회귀 결과를 DL 1 km 격자(적색)와 동일 유형 내 측정소 관측(청색)으로 비교한 것이다(JJA 주간, 10~16 KST). 단, 산림 유형 비교에는 모델 학습 도메인의 경계 외부 10 km 이내에 위치한 강원 교외대기 측정소 4개소를 선정하였으며 DL 추론과 비교할 측정소가 매우 희소하여 직접 비교에 한계가 있다.

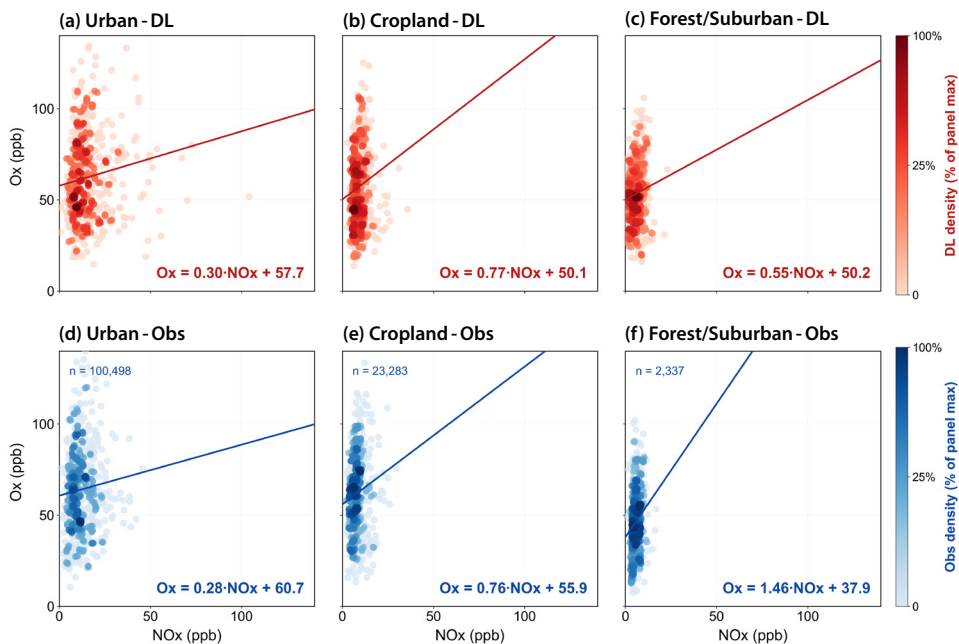


Fig. 4. NO_x-O_x scatter plots by land cover type (JJA daytime 10~16 KST, 2023). Top row (red, DL 1 km grid cells): (a) Urban - DL, (b) Cropland - DL, (c) Forest/Suburban - DL. Bottom row (blue, AirKorea station observations): (d) Urban - Obs, (e) Cropland - Obs, (f) Forest/Suburban - Obs. Relative density colorbars (low/mid/high) shown on the right of each row. Regression equations annotated per panel. Land cover masks: Urban (IGBP 13; 166 obs stations), Cropland (IGBP 12, 14; 39 stations), Forest/Suburban (DL: IGBP 1, 4, 5 forest grid; Obs: 4 Gangwon suburban stations within station buffer).

DL 격자의 토지피복 유형별 회귀 기울기는 농경지(0.77) > 산림(0.55) > 도시(0.30) 순으로 나타났다. 동일 유형 내 측정소 관측 기울기는 농경지(39개소, 0.76), 도시(166개소, 0.28), 산림(강원 교외 4개소, 1.46) 순으로 산출되었다. 도시와 농경지에서는 DL과 관측의 기울기가 각각 0.30과 0.28, 0.77과 0.76으로 높은 일치도를 보여 DL 농도장이 관측의 NO_x-O_x 반응 구조를 잘 재현하고 있음이 확인된다. 도시와 같은 고 NO_x 환경에서는 NO_x 증가에 따른 순 O_x 생성 효율이 제한되어 기울기가 낮게 나타나는 반면, 상대적으로 저NO_x인 농경지와 산림에서는 NO_x 증가가 광화학 O_x 생성으로 보다 효율적으로 연결되어 기울기가 크게 나타난다.

산림 격자(DL 0.55)와 강원 교외 측정소 4개소(1.46)의 약 2.6배 기울기 차이는 다음 세 가지 요인의 복합으로 해석된다. 첫째, 산림 토지피복은 격자 내부에 직접 대응되는 측정소가 극히 희소하여 NO_x-O_x 회귀의 관측 기반 직접 검증이 본질적으로 어렵다. 측정소 기원 메시지가 거리 감쇠(식 (2))로 약화되고 인위적 NO_x 배출도 희박하여 GNN이 학습하는 국지 간차 δ (식 (3))의 신호가 자연스럽게 작아지므로 산림 격자의 DL 출력이 CTM 배경장 보정 항($\alpha \cdot \text{CTM} + \beta$)에 가까워지는 경향이 발생한다. 둘째, 본 연구는 차선의 비교 대상으로 도메인 경계 완충 측정소를 포함하여 강원 교외대기 측정소 4개소를 선정하였다. 이 측정소들은 도시대기·도로변 측정소 대비 NO_x 배출원에서 멀고 산림 토지피복에 인접하여 생물기원 휘발성유기화합물(Biogenic VOC, BVOC)이 풍부한 산림 화학 환경을 가장 가깝게 반영할 수 있는 위치이다. 다만 표본이 4개소로 매우 적어 통계적 안정성이 낮고 정주지·도로 등 소규모 인위적 배출원의 영향을 받을 수 있어 인위적 배출이 통제된 순수 산림 화학 환경의 완전한 대리지표로 보기는 어렵다. 또한 점 자료인 측정소와 면 평균인 1km 격자 사이의 본질적 공간 대표성 차이도 기울기 차이에 영향을 미친다. 셋째, 강원 교외 관측의 높은 기울기는 BVOC가 풍부한 환경에서 NO_x 첨가가 광화학 O₃ 생성을 강화하는 잘 알려진

패턴을 반영한다. 생물기원 이소프렌·모노테르펜은 OH 라디칼과 반응하여 과산화라디칼(RO₂)을 생성하고 RO₂가 NO를 NO₂로 산화시키는 사슬을 통해 광화학 O₃ 생성에 기여하기 때문이다(Yu and Son, 2021). 그러나 본 데이터 기반 모델은 BVOC 매개 NO_x-VOC 비선형 광화학 반응을 명시적 화학 반응식이 아닌 입력 변수와 측정소 관측의 관계로만 학습한다. 따라서 학습 도메인 내 산림 측정소 표본의 절대적 희소성과 결합하여 BVOC 풍부 환경의 NO_x-O₃ 비선형 응답을 충분히 포착하지 못하고 첫째 요인의 CTM 배경장 의존 효과와 결합되어 산림 영역 NO_x-O_x 기울기가 과소 추론될 가능성이 있다. 약 2.6배의 차이는 이러한 BVOC 매개 광화학 학습 한계가 정량적으로 드러난 결과로 해석할 수 있다.

DL 격자의 광역 배경 O_x (O_x 절편)는 도시(57.7 ppb) > 산림(50.2 ppb) $\approx$ 농경지(50.1 ppb) 순으로 나타났다. 도시에서 O_x 절편이 높게 나타난 것은 서해 방면의 광역 수송 영향권에 위치하여 배경 O_x 수준이 높고(Kim, 2025), NO 적정에 의한 손실에도 불구하고 국지 광화학 생성이 추가적으로 기여했을 가능성이 있다. 반면 농경지와 산림 O_x 절편이 낮은 것은 높은 회귀 기울기에 수반되는 구조적 특성도 있지만 내륙 지역의 상대적으로 낮은 광역 O_x 유입과 광화학 활성기 주간에 산림 지역의 활발한 수관층 건성침적 등 물리적 요인과 일관되는 결과로 해석할 수 있다.

3.2.3 주말효과 및 도시-산림 지역의 시간적 구조 비교

주말-주중 O₃ 차이는 배출 활동의 주기적 변화를 활용하는 자연실험적 진단으로 도시 환경의 NO_x 적정 효과를 검토하는 데 널리 사용되어 왔다(Tian *et al.*, 2025; Sicard *et al.*, 2020). 주말에는 교통량 감소로 NO_x 배출이 주중보다 자연적으로 줄어들며, 이에 따른 O₃ 농도 변화는 인위적 통제 실험 없이도 배출 변동에 대한 O₃ 반응 방향을 간접적으로 관찰할 수 있는 자연실험적 상황을 제공한다. 그림 5는 2023년 JJA 주간의 주말-주중 O₃ 및 NO₂ 차이의 공간 분포를 나타낸다. 남한 육상 평균에서는 $\Delta \text{O}_3 = +4.93$ ppb, $\Delta \text{NO}_2 = -0.64$

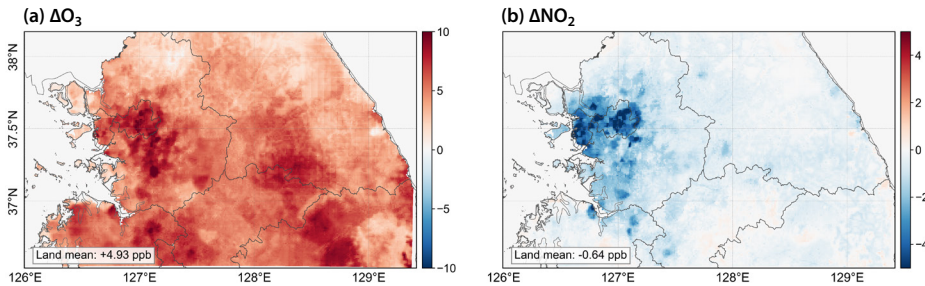


Fig. 5. Weekend minus weekday difference for JJA daytime (10~16 KST) 2023: (a) ΔO_3 and (b) ΔNO_2 (ppb; color scale). North Korea and ocean grids are masked (side administrative boundary $\cap$ IGBP $\neq$ 20), consistent with §3.2.1 and §3.2.2 analysis domain. Annotated land mean values are computed over South Korea land grids only.

