

지자체 탄소중립계획 지원을 위한 지역 수준 탄소배출모델 개발 및 활용성 평가

정유정* · 송철호** · 조현우*** · 고영진**** · 이우균*****†

*고려대학교 환경생태공학과 석사과정학생, **고려대학교 오정리질리언스연구원 연구교수, ***고려대학교 오정리질리언스연구원 박사후연구원 / 국제응용시스템분석연구소 박사후연구원, ****고려대학교 환경생태공학과 박사과정학생, *****고려대학교 환경생태공학부 교수

Development and applicability assessment of local-scale carbon emission models to support carbon neutrality plans of local governments

Jeong, Yujeong* · Song, Cholho** · Jo, Hyeon-Woo*** · Ko, YoungJin**** and Lee, Woo-Kyun*****†

*M.S. Student, Dept. of Environmental Science and Ecological Engineering, Korea University, Seoul, Korea

**Research Professor, OJong Resilience Institute (OJERI), Korea University, Seoul, Korea

***Post-doc, OJong Resilience Institute (OJERI), Korea University, Seoul, Korea /

Postdoctoral Fellowship, Agriculture Forestry and Ecosystem Services (AFE) Group, Biodiversity and Natural Resources (BNR) Program, International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, A-2361 Laxenburg, Austria

****Ph.D. Student, Dept. of Environmental Science and Ecological Engineering, Korea University, Seoul, Korea

*****Professor, Division of Environmental Science and Ecological Engineering, Korea University, Seoul, Korea

ABSTRACT

Achieving carbon neutrality by 2050 necessitates significant contributions from local governments. To accelerate the creation and execution of local government carbon neutral plans, ongoing research focuses on spatial carbon emission data. However, in the Republic of Korea, policy-related use of spatial carbon emission data remains limited. This study aims to develop a spatial regression model for carbon emissions using machine learning-based ridge regression and regional greenhouse gas inventories to aid basic local governments in their decision-making regarding carbon neutrality and regional carbon mitigation. Input data were created by disaggregating subdivision land cover maps and facility-level national statistics to the local government level. Group K-fold cross-validation and area-based scaling were applied to improve generalization of the model. Two local carbon emission prediction models (the ‘LCE model’) were developed based on regional GHG inventory in Si·Do-level (the ‘Level-1 model’) and LCE model based on Regional GHG inventory in Si·Gun·Gu-level (the ‘Level-2 model’). The Level-2 model exhibited higher accuracy, with R² values of 0.84 and 0.66 at Si·Gun·Gu-level for 2019 and 2020, respectively, and at 0.93 and 0.76 at Si·Do-level. The carbon emission maps generated by the Level-2 model exhibited higher accuracy than the Level-1 model. This study highlights the cost effectiveness of machine learning-based spatial regression models for carbon emissions compared to IPCC (Intergovernmental Panel on Climate Change) methods and fuel-based models. The machine learning-based methodology and its detailed emission maps are expected to provide timely scientific evidence for developing and assessing carbon neutrality plans at Si·Gun·Gu-level, delivering granular information on carbon emissions down to the Dong·Ri-level

Key words : Carbon Emission Model, Carbon Emission Map, Carbon Neutrality Plan, Carbon Neutrality, Local Government, Ridge Regression, Machine Learning

†Corresponding author : leewk@korea.ac.kr (145 Anam-ro, Seongbuk-gu, Seoul, 02841, Korea. Tel. +82-2-3290-3470)

ORCID 정유정 0009-0001-4462-0266
송철호 0000-0002-8491-9545
조현우 0000-0001-6127-883X

고영진 0000-0002-9345-3214
이우균 0000-0002-2188-359X

1. 서론

2050년 국가 탄소중립(Carbon Neutrality) 달성을 위해 국가 및 지방정부는 온실가스 순배출량 '0'을 달성해야 한다. 이를 위한 탄소중립계획 수립 및 달성에 관한 논의가 국제 사회(Global)부터 지역(Local)에 이르기까지 다양한 공간적 수준에서 이뤄지고 있다(Rhoden et al., 2021; UNFCCC, 2018; Wu et al., 2022). 특히, 지방정부는 탄소중립 이행의 핵심 주체로 인정되고 있으며, 온실가스 감축에 관한 상향식 및 하향식 의사결정의 연결부이자 배출원에 대한 직·간접적 관리 권한을 지녔기에 이들의 온실가스 감축 활동이 국가 탄소중립 달성 여부의 중요한 요인으로 여겨진다(Castán Broto et al., 2022; Huovila et al., 2022; ICLEI East Asia, 2022; Li et al., 2021; Park et al., 2022; Rankl et al., 2023).