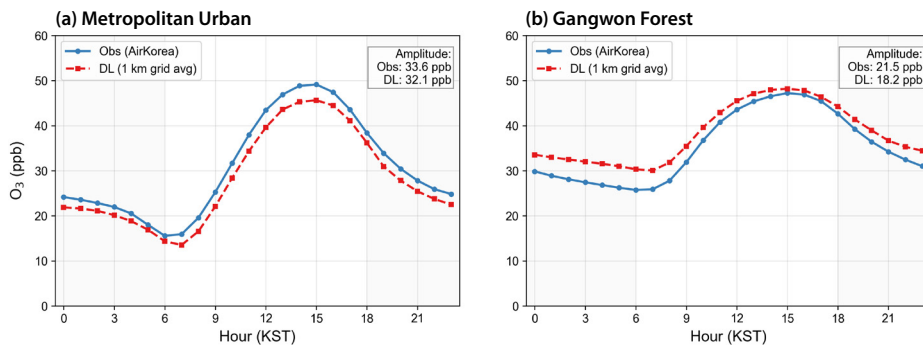


Fig. 6. DL-inferred vs observed O₃ diurnal patterns, full year 2023: (a) Metropolitan urban (DL: 2,356 grid cells, Obs: 123 urban stations in Seoul, Incheon, and Gyeonggi with grid IGBP 13) and (b) Gangwon forest (DL: 10,966 grid cells, Obs: 4 Gangwon suburban stations within station buffer)

ppb로 나타나, 주말 NO_x 감소에 따라 O₃가 전반적으로 증가하는 경향이 확인되었다. 서울 도시에서는 $\Delta O_3 = +8.0$ ppb, $\Delta NO_2 = -3.8$ ppb로 강한 주말효과를 보였다. 이는 NO_x 과잉 환경에서 주말 NO_x 감소가 O₃ 적정을 약화시켜 오히려 O₃ 농도를 상승시키는 대기화학적으로 기대되는 패턴과 일관된다(Tian *et al.*, 2025). 반면 강원 산림에서는 $\Delta O_3 = +4.2$ ppb, $\Delta NO_2 = -0.34$ ppb로 나타나 서울 도시의 약 절반 수준에 머물렀는데 이는 NO_x 배출이 희소한 배경 환경에서는 주중-주말 배출 차이 자체가 작아 주말효과가 약화되는 특성을 반영한다. 이 결과는 3.2.2절의 도시 NO_x-O_x 회귀 기울기 저하(0.30)와 동일한 비선형 구조를 상호 보완적인 관점에서 확인한 것으로 고NO_x 환경의 적정 포화가 주말 NO_x 감소 조건에서는 O₃ 상승이라는 역방향

응답으로 관찰됨을 뒷받침한다. 한편 남한 육상 격자의 100.0%에서 $\Delta O_3 > 0$, 91.0%에서 $\Delta NO_2 < 0$ 으로 나타나 주말 NO_x 감소와 그에 대응하는 O₃ 변화가 분석 영역 전역에서 관찰되었다.

이러한 DL 추론장의 주말효과 패턴은 실제 관측에서도 확인되었다. 서울 도시대기 측정소 25개소에서 $\Delta O_3 = +8.5$ ppb, $\Delta NO_2 = -3.9$ ppb로 DL 격자 분석($\Delta O_3 = +8.0$ ppb)과 양호하게 일치하였으며 전 측정소에서 $\Delta O_3 > 0$ 이었다. 강원 도메인 내 측정소 30개소에서는 $\Delta O_3 = +3.6$ ppb로 서울의 약 43% 수준에 그쳐 도시-배경 간 주말효과 대비가 관측에서도 재현되었다. 이러한 DL-관측 간 정합성은 DL 농도장이 배출 변동에 대응하는 농도 응답 구조를 측정소 외 영역에서도 공간적으로 재현하고 있음을 시사한다. 도심 ΔO_3 의

증가 방향은 수도권 지역의 오존 증가 분석(Colombi *et al.*, 2023) 및 도시 환경의 주말효과 종합 분석(Tian *et al.*, 2025; Sicard *et al.*, 2020)에서도 일관되게 보고된 양상이다. 다만 본 분석은 기상과 배출 효과를 명시적으로 분리하는 절차(Ajayi *et al.*, 2025; Koo *et al.*, 2012)를 적용하지 않았으므로 도심 ΔO_3 의 정량값에는 기상 변동에 의한 잔여 영향이 일부 남을 가능성이 있다.

수도권 도시와 강원 산림의 O_3 일변화 구조에서도 두 환경의 산화제 분배 특성이 뚜렷하게 대비되었다(그림 6). 수도권에서는 오전 출퇴근 시간대의 NO_x 적정에 의한 O_3 최저와 오후 UV 극대 시의 광화학 생성에 의한 최대가 번갈아 나타나는 전형적인 적정-광분해 일변화(Kim *et al.*, 2016)가 DL과 관측 간에 양호하게 일치하였다(관측 진폭 33.6 ppb, DL 진폭 32.1 ppb). 강원 산림에서는 NO_x 배출이 희소하고 생물기원 VOC가 우세한 환경으로(Yu and Son, 2021) 적정에 의한 오전 O_3 감소가 현저히 약하고 일변화 진폭(관측 21.5 ppb, DL 18.2 ppb)이 도시의 약 60% 수준(관측 64%, DL 57%)으로 완만한 배경 O_3 우세 패턴을 보였다.

수도권에서는 DL이 관측보다 체계적으로 낮은 농도(오후 peak 기준 약 3.5 ppb)를 추론한 반면 강원 산림에서는 DL이 관측보다 일평균 약 3.3 ppb 높게 추론하였다. 이는 수도권의 경우 도시 환경의 간헐적 고농도 peak를 학습 단계에서 충분히 포착하지 못한 딥러닝 모델의 한계가 시간 평균에 누적된 결과이며 강원 지역의 경우 교외대기 측정소 4개소와 산림 격자 10,966개

사이의 공간 대표성 차이가 작용한 것으로 판단된다. 강원 산림 DL 격자(10,966개)는 배출원에서 더 떨어진 산림 내부 격자를 다수 포함하므로 격자 평균에서는 NO_x 적정 효과가 측정소 위치 대비 약화되어 O_3 가 상대적으로 높게 추론된 것으로 판단된다.

3.3 1 km 해상도 지표 산화제 분배 구조의 외부 정합성 평가

3.3.1 1 km 지표 산화제 공간 분포

도심의 강한 NO_x 적정 효과가 배경 O_x 우세 상태로 전환되는 공간 구매는 수 km 이내의 좁은 범위에서 형성된다. 따라서 이러한 산화제 분배의 공간적 전이 대역을 분해하고 도시-교외 화학 대비를 분석하기 위해서는 수 km 미만의 고해상도 농도장이 요구된다. 그림 7은 2023년 광화학 활성기(5~9월) 주간(10~16 KST) 평균 분배비 $r (=NO_2/O_x)$ 에 기반한 1 km 지표 산화제 분배 지도를 나타낸다. 이 지도는 배출 변동에 대한 O_3 반응을 나타내는 민감도 지도가 아니라 지표에서의 NO_2 - O_3 분배 구조를 보여주는 진단 지도이다. 해양·북한을 제외한 남한 육상 41,432개 격자에서 저적정 O_3 우세 영역($r < 0.2$, 광역 배경 O_x 우세)이 96.4%, 전이 영역($0.2 \leq r < 0.5$)이 3.6%, 고적정 NO_2 우세 영역($r \geq 0.5$, 강한 NO_x 적정)이 0.0% (서울 도심 핵심부 소수 격자)를 차지하였다.

서울 도심 핵심부는 도메인 내에서 r 이 가장 높게 나타난 영역으로(격자 최댓값 ≈ 0.50 , 전이 구간 상한 부근), NO_x 적정이 가장 강하게 작동하는 고적정 NO_2

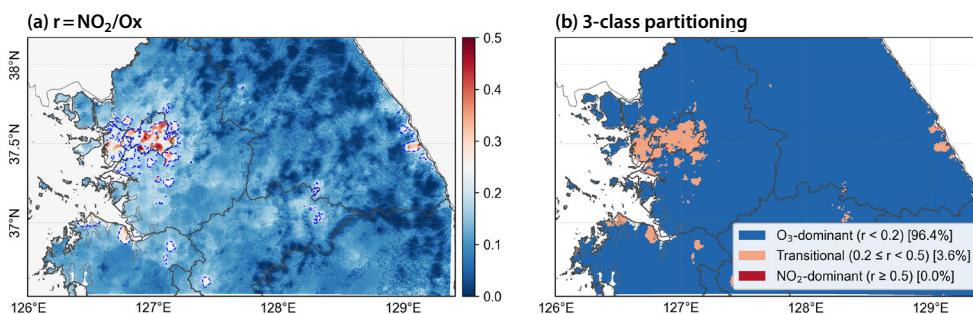


Fig. 7. 1 km surface oxidant partitioning map for May~Sep daytime (10~16 KST) 2023: (a) continuous $r (=NO_2/O_x)$; color scale) and (b) 3-class surface oxidant partitioning.

우세 영역에 해당한다. 전이 영역(1,497개 격자)은 수도권에 약 75%가 집중되었으며 인천 송도·남동공단 및 경기 시화 산업단지, 충남 당진·서산 화력·석유화학단지, 강원권·동해안(동해·삼척) 화력·산업단지 등 NO_x 배출이 집중된 도시·산업·항만 영역과 공간적으로 일치하였다. 대부분의 도메인(배경, 산림, 교외)은 저적정 O₃ 우세로 분류되었다. 이러한 분배 구조의 공간 분포는 Oak *et al.* (2025)이 보고한 한국의 대부분이 NO_x 제한 또는 전이이고 VOC 제한은 주로 중앙 수도권에 남아 있다는 최근 결과, 그리고 Lee *et al.* (2024b)이 OH 반응성 비 분석에서 인천이 전이 경계에 근접하였음을 보고한 결과와 방향적으로 일관된다. 다만 본 *r* 기반 분배 분류는 배출 민감도 체제 분류가 아니므로 선행 문헌의 VOC 제한·NO_x 제한·전이 용어를 직접 귀속시키지 않고 분배 구조와 민감도 체제 사이의 공간적 방향성 대응 수준에서 해석한다.

DL 추론 농도장은 서울 도심 내부의 *r* 공간 구배를 1 km 해상도에서 재현하였으며 고적정 NO₂ 우세와 저적정 O₃ 우세 사이의 전이 영역은 NO_x 배출이 집중된 도시·산업 지역을 따라 형성되는 경향을 보였다. 이는 DL 모델이 배출 강도와 적정-분배 관계의 공간적 연계를 효과적으로 포착하고 있음을 시사한다.

3.3.2 위성 FNR을 이용한 오존 생성 민감도 체제 비교

본 절에서는 위성 FNR(컬럼 적분 산출물)과 DL 추론 *r*(=NO₂/O_x, 지표 농도 산출물)이라는 서로 다른 물리량이 산화제 분배 구조의 공간 방향성에서 상호 일관됨을 확인하는 다중 자료 정합성 점검을 수행한다. 그림 8은 두 위성(GEMS, TROPOMI)의 FNR 공간 분포와 DL 추론 *r*의 교차 비교를 2023년 5~9월로 통일하여 제시한 것이다. GEMS는 관측 시간대(주간 일 8~9회) 중 10~16 KST 관측을 평균하여 *r* 산출 시

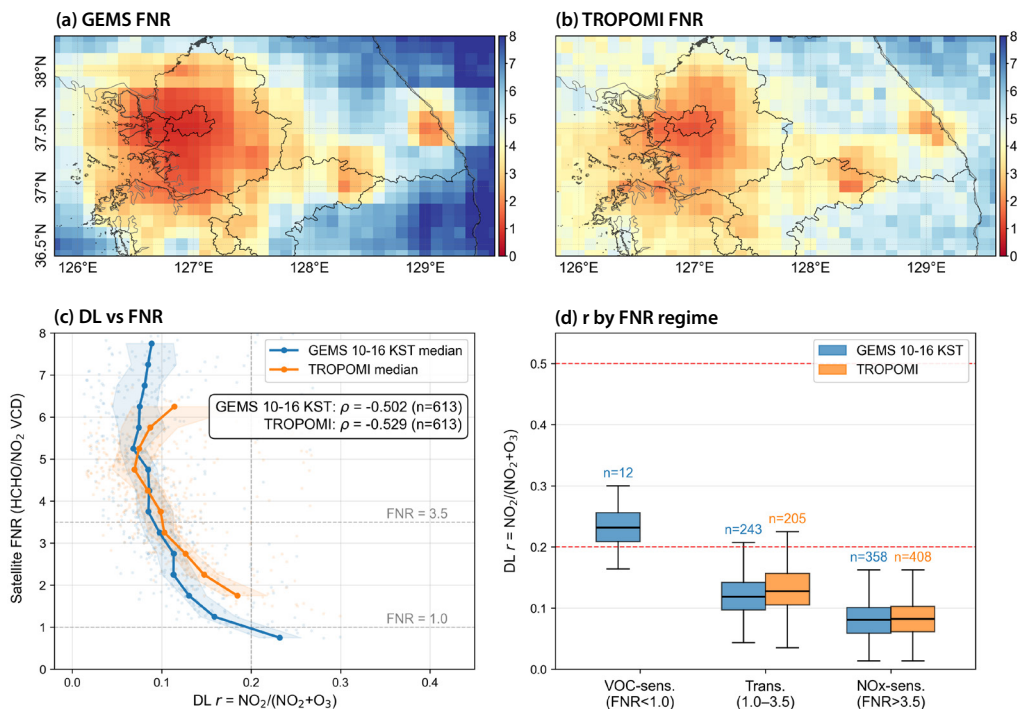


Fig. 8. Dual-satellite FNR and DL *r* (=NO₂/O_x) comparison (May~Sep 2023): (a) GEMS FNR 10~16 KST average, (b) TROPOMI FNR, (c) DL *r* vs dual-source FNR scatter with FNR-bin median *r* curves and IQR bands (n=613 jointly valid 0.1° grid cells), (d) DL *r* distribution conditioned on satellite FNR regime class using GEMS 10~16 KST and TROPOMI (thresholds: FNR < 1/1.0~3.5/> 3.5).

간장과 일치시켰으며 TROPOMI는 극궤도에서 일 1 회 오후(약 13:30 KST)에 한반도 상공을 통과하므로 별도의 시간 선별 없이 해당 시점의 관측값을 사용하였다. 그림 8a에 GEMS 10~16 KST 평균, 그림 8b에 TROPOMI FNR의 공간 분포를 나타내었다. 두 위성 모두 서울 도심에서 FNR이 최저(VOC 민감 방향)를 보이고 강원 산림으로 갈수록 증가(NO_x 민감 방향)하는 도시-배경 정도 구배를 일관되게 포착하였다.

그림 8c에 DL 추론 분배비 r ($=\text{NO}_2/\text{Ox}$)과 두 위성 FNR 산출물의 관계를 산점도와 FNR 구간별 median r 곡선으로 제시하였다. GEMS 10~16 KST 평균($\rho = -0.502$), TROPOMI($\rho = -0.529$) 모두에서 r -FNR의 단조 감소 관계가 확인되었다. 두 위성의 median r 곡선이 FNR 범위 전반에서 유사한 형태를 보이는 것은 관측 시간대와 관측 기하에 관계없이 r -FNR 단조 관계가 강건함을 확인한다.

그림 8d에는 GEMS 10~16 KST FNR 민감도 체제별(VOC 민감: FNR<1, 전이: 1~3.5, NO_x 민감: >3.5) DL r 분포를 TROPOMI와 비교하여 제시하였다. GEMS FNR이 VOC 민감도로 분류한 격자에서 DL r 의 중앙값은 0.232로 본 연구의 분배 전환 지점($r=0.2$) 부근에 위치함을 확인하였으며 이는 위성 민감도 체제와 지표 분배 구조가 공간적으로 대응함을 시사한다. TROPOMI에서는 FNR<1에 해당하는 격자가 관찰되지 않았는데 이는 극궤도 위성의 오후 1회 관측 특성상 FNR이 전반적으로 높게 산출되는 경향에 기인한 것으로 추정된다.

표 4에 지상 관측, DL 모델, 위성 FNR의 산화제 분배 분류를 교차 비교하였다. 지상 측정소 r 과 DL r 의 3분류 일치율은 90.2%로 높게 나타났다. 지표-컬럼 간(Surface-Column) 일치율은 57.7~67.0%로 낮았으나

이는 컬럼 적분(FNR)과 지표 농도(r)라는 본질적으로 다른 물리량 사이의 구조적 차이에 기인한다(Souri et al., 2023). 반면 두 위성 간(Column-Column) 일치율은 87.6% ($\rho = +0.882$)로 관측 시간대와 관측 기하가 다름에도 독립 위성 산출물 간 양호한 정합성을 보였다.

위성 FNR은 컬럼 적분 산출물로서 HCHO·NO₂ 컬럼 산출 알고리즘에 내재된 오차, 컬럼-경계층 변환의 가정 오차, 평균화 커널(averaging kernel, AK)의 비식별성에 따른 경계층 감도 저하(NO₂와 HCHO 모두 도시 핵심부 대기경계층 농도 변동이 과소 표현될 수 있으며, FNR 비율 편향 방향은 두 종의 수직 프로파일 차이와 화학적 피드백을 함께 고려하여 해석해야 함), 픽셀 해상도(수 km)에 따른 1km 미세 구배 분해 한계, 맑은 날 편향(구름 분율 필터링), 운량·시야각·태양천정각에 따른 0.1° 격자별 산출 표본 수의 시공간 비균질성 등의 한계를 갖는다. 반면 본 연구의 DL 산화제 분배 r 은 지표 농도를 직접 산출하므로 컬럼-지표 변환 오차로부터 자유롭고 1km 격자에서 미세 구배를 분해하며 광역 측정소 관측을 학습 신호로 활용하므로 위성 자료 단독 분석보다 지표 분배 구조의 절대 수준 추정에서 더 높은 신뢰성을 가질 수 있다. 두 자료원은 상호 보완적으로 활용 가능하며 측정소가 부재한 영역에서는 위성 FNR이 DL r 공간 패턴의 독립 검증 자료원으로 기능하고 두 자료원의 민감도 체제 분류가 일치하는 0.1° 격자는 신뢰도 높은 진단 영역으로 활용할 수 있을 것으로 판단된다.

3.3.3 동서 횡단면 분석을 통한 민감도 전이 구조

앞서 분석한 공간적 산화제 분배 구조(r)와 민감도 체제(FNR) 분류 결과를 종합하여 서해-인천-서울-

Table 4. Cross-source pairwise classification agreement and directional consistency (May~Sep daytime, 2023).

Category	Comparison pair	Agreement (%)	Spearman ρ	n
Surface-Surface	Station r vs DL r	90.2	+ 0.957	305
Surface-Column	DL r vs GEMS FNR	57.7	-0.502	613
	DL r vs TROPOMI FNR	67.0	-0.529	613
Column-Column	GEMS 10~16 KST vs TROPOMI FNR	87.6	+ 0.882	780

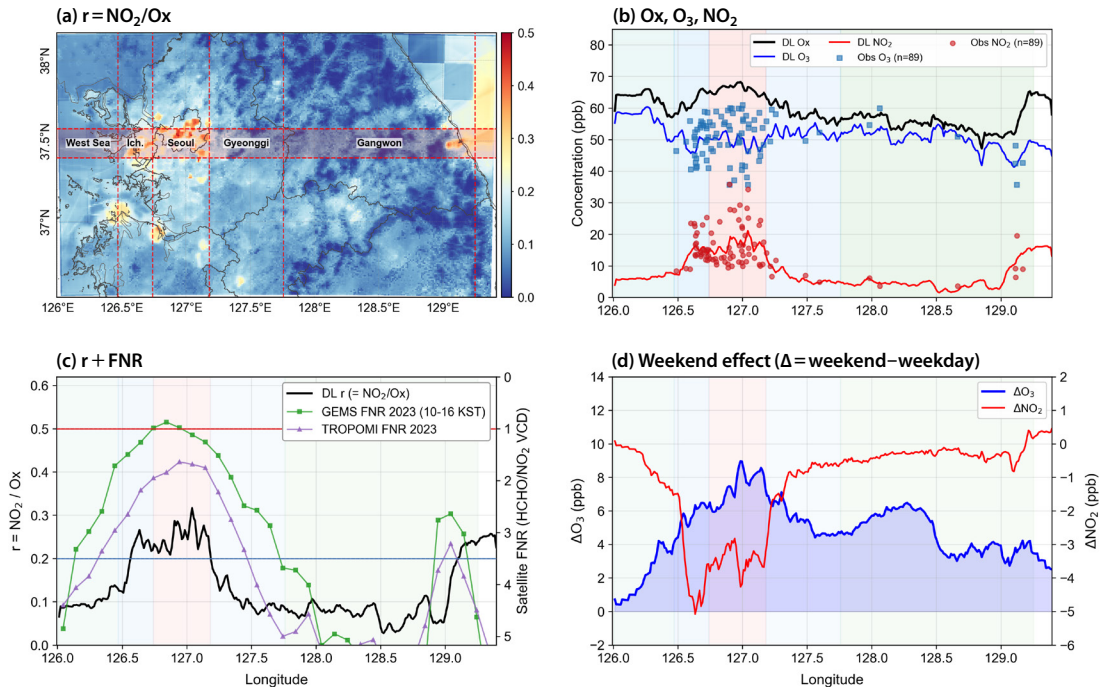


Fig. 9. W-E transect along 37.4~37.6°N: (a) 1 km r ($= \text{NO}_2/\text{Ox}$; May~Sep daytime) with transect band, (b) band-averaged DL NO_2 , O_3 , Ox profiles with in-band AirKorea observations, (c) band-averaged r (left axis) and dual-satellite FNR (right axis, inverted; GEMS 10~16 KST and TROPOMI, 0.1° bins; threshold lines at FNR = 1.0 and 3.5), (d) band-averaged weekend minus weekday ΔO_3 (left) and ΔNO_2 (right) (JJA daytime).

경기-강원을 관통하는 37.4~37.6°N 밴드의 동서 횡단면으로 분배 구조와 민감도 체제의 공간 전이를 통합적으로 분석하였다(그림 9). 본 절에서 제시하는 모든 밴드 프로파일 수치는 37.4~37.6°N 위도 범위에 대해 평균한 경도별 값이다.

밴드 평균 농도 프로파일에서 Ox ($= \text{NO}_2 + \text{O}_3$)는 도시(인천-서울, ~64 ppb)에서 강원(~56 ppb)으로 갈수록 감소하나 횡단면 300개 경도 격자에 대한 공간 변동계수(CV)는 Ox 가 8.2% (평균 59.1 ppb, 표준편차 4.8 ppb)로 NO_2 의 62.0% (평균 8.1 ppb, 표준편차 5.0 ppb)에 비해 현저히 억제되어 있다. 특히 도시 구간에서 NO_2 와 O_3 는 뚜렷한 반비례 관계를 보여, Ox 내 분배가 적정에 의해 조절되고 있음이 확인되었으며 이는 3.2.1절의 도시 Ox 보존성 결과와 일관됨을 보여준다. 분배 구조 지표 r ($= \text{NO}_2/\text{Ox}$)은 서울 핵심부(126.8~127.2°E)에서 최대 0.32(전이)에 도달한 뒤, 서울 동단

(~127.2°E, $r \approx 0.18$)에서 경기 동부(127.6°E, $r < 0.10$)까지 급감하였다.

동일 횡단면에서 GEMS FNR (0.1° 경도 구간 평균, 5~9월 10~16 KST, $\text{QF} = 0$)을 중첩한 결과, FNR은 서울 서부에서 횡단면 최저값(~0.88)에 도달한 뒤 강원으로 갈수록 증가(5 이상)하여 r 과 역방향으로 일관된 경도 구배를 나타냈다. 오존 생성 민감도 3단 분류($\text{FNR} < 1$: VOC 민감, 1~3.5: 전이, > 3.5 : NO_x 민감)를 적용하면 서울 서부가 VOC 민감에, 서울 중·동부 및 수도권 외곽이 전이에, 강원권 심부는 NO_x 민감 방향의 양상을 보였다. 이는 3.3.2절의 FNR- r Spearman 상관($\rho = -0.502$)을 공간적으로 확인하는 것이다.

다만 도시 고 NO_x 환경에서는 컬럼-경계층 변환 오차(Souri *et al.*, 2023)와 고 NO_x 조건에서 NO_2 가 HCHO 생성을 억제하는 화학적 피드백(Souri *et al.*, 2020)에 의해 FNR이 실제보다 낮게 산출되어 VOC 민

감 방향으로 편향될 수 있으므로 절대값 해석에는 유의가 필요하다. 동일 횡단면의 주말-주중 O_3 차이(그림 9d)에서도 서울 도심에서 $\Delta O_3 \approx +9$ ppb, $\Delta NO_2 \approx -5$ ppb로 강한 주말효과가 나타나고 강원으로 갈수록 약화되어 r 구배 및 FNR 구배와 공간적으로 일관된 정도 패턴을 보였다.

3.4 다년도 평가

산화제 분배 구조의 다년도 일관성을 확인하기 위해 2022년 자료로 학습한 모델을 각 연도의 입력 자료에 적용하여 2020~2023년(4개년)의 고해상도 농도 추론을 수행하였다(그림 10).

본 연구는 연산자원의 제약으로 다년도 동시 학습 대신 2022년 단년도 학습 후 각 연도의 입력 자료에 적용하는 방식의 다년도 추론을 수행하였다(보충자료 표 S1). 이 설정에서 학습 시기와 다른 연도로의 적용 가능성은 다음 두 가지 가정에 의존한다. 첫째, 단년도 학습이라 하더라도 1년 자료는 4계절에 걸친 다양한

대기환경 현상을 포괄하므로 모델이 학습한 입력-농도 관계(CTM 배경장의 편향 보정 $\alpha \cdot \beta$, 배출-농도 비 선형성, 측정소 관측의 공간 외삽)가 학습 연도 외에서도 유지된다. 둘째, 측정소 관측 자료를 매 연도 입력으로 활용하므로 모델은 연도별 NO_x 배출 변화·기상 조건 변화를 관측 정보를 통해 흡수할 수 있다.

수도권 도시의 r ($=NO_2/Ox$) 중앙값은 4개년에 걸쳐 감소하는 경향을 보였으며 이는 동기간 CAPSS NO_x 배출량 감소 추세와 방향적으로 일관된다. CAPSS 시도별 배출량 통계(NAEIRC, 2025) 기준 수도권(서울·인천·경기) NO_x 배출량은 2020년 약 262 kt에서 2023년 약 212 kt로 약 19% 감소하였고 r 추세는 이와 방향적으로 부합하였다. 반면 강원도 NO_x 배출량은 동기간 약 9% 감소(59 kt $\rightarrow$ 54 kt)에 그쳐 인위적 NO_x 배출이 희소한 배경 환경에서 r 이 일관되게 낮은 수준을 유지하였으며(그림 10d), 이러한 도시-배경 간 NO_x 배출량 저감 격차는 r 구배 변화 방향과 부합한다. 도시-산림 간 r 구배는 Δr 0.174 $\rightarrow$ 0.134로 23%

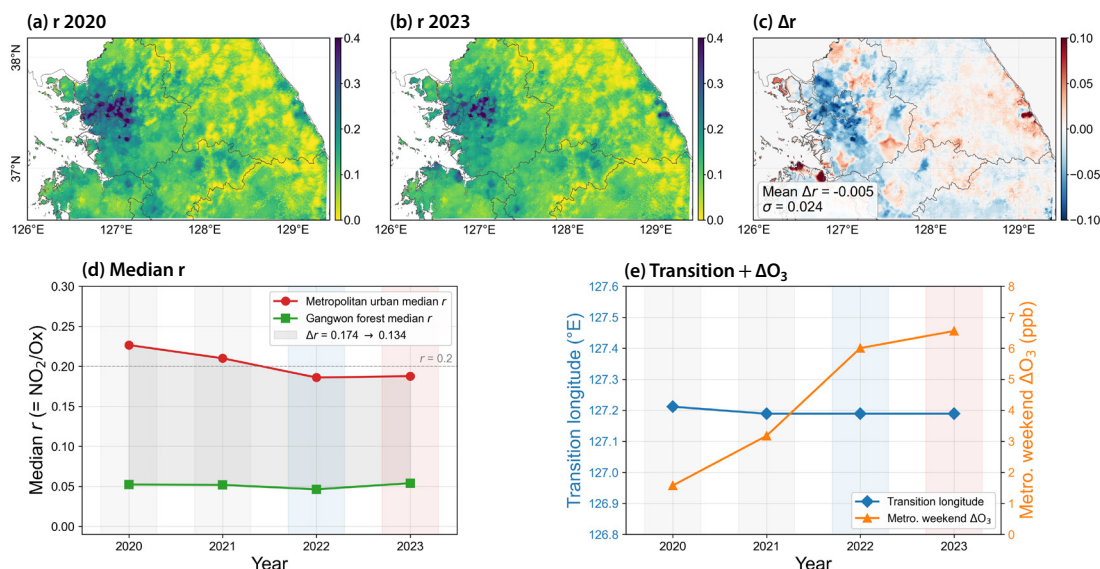


Fig. 10. Interannual variation of surface oxidant partitioning structure (2020~2023): (a) 2020 and (b) 2023 May~Sep daytime r ($=NO_2/Ox$) maps over South Korea land grids (sido administrative boundary $\cap$ IGBP $\neq 20$; North Korea and ocean masked), (c) pixel-wise Δr (2023 minus 2020) over the same domain with annotated domain-mean and standard deviation, (d) metropolitan urban and Gangwon forest median r with Δr band (4 years), (e) transition longitude ($r=0.2$ boundary) and metropolitan weekend ΔO_3 .

축소되었으며 이 또한 동기간 수도권 NO_x 배출량 감소 방향과 부합하는 관찰이다.

다만 GEOS-CF 배경장의 연도별 편향 변동과 CAPSS 상향식 산정의 활동량·배출계수 갱신에 따른 연도별 불확실성 차이가 입력 자료에 내재되어 있어 이러한 방향적 부합의 절대 수준 해석에는 잔여 한계가 따른다. 연산자원이 제약되지 않을 경우 다년도 자료의 동시 학습을 통해 학습 도메인을 입력 분포 전체로 확장함으로써 이러한 한계 완화가 가능할 것으로 판단된다.

수도권의 분배 전환 정도($r=0.2$ 등치선)는 4개년에 걸쳐 소폭 서진하는 경향을 나타내어 고적정 NO₂ 우세 영역이 서울 중심부로 수축하는 것으로 해석된다(그림 10e). 이러한 경향은 Lee *et al.* (2024b)이 2016~2018년 수도권에서 OH 반응성 비 분석을 통해 서울의 NO_x 포화가 약화되고 인천이 전이 경계에 근접하였음을 보고한 결과의 연장선에 있으며 본 연구의 2020~2023년 추론 결과는 이 추세가 이후에도 지속되고 있음을 시사한다. 한편 수도권 주말 ΔO_3 는 2020~2021년에 상대적으로 낮았는데, 이는 COVID-19 기간 주중-주말 활동 격차 축소로 주말효과가 약화된 선행 사례(Gough and Anderson, 2022; Wu *et al.*, 2022)와 부합하며 2022~2023년에는 활동 회복과 함께 주말 효과가 증가하는 경향이 관찰되었다. 국내에서는 관련 연구가 제한적이나, 본 연구의 2020~2021년 결과 또한 활동 변동의 영향 가능성을 시사하는 국내 사례로 볼 수 있다. 다만 2020~2021년은 COVID-19 활동 변화로 인해 학습 데이터 분포의 외삽 영역에 해당하므로 본 절의 결과는 NO_x 배출 변화의 직접 증거가 아니라 산화제 분배 구조의 다년도 일관성 점검 차원에서 제시된다.

3.5 한계와 후속 과제

본 연구의 주요 한계는 다음과 같다. 첫째, 위성 FNR(컬럼)과 지상 r (지표 농도)은 본질적으로 다른 물리량으로 절대값 일치가 아닌 방향성 중심의 비교만 가능하며 위성의 맑은 날 편향과 픽셀 해상도(수

km) 한계로 1 km 세부 경계의 직접 검증이 불가능하다. 위성 산출에는 평균화 커널 비식별성에 따른 경계층 감도 저하, 운량·태양천정각에 따른 0.1° 격자별 표본의 시공간 비균질성, 도시 고NO_x 환경에서 NO₂가 HCHO 생성을 억제하는 화학적 피드백(3.3절)도 내재하여 절대값 해석에 추가적 제약이 따른다. 또한 FNR 3분류 임계값(1/3.5)은 선행 문헌의 경험적 기준이고 r 3분류 임계값(0.2/0.5)은 본 연구에서 도입한 운영적 기준(2.3.3절)으로 두 기준 모두 보편적 화학 임계값으로 확립된 것은 아니므로 임계값에 의한 경계 구분을 절대적으로 해석하는 데에는 한계가 있다. 둘째, 수도권 도시에서 O₃의 체계적 저평가(3.2.3절)는 도시 환경의 높은 시공간 변동성과 간헐적 고농도 현상을 학습 과정에서 충분히 포착하지 못하는 데이터 기반 모델의 구조적 특성과 격자 평균·지점 측정 간 공간 대표성 차이가 함께 기여한 것으로 판단된다. 이러한 O₃ 저평가는 O_x 과소를 통해 도시 지역 r 의 소폭 과대 방향으로 작용할 수 있으며 공간 교차 검증에서의 NO₂ R² 하락(0.539)과 함께 관측 희소 지역의 분배비 산출 불확실성을 구성한다. 강원 산림에서 관찰된 O₃ 과대평가는 4개소 측정소 표본의 공간 대표성 제약뿐 아니라 데이터 기반 모델이 BVOC 매개 NO_x-VOC 비선형 광화학을 입력 변수 관계로만 학습하는 구조적 한계(3.2.2절)와도 연관된다. 또한 해양 격자에서 관찰되는 격자 경계 불연속(3.2.1절)은 측정소 부재에 따른 공간 보정 한계와 학습 도메인 경계 효과를 반영한다. 이러한 추론 품질의 공간적 불균일성을 완화하기 위해 위성 보조 손실 도입이나 광역 배경장과 국지 세부 구조를 분리 학습하는 다단계 아키텍처로의 고도화가 필요하다. 셋째, 인위적 배출 입력인 CAPSS는 상향식 산정에 내재하는 배출원별 활동량·배출계수의 불확실성을 포함하며 1 km 격자 수준에서의 공간 배분 오차가 고해상도 농도 추론에 직접 전파될 수 있다. 넷째, R²/RMSE에 의한 직접 검증은 309개 측정소 위치(도메인 54,000 격자의 약 0.6%)에서만 가능하며 본 연구의 화학적 경향 비교는 나머지 격자의 공간 패턴을 해석하기 위한 보조 근거이지 미관측 지역의

절대 농도 정확도를 입증하는 것은 아니다. 또한 본 산출장은 배출 섭동에 대한 농도 응답을 직접 산정하지 않으므로 특정 배출원에 대한 저감 효과를 정량적으로 추정하는 데에는 한계가 있다. 다섯째, 본 연구는 주말효과 분석에서 기상과 배출 효과를 명시적으로 분리하지 않았다. 따라서 도심 ΔO_3 의 정량값에는 기상 변동에 의한 잔여 영향이 일부 남을 가능성이 있으며 광역 기상 정규화 또는 병행 CTM 시뮬레이션을 활용한 정량화는 후속 연구의 과제에 해당한다. 여섯째, 본 연구는 연산자원 제약으로 2022년 단년도 학습 후 각 연도 입력에 적용하는 방식으로 다년도 추론을 수행하였으며(3.4절) 특히 2020~2021년은 COVID-19 활동 변화로 학습 분포의 외삽 영역에 해당하여 직접 증거로 해석하기에는 한계가 있다.

본 연구의 적응형 편향 보정(식 (3))의 $\alpha \cdot \beta$ 는 학습 가능 형식이지만 전지구 CTM(GEOS-CF)의 1 km 추론에 대한 해상도 격차와 GEOS-CF 자체 편향으로 인해 식별 가능성이 제한적이며 실제 농도장의 1 km 변동은 잔차 δ 가 분해한다. 보다 신뢰성 있는 1 km 농도장 추론을 위해서는 향후 한반도 자료동화 산출물 등 관측 정합성과 공간 해상도가 보강된 배경장을 활용하여 적응형 편향 보정의 식별 가능성과 물리적 해석 가능성을 재검점하는 후속 연구가 필요하다. 또한 토지피복별 Ox 보존성 평가에서 광역 배경 구매와 국지적 정적 효과를 정량적으로 분리하기 위한 사전 공간 추세 제거·하위 영역 단위 분석 등 방법론적 정제도 후속 과제로 남는다(3.2.1절).

4. 결 론

본 연구는 그래프 신경망 기반 1 km NO_2 - O_3 DL 추론 농도장이 측정소 위치 밖에서도 대기화학적으로 일관된 공간 구조를 표현하는지를 평가하였다. 이를 위해 수도권·강원권을 대상으로 광화학적으로 활발한 기간을 중심으로 다각적 분석을 수행하였다.

주요 결과는 세 가지로 요약할 수 있다. 첫째, $NO_2 \cdot$

$O_3 \cdot NO$ 의 시간 독립 검증과 5-fold 공간 교차 검증 모두에서 관측과의 일치도가 확보되어(대표적으로 O_3 시간 독립 $R^2=0.926$, 공간 교차 $R^2=0.732$) 후속 화학적 정합성 평가의 전제가 되는 관측 재현력이 확인되었다. 둘째, Ox 보존성, NOx - Ox 회귀, 주말효과, 도시-산립 O_3 시간 구조, 1 km 분배비 $r(=NO_2/Ox)$ 공간 구매가 모두 수도권 도시와 강원 산림 간 상이한 산화제 분배 경향을 일관되게 드러내었고 서울 도심 핵심부에 국한된 전이 영역의 공간적 집중은 한국의 대부분 지역이 NOx 제한에 가깝다는 최근 보고와도 방향적으로 부합하였다. 다중 자료 교차 정합성에서 DL r 은 지상 관측 분류와 약 90%의 일치율을 보였고 두 위성 FNR(GEMS, TROPOMI)과의 Spearman ρ 가 모두 -0.5 수준의 방향적 일관성을 유지하였다. 셋째, 2020~2023년 다년도 평가에서 수도권 r 중앙값이 NOx 배출량 감소 추세와 일관되게 감소하고 분배 전환 정도($r=0.2$ 등치선)가 서울 중심부로 소폭 서진하는 경향이 나타났으며 COVID-19 기간의 주말효과 약화와 2022~2023년 회복이 함께 관찰되어 배출 저감과 사회적 활동 변화가 산화제 분배 구조에 반영됨을 확인하였다.

본 연구는 다음 세 측면에서 방법론적·응용적 기여를 갖는다. 첫째 평가 방법론 측면에서, 기존 통계 지표(R^2 ·RMSE) 중심의 평가를 Ox 보존성· NOx - Ox 회귀·주말효과·다중 자료 교차 정합성을 포괄하는 화학적 정합성 검증으로 확장하였으며 1 km 해상도에서 측정소 밖 격자의 분배 구조 일관성을 지상·DL·위성 교차 비교로 점검하는 절차를 정립하였다. 둘째 모델링 방법론 측면에서, 광역 물리 사전정보(CTM·기상의 5 km 보간), 국지 정적 자료(지형·배출·토지피복의 1 km 격자), 비균질·공간 회소 정답(322개 측정소 점관측)의 서로 다른 공간 단위·해상도·밀도를 단일 그래프 신경망 위에서 융합하여 1 km 농도장을 산출한 사례는 비균질한 측정망 환경에서 다중 자료 융합형 고해상도 농도 추론의 방법론적 참고가 될 수 있다. 셋째 응용 측면에서, 도출된 r 분배 지도는 도시-배경 간 화학 환경 차이를 1 km 분해능에서 가시화함으로써

써 시·군·구 단위 NO_x 저감 우선순위나 광역 배경 영향이 큰 지역 식별 등 후속 정책 검토의 보조 근거로 활용될 수 있다.

본 연구는 별도의 고해상도 자료동화 산출물이나 국지 확산 모델링 없이 기 배포된 자료만으로 1 km 농도장을 산출하였으며 향후 한반도 고해상도 자료동화 농도장 등 품질이 보장된 사전 정보가 가용해질 경우 적응형 편향 보정의 식별 가능성과 추론 정확도가 함께 개선될 가능성이 있다. 또한 연산자원 제약이 해소되는 경우 다년도 학습으로 시간 외삽 한계를 완화할 수 있으며 광역 배경장과 국지 세부 구조를 분리 학습하는 다단계 아키텍처나 위성 보조 손실의 도입을 통해 도시 환경의 시공간 변동성 포착력을 추가로 강화할 여지가 있다. 본 연구의 고해상도 DL 추론 농도장은 이러한 보완을 통해 관측 희소 지역 해석의 보조 자료로 활용 범위를 확장해 나가는 기초가 될 것으로 기대된다.

감사의 글

이 성과는 정부(기후에너지환경부)의 재원으로 한국환경산업기술원의 미세먼지관리 특성화대학원 사업의 지원을 받아 수행된 연구임.

References

- Ajayi, T., Mirrezaei, M.A., Arellano, A.F., Robinson, E.S., Sorooshian, A. (2025) A long-term (2001-2022) examination of surface ozone concentrations in Tucson, Arizona, *Environmental Science: Atmospheres*, 5, 1326-1340. <https://doi.org/10.1039/d5ea00072f>
- Bae, K., Song, C.-K., Van Roozendaal, M., Richter, A., Wagner, T., Merlaud, A., Pinardi, G., Friedrich, M.M., Fayt, C., Dimitropoulou, E., Lange, K., Bösch, T., Zilker, B., Latsch, M., Behrens, L.K., Ziegler, S., Ripperger-Lukosiunaite, S., Kuhn, L., Lauster, B., Reischmann, L., Uhlmannsieck, K., Cede, A., Tiefengraber, M., Gebetsberger, M., Park, R.J., Lee, H., Hong, H., Chang, L.-S., Jeon, K. (2025) Validation of GEMS operational v2.0 total column NO₂ and HCHO during the GMAP/SUJQA campaign, *Science of the Total Environment*, 974, 179190. <https://doi.org/10.1016/j.scitotenv.2025.179190>
- Baek, S.-H., Lee, H.-J., Kim, C.-H. (2023) Analysis of photochemical ozone formation regime in Busan - comparative study on Busan vs. Seoul metropolitan area (IV), *Journal of Environmental Science International*, 32, 285-301, (in Korean with English abstract). <https://doi.org/10.5322/jesi.2023.32.5.285>
- Brody, S., Alon, U., Yahav, E. (2022) How attentive are graph attention networks?, *Proceedings of the 10th International Conference on Learning Representations*. <https://doi.org/10.48550/arXiv.2105.14491>
- Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P. (2017) Geometric deep learning: Going beyond Euclidean data, *IEEE Signal Processing Magazine*, 34(4), 18-42. <https://doi.org/10.1109/MSP.2017.2693418>
- Choi, H., Hwang, S., Kang, E., Kim, Y., Yang, S., Malik, S., Keya, J.N., Lee, S., Im, J. (2025) Satellite-based air quality monitoring using artificial intelligence: Research trends and future perspectives, *Journal of Korean Society for Atmospheric Environment*, 41, 179-198, (in Korean with English abstract). <https://doi.org/10.5572/kosae.2025.41.2.179>
- Clapp, L.J., Jenkin, M.E. (2001) Analysis of the relationship between ambient levels of O₃, NO₂ and NO as a function of NO_x in the UK, *Atmospheric Environment*, 35(36), 6391-6405. [https://doi.org/10.1016/S1352-2310\(01\)00378-8](https://doi.org/10.1016/S1352-2310(01)00378-8)
- Colombi, N.K., Jacob, D.J., Yang, L.H., Zhai, S., Shah, V., Grange, S.K., Yantosca, R.M., Kim, S., Liao, H. (2023) Why is ozone in South Korea and the Seoul metropolitan area so high and increasing?, *Atmospheric Chemistry and Physics*, 23, 4031-4044. <https://doi.org/10.5194/acp-23-4031-2023>
- De Smedt, I., Theys, N., Yu, H., Danckaert, T., Lerot, C., Compennolle, S., Van Roozendaal, M., Richter, A., Hilboll, A., Peters, E., Pedergnana, M., Loyola, D., Beirle, S., Wagner, T., Eskes, H., van Geffen, J., Boersma, K.F., Veefkind, P. (2018) Algorithm theoretical baseline for formaldehyde retrievals from S5P TROPOMI and from the QA4ECV project, *Atmospheric Measurement Techniques*, 11, 2395-2426. <https://doi.org/10.5194/amt-11-2395-2018>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N. (2021) An image

- is worth 16×16 words: Transformers for image recognition at scale, International Conference on Learning Representations (ICLR). <https://doi.org/10.48550/arXiv.2010.11929>
- Duncan, B.N., Yoshida, Y., Olson, J.R., Sillman, S., Martin, R.V., Lam-sal, L., Hu, Y., Pickering, K.E., Retscher, C., Allen, D.J., Crawford, J.H. (2010) Application of OMI observations to a space-based indicator of NO_x and VOC controls on surface ozone formation, *Atmospheric Environment*, 44, 2213-2223. <https://doi.org/10.1016/j.atmosenv.2010.03.010>
- Friedl, M., Sulla-Menashe, D. (2022) MODIS/Terra+Aqua land cover type yearly L3 global 500m SIN grid V061, NASA Land Processes Distributed Active Archive Center. <https://doi.org/10.5067/MODIS/MCD12Q1.061> (accessed on Feb. 11, 2026).
- García-Yee, J.S., Torres-Jardón, R., Barrera-Huertas, H., Castro, T., Peralta, O., García, M., Gutiérrez, W., Robles, M., Torres-Jaramillo, J.A., Ortiz-Alvarez, A., Ruiz-Suárez, L.G. (2018) Characterization of NO_x-O_x relationships during daytime interchange of air masses over a mountain pass in the Mexico City megalopolis, *Atmospheric Environment*, 177, 100-110. <https://doi.org/10.1016/j.atmosenv.2017.11.017>
- Gomez Gonzalez, C.A. (2023) DL4DS - deep learning for empirical downscaling, *Environmental Data Science*, 2, e3. <https://doi.org/10.1017/eds.2022.26>
- Gough, W.A., Anderson, V. (2022) Changing air quality and the ozone weekend effect during the COVID-19 pandemic in Toronto, Ontario, Canada, *Climate*, 10, 41. <https://doi.org/10.3390/cli10030041>
- Granier, C., Darras, S., Denier van der Gon, H., Doubalova, J., Elguindi, N., Galle, B., Gauss, M., Guevara, M., Jalkanen, J.-P., Kuenen, J., Lioussé, C., Quack, B., Simpson, D., Sindelarova, K. (2019) The Copernicus Atmosphere Monitoring Service global and regional emissions (april 2019 version), ECMWF Copernicus Atmosphere Monitoring Service, CAMS-GLOB-BIO v3.1. <https://doi.org/10.24380/d0bn-kx16> (accessed on Feb. 11, 2026).
- Hamilton, W.L., Ying, R., Leskovec, J. (2017) Inductive representation learning on large graphs, *Advances in Neural Information Processing Systems* 30, Curran Associates, pp. 1024-1034. <https://doi.org/10.48550/arXiv.1706.02216>
- Han, J., Liu, H., Xiong, H., Yang, J. (2023) Semi-supervised air quality forecasting via self-supervised hierarchical graph neural network, *IEEE Transactions on Knowledge and Data Engineering*, 35(5), 5230-5243. <https://doi.org/10.1109/TKDE.2022.3149815>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R.J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P. de, Rozum, I., Vamborg, F., Villaume, S., Thépaut, J.-N. (2020) The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999-2049. <https://doi.org/10.1002/qj.3803>
- Hoek, G., Beelen, R., Hoogh, K. de, Vienneau, D., Gulliver, J., Fischer, P., Briggs, D. (2008) A review of land-use regression models to assess spatial variation of outdoor air pollution, *Atmospheric Environment*, 42(33), 7561-7578. <https://doi.org/10.1016/j.atmosenv.2008.05.057>
- Hu, Z., Dong, Y., Wang, K., Sun, Y. (2020) Heterogeneous graph transformer, in *Proceedings of The Web Conference 2020 (WWW '20)*, ACM, pp. 2704-2710. <https://doi.org/10.1145/3366423.3380027>
- Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J.J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., Suttie, M. (2019) The CAMS reanalysis of atmospheric composition, *Atmospheric Chemistry and Physics*, 19(6), 3515-3556. <https://doi.org/10.5194/acp-19-3515-2019>
- Jenkin, M.E. (2004) Analysis of sources and partitioning of oxidant in the UK - part 1: The NO_x-dependence of annual mean concentrations of nitrogen dioxide and ozone, *Atmospheric Environment*, 38(30), 5117-5129. <https://doi.org/10.1016/j.atmosenv.2004.05.056>
- Jeong, M., Kim, W.J., Heo, J., Kim, S.-Y. (2022) Exploration of the spatial and temporal patterns of surface ozone concentrations for development of ozone prediction model in South Korea, *Journal of Korean Society for Atmospheric Environment*, 38, 100-125, (in Korean with English abstract). <https://doi.org/10.5572/kosae.2022.38.1.100>
- Jin, X., Fiore, A., Boersma, K.F., De Smedt, I., Valin, L. (2020) Inferring changes in summertime surface ozone-NO_x-VOC chemistry over U.S. Urban areas from two decades of satellite and ground-based observations, *Environmental Science & Technology*, 54, 6518-6529. <https://doi.org/10.1021/acs.est.9b07785>
- Kaiser, J.W., Heil, A., Andreae, M.O., Benedetti, A., Chubarova, N.,

- Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M.G., Suttie, M., van der Werf, G.R. (2012) Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, *Bio-geosciences*, 9(1), 527-554. <https://doi.org/10.5194/bg-9-527-2012>
- Keller, C.A., Knowland, K.E., Duncan, B.N., Liu, J., Anderson, D.C., Das, S., Lucchesi, R.A., Lundgren, E.W., Nicely, J.M., Nielsen, E., Ott, L.E., Saunders, E., Strode, S.A., Wales, P.A., Jacob, D.J., Pawson, S. (2021) Description of the NASA GEOS Composition Forecast Modeling System GEOS-CF v1.0, *Journal of Advances in Modeling Earth Systems*, 13(4), e2020MS002413. <https://doi.org/10.1029/2020MS002413>
- Kim, B.-U., Kim, H.C., Kim, S. (2019) Review on ozone management in US and recommendations for domestic ozone control, *Journal of Korean Society for Atmospheric Environment*, 35, 370-394, (in Korean with English abstract). <https://doi.org/10.5572/kosae.2019.35.3.370>
- Kim, C.H. (2025) Overview and prospects of long-range transport of air pollutants from the perspectives of modeling studies, *Journal of Korean Society for Atmospheric Environment*, 41(2), 161-178, (in Korean with English abstract). <https://doi.org/10.5572/KOSAE.2025.41.2.161>
- Kim, D.-S., Jeong, J., Ahn, J. (2016) Characteristics in atmospheric chemistry between NO, NO₂ and O₃ at an urban site during MAPS (Megacity Air Pollution Study)-Seoul, Korea, *Journal of Korean Society for Atmospheric Environment*, 32(4), 422-434, (in Korean with English abstract). <https://doi.org/10.5572/KOSAE.2016.32.4.422>
- Kim, J., Jeong, U., Ahn, M.-H., Kim, J.H., Park, R.J., Lee, H., Song, C.H., Choi, Y.-S., Lee, K.-H., Yoo, J.-M., Jeong, M.-J., Park, S.K., Lee, K.-M., Song, C.-K., Kim, S.-W., Kim, Y.J., Kim, S.-W., Kim, M., Go, S., Liu, X., Chance, K., Chan Miller, C., Al-Saadi, J., Veihelmann, B., Bhartia, P.K., Torres, O., Abad, G.G., Haffner, D.P., Ko, D.H., Lee, S.H., Woo, J.-H., Chong, H., Park, S.S., Nicks, D., Choi, W.J., Moon, K.-J., Cho, A., Yoon, J., Kim, S., Hong, H., Lee, K., Lee, H., Lee, S., Choi, M., Veefkind, P., Levelt, P.F., Edwards, D.P., Kang, M., Eo, M., Bak, J., Baek, K., Kwon, H.-A., Yang, J., Park, J., Han, K.M., Kim, B.-R., Shin, H.-W., Choi, H., Lee, E., Chong, J., Cha, Y., Koo, J.-H., Irie, H., Hayashida, S., Kasai, Y., Kanaya, Y., Liu, C., Lin, J., Crawford, J.H., Carmichael, G.R., Newchurch, M.J., Lefer, B.L., Herman, J.R., Swap, R.J., Lau, A.K.H., Kurosu, T.P., Jaross, G., Ahlers, B., Dobber, M., McElroy, C.T., Choi, Y. (2020) New era of air quality monitoring from space: Geostationary Environment Monitoring Spectrometer (GEMS), *Bulletin of the American Meteorological Society*, 101, E1-E22. <https://doi.org/10.1175/bams-d-18-0013.1>
- Kipf, T.N., Welling, M. (2017) Semi-supervised classification with graph convolutional networks, *International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/1609.02907>
- Koo, B., Jung, J., Pollack, A.K., Lindhjem, C., Jimenez, M., Yarwood, G. (2012) Impact of meteorology and anthropogenic emissions on the local and regional ozone weekend effect in Midwestern US, *Atmospheric Environment*, 57, 13-21. <https://doi.org/10.1016/j.atmosenv.2012.04.043>
- Lange, K., Richter, A., Bösch, T., Zilker, B., Latsch, M., Behrens, L.K., Okafor, C.M., Bösch, H., Burrows, J.P., Merlaud, A., Pinardi, G., Fayt, C., Friedrich, M.M., Dimitropoulou, E., Van Roozendaal, M., Ziegler, S., Ripperger-Lukosiun-aite, S., Kuhn, L., Lauster, B., Wagner, T., Hong, H., Kim, D., Chang, L.-S., Bae, K., Song, C.-K., Park, J.-U., Lee, H. (2024) Validation of GEMS tropospheric NO₂ columns and their diurnal variation with ground-based DOAS measurements, *Atmospheric Measurement Techniques*, 17, 6315-6344. <https://doi.org/10.5194/amt-17-6315-2024>
- Lee, G., Park, J.-H., Kim, D.-G., Koh, M.S., Lee, M., Han, J.-S., Kim, J.-C. (2020) Current status and future directions of tropospheric photochemical ozone studies in Korea, *Journal of Korean Society for Atmospheric Environment*, 36, 419-441, (in Korean with English abstract). <https://doi.org/10.5572/kosae.2020.36.4.419>
- Lee, G.T., Park, R.J., Kwon, H.-A., Ha, E.S., Lee, S.D., Shin, S., Ahn, M.-H., Kang, M., Choi, Y.-S., Kim, G., Lee, D.-W., Kim, D.-R., Hong, H., Langerock, B., Vigouroux, C., Lerot, C., Hendrick, F., Pinardi, G., De Smedt, I., Van Roozendaal, M., Wang, P., Chong, H., Cho, Y., Kim, J. (2024a) First evaluation of the GEMS formaldehyde product against TROPOMI and ground-based column measurements during the in-orbit test period, *Atmospheric Chemistry and Physics*, 24, 4733-4749. <https://doi.org/10.5194/acp-24-4733-2024>
- Lee, K.-H., Shin, S.-K. (2025) Impact of monitoring site density on air quality time series: Trend variability across spatial resolutions, *Journal of Korean Society for Atmospheric Environment*, 41, 845-861, (in Korean with English abstract). <https://doi.org/10.5572/kosae.2025.41.5.845>
- Lee, S.-H., Lee, J.-S., Lee, H.-M., Kwon, H.-A., Nowlan, C.R., González Abad, G., Park, R.J., Park, J.-H., Kang, S., Shin,

- H.-J., Kim, D.-G., Park, J.-M. (2024b) Understanding ozone formation and carbonyl contributions in the Seoul Metropolitan Area, *Aerosol and Air Quality Research*, 24, 230227. <https://doi.org/10.4209/aaqr.230227>
- Li, L., Wang, J., Franklin, M., Yin, Q., Wu, J., Camps-Valls, G., Zhu, Z., Wang, C., Ge, Y., Reichstein, M. (2023) Improving air quality assessment using physics-inspired deep graph learning, *npj Climate and Atmospheric Science*, 6, 152. <https://doi.org/10.1038/s41612-023-00475-3>
- Martin, R.V., Fiore, A.M., van Donkelaar, A. (2004) Space-based diagnosis of surface ozone sensitivity to anthropogenic emissions, *Geophysical Research Letters*, 31(6), L06120. <https://doi.org/10.1029/2004GL019416>
- Mijling, B. (2020) High-resolution mapping of urban air quality with heterogeneous observations: A new methodology and its application to Amsterdam, *Atmospheric Measurement Techniques*, 13(8), 4601-4617. <https://doi.org/10.5194/amt-13-4601-2020>
- NASA JPL (2020) NASADEM merged DEM global 1 arc second V001, NASA Land Processes Distributed Active Archive Center. https://doi.org/10.5067/MEASURES/NASADEM/NASADEM_HGT.001 (accessed on Feb. 11, 2026).
- National Air Emission Inventory and Research Center (NAEIRC) (2025) Air pollutant emission statistics (CAPSS): Annual inventory by sigungu, source sub-classification, and fuel type, 2020-2023, (in Korean). <https://www.air.go.kr> (accessed on Feb. 11, 2026).
- National Institute of Environmental Research (NIER) (2025) Annual report of air quality in Korea 2024, (in Korean). https://www.airkorea.or.kr/web/detailViewDown?pMENU_NO=125 (accessed on Feb. 11, 2026).
- Oak, Y.J., Jacob, D.J., Pendergrass, D.C., Dang, R., Colombi, N.K., Chong, H., Lee, S., Kuk, S.K., Kim, J. (2025) Air quality trends and regimes in South Korea inferred from 2015-2023 surface and satellite observations, *Atmospheric Chemistry and Physics*, 25, 3233-3252. <https://doi.org/10.5194/acp-25-3233-2025>
- Oh, I., Hwang, M.-K., Bang, J.-H., Yang, W., Kim, S., Lee, K., Seo, S., Lee, J., Kim, Y. (2021) Comparison of different hybrid modeling methods to estimate intraurban NO₂ concentrations, *Atmospheric Environment*, 244, 117907. <https://doi.org/10.1016/j.atmosenv.2020.117907>
- Park, J., Kim, E., Kang, Y.-H., Kim, S. (2024) Assessment of provincial air quality based on air quality index during 2016-2022, *Journal of Korean Society for Atmospheric Environment*, 40, 225-241, (in Korean with English abstract). <https://doi.org/10.5572/kosae.2024.40.2.225>
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat (2019) Deep learning and process understanding for data-driven Earth system science, *Nature*, 566(7743), 195-204. <https://doi.org/10.1038/s41586-019-0912-1>
- Ryu, Y.-H., Min, S.-K. (2021) Long-term evaluation of atmospheric composition reanalyses from CAMS, TCR-2, and MERRA-2 over South Korea: Insights into applications, implications, and limitations, *Atmospheric Environment*, 246, 118062. <https://doi.org/10.1016/j.atmosenv.2020.118062>
- Ryu, Y.-H., Min, S.-K., Hodzic, A. (2021) Recent decreasing trends in surface PM_{2.5} over East Asia in the winter-spring season: Different responses to emissions and meteorology between upwind and downwind regions, *Aerosol and Air Quality Research*, 21(6), 200654. <https://doi.org/10.4209/aaqr.200654>
- Shen, S., Li, C., van Donkelaar, A., Jacobs, N., Wang, C., Martin, R.V. (2024) Enhancing global estimation of fine particulate matter concentrations by including geophysical a priori information in deep learning, *ACS ES&T Air*, 1(5), 332-345. <https://doi.org/10.1021/acsestair.3c00054>
- Sicard, P., Paoletti, E., Agathokleous, E., Araminiené, V., Proietti, C., Coulibaly, F., De Marco, A. (2020) Ozone weekend effect in cities: Deep insights for urban air pollution control, *Environmental Research*, 191, 110193. <https://doi.org/10.1016/j.envres.2020.110193>
- Souri, A.H., Johnson, M.S., Wolfe, G.M., Crawford, J.H., Fried, A., Wisthaler, A., Brune, W.H., Blake, D.R., Weinheimer, A.J., Verhoelst, T., Compernelle, S., Pinardi, G., Vigouroux, C., Langerock, B., Choi, S., Lamsal, L., Zhu, L., Sun, S., Cohen, R.C., Min, K.-E., Cho, C., Philip, S., Liu, X., Chance, K. (2023) Characterization of errors in satellite-based HCHO/NO₂ tropospheric column ratios with respect to chemistry, column-to-PBL translation, spatial representation, and retrieval uncertainties, *Atmospheric Chemistry and Physics*, 23, 1963-1986. <https://doi.org/10.5194/acp-23-1963-2023>
- Souri, A.H., Nowlan, C.R., Wolfe, G.M., Lamsal, L.N., Chan Miller, C.E., Abad, G.G., Janz, S.J., Fried, A., Blake, D.R., Weinheimer, A.J., Diskin, G.S., Liu, X., Chance, K. (2020) Revisiting the effectiveness of HCHO/NO₂ ratios for inferring ozone sensitivity to its precursors using high resolution airborne remote sensing observations in a high ozone episode during the KORUS-AQ campaign, *Atmospheric Environment*, 224, 117341. <https://doi.org/10.1016/j.atmosenv.2020.117341>
- Terroso-Saenz, F., Morales-García, J., Muñoz, A. (2024) Nationwide

- air pollution forecasting with heterogeneous graph neural networks, *ACM Transactions on Intelligent Systems and Technology*, 15(1), 1-19. <https://doi.org/10.1145/3637492>
- Tian, Y., Wang, S., Jin, X. (2025) Global patterns and trends in ground-level ozone chemical formation regimes from 1996 to 2022, *Atmospheric Chemistry and Physics*, 25, 9127-9149. <https://doi.org/10.5194/acp-25-9127-2025>
- van Geffen, J., Eskes, H., Compernelle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., Linden, M. ter, Ludewig, A., Boersma, K.F., Veefkind, J.P. (2022) Sentinel-5P TROPOMI NO₂ retrieval: Impact of version v2.2 improvements and comparisons with OMI and ground-based data, *Atmospheric Measurement Techniques*, 15, 2037-2060. <https://doi.org/10.5194/amt-15-2037-2022>
- Veefkind, J.P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H.J., Haan, J.F. de, Kleipool, Q., Weele, M. van, Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., Levelt, P.F. (2012) TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, *Remote Sensing of Environment*, 120, 70-83. <https://doi.org/10.1016/j.rse.2011.09.027>
- Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., Bengio, Y. (2018) Graph attention networks, *Proceedings of the 6th International Conference on Learning Representations*. <https://doi.org/10.48550/arXiv.1710.10903>
- Wang, X., Ji, H., Shi, C., Wang, B., Ye, Y., Cui, P., Yu, P.S. (2019) Heterogeneous graph attention network, *The World Wide Web Conference (WWW '19)*, ACM, pp. 2022-2032. <https://doi.org/10.1145/3308558.3313562>
- Wu, L., Xie, J., Kang, K. (2022) Changing weekend effects of air pollutants in Beijing under 2020 COVID-19 lockdown controls, *npj Urban Sustainability*, 2(1), 23. <https://doi.org/10.1038/s42949-022-00070-0>
- Yeo, M.J., Kim, Y.P. (2021) Long-term trends of surface ozone in Korea, *Journal of Cleaner Production*, 294, 125352. <https://doi.org/10.1016/j.jclepro.2020.125352>
- Yu, S.J., Son, Y.S. (2021) A review on characteristics of biogenic volatile organic compounds emitted from major tree species in South Korea, *Journal of Korean Society for Atmospheric Environment*, 37(5), 695-709, (in Korean with English abstract). <https://doi.org/10.5572/KOSAE.2021.37.5.695>

Authors Information

정필수 (안양대학교 일반대학원 환경공학과 박사수료)
(pala77@empas.com)

손승민 (안양대학교 일반대학원 환경공학과 박사수료)
(eia2000@empas.com)

왕경희 (안양대학교 일반대학원 환경공학과 산학협력중점교수)
(kyunghui96@gmail.com)

김현수 (안양대학교 일반대학원 환경공학과 박사수료)
(gustn1116@gmail.com)

최대련 (안양대학교 환경에너지공학과 조교수)
(drchoi@anyang.ac.kr)

윤희영 (안양대학교 환경에너지공학과 조교수)
(huiyoung@anyang.ac.kr)

Supplementary Materials

Table S1. Training hyperparameters of AMR-HGN.

Parameter	Value	Note
Learning rate	3×10^{-4}	Cosine annealing
Batch size	2 (gradient accumulation 4)	Effective batch = 8
Epochs	30	Early stopping patience = 10
Hidden dimension	128	All nodes shared
GNN layers	2 (bottom-up) + 2 (top-down)	4 layers total
Station dropout	$p = 0.15$	Observation masking (tested 0.1~0.2)
Emission dropout	$p = 0.05$	CAPSS input masking (tested 0.05~0.2)
Station decay σ	30.0 km	Gaussian distance decay radius
Station buffer	10.0 km	Boundary stations outside domain
Total stations	322 (309 + 13 buffer)	Domain + boundary buffer
λ_{grid}	0.5	Station-grid consistency
λ_{emis}	1.0	Emission auxiliary; $w(e) = \min(1, e/5)$ warmup over first 5 epochs
Hardware	NVIDIA RTX 5090 (VRAM 32 GB), system RAM 128 GB	Single GPU; 2022-only training

Table S2. Deep learning model dual validation performance for all 7 pollutants.

Pollutant	In-sample R^2 (2022)	Temporal holdout R^2 (2023)	5-fold spatial holdout R^2 (2023)	Spatial holdout RMSE
PM _{2.5}	0.910	0.847	0.510	10.87 $\mu\text{g}/\text{m}^3$
PM ₁₀	0.916	0.903	0.389	20.87 $\mu\text{g}/\text{m}^3$
NO ₂	0.910	0.892	0.539	8.84 ppb
O ₃	0.944	0.926	0.732	10.55 ppb
CO	0.664	0.589	0.447	0.14 ppm
SO ₂	0.415	0.277	0.113	1.04 ppb
NO	0.883	0.842	0.590	9.93 ppb

R^2 , RMSE from grid readout at station-collocated 1 km grid cells
Train: 2022, Eval: 2023