탄소중립 이행에 대한 지방정부의 기여를 확대하기 위해서는 지역 탄소중립계획 수립 및 실질적 이행이 선행되어야 하며, 계획 수립을 지원하기 위한 행정구역 단위 혹은 픽셀 수준의 탄소배출 공간자료(Carbon emission spatial product) 구축의 필요성이 강조되고 있다. 탄소배출 공간자료는 주요 배출원의 위치 및 배출량 핫스팟(Hot spot) 파악을 위한 정보를 제공하며, 탄소감축을 고려한 지역 토지이용계획의 기초자료로 활용될 수 있다는 점에서 필수적이다. 타 국가에서는 지역 혹은 보다 세밀한 수준으로 탄소배출지도가 구축되고 있다(Allam et al., 2022; Kang et al., 2021; Kim et al., 2024). 미국은 Greenhouse Gas Reporting Program (GHGRP)를 통해 북미 전역 이산화탄소 배출량을 시설 단위, 주(State) 단위, 유역 단위 등 다양한 공간 규모로 표출하는 지도시스템을 구축하였으며, 영국은 국가 인벤토리 자료의 배출량을 공간적으로 분할하여 1 km × 1 km 해상도의 인벤토리 배출 지도서비스를 제공하고 있다(EPA, 2022; Tsagatakis et al., 2023). 해당 시스템들을 통해 탄소배출량 모니터링 및 데이터베이스화는 가능하나 공간자료의 정책 활용 관련 논의는 아직 미흡한 실정이다(Shin et al., 2023).

현재 우리나라도 탄소중립 달성을 위해 기초지자체의 역할을 확대하고 공간 단위 탄소 정보 구축 및 탄소 관리를 위해 노력하고 있다. 지역 주도형 탄소감축을 실현하고자 「기후위기 대응을 위한 탄소중립·녹색성장 기본법(탄소중립기본법)」을 통해 광역 및 기초지자체의 탄소중립기본계획 수립을 의무화하고 탄소중립도시 지정 등과 같은 공간 계획과 관련된 탄소중립 이행을 법제화하였다.

탄소중립기본법 시행령 제39조는 온실가스 통계와 함께 공간 단위 온실가스 배출량 및 흡수량 정보 구축에 관한 법적 토대를 마련하였고 국토교통부는 2023년부터 온실가스 배출량·흡수량 정보를 탄소공간지도 시스템을 통해 배포하고 있다. 해당 시스템은 격자 단위로 배출량 공간 자료를 제공하고 있지만 전체 배출 부문(Sector) 중 건물 및 수송 부문 배출량만을 반영하기에 국가 온실가스 인벤토리에서 명시하는 주요 배출원에 대한 모든 정보를 아직 포괄하지 못하고 있으며 토지피복 정보가 반영되지 않았기에 지역 토지이용계획 수립과의 연계에 한계가 있다.

지역 탄소배출량을 공간적으로 예측하기 위한 연구 또한 활발히 진행되고 있다. 미국의 Vulcan 프로젝트와 Hestia 프로젝트는 대기 관측자료 및 모형을 적용하는 상향식 방법론과 구축된 온실가스 인벤토리를 프록시 자료(Proxy data)로 다운스케일링(Downscaling)하는 하향식 방법론을 적용하여 미국 전역 혹은 미국의 특정 주(State)의 탄소공간자료(Spatially-explicit Carbon Emission)를 활발히 구축하고 있다(Gurney et al., 2012; Gurney, Patarasuk et al., 2019; Huang et al., 2022; Khazini et al., 2022; Patarasuk et al., 2016; Roest et al., 2020). Li et al. (2017)과 Tsagatakis et al. (2023)은 사회경제적 통계 자료와 IPCC 온실가스 배출량 산정방법을 탄소배출량을 추정하는 다음, 토지피복 단위로 배출량을 분할하거나 시설 단위 배출량을 격자 단위로 집계하여 탄소배출량을 공간화했으며, Han et al. (2018)은 야간불빛 위성영상을 활용하여 배출량을 픽셀 단위로 변환했다. 유럽연합을 대상으로 IIASA-GAINS (Greenhouse gas Air pollution Interaction and Synergies)모형, STEAM (Ship Traffic Emissions Assessment Model) 모형, GFAS (Global Fire Assimilation System) 등 부문별 계수 기반 모형을 적용해 유럽의 온실가스 배출량을 모의하고 이를 인구밀도와 토지피복도를 통해 격자 단위로 배출량을 분해하여 지도화한 연구도 진행되었다(Kuenen et al., 2022). 다만, IPCC 산정방법론과 경제 및 연료 통계자료 기반의 탄소배출량 예측은 다량의 국가와 지역 통계자료, 부문별 배출계수 혹은 모형을 요구하기에 전 부문의 배출량 산정에 많은 자원이 투입되어야 한다. 또한, 배출량 맵핑(Mapping)을 위한 별도의 데이터 수집 및 분석 과정을 거쳐하는 번거로움이 존재한다. 이를 극복하기 위해 인공지능기법을 적용하여 보다 적은 데이터와 간략한 분석 과정으로 탄소배출량을 예측하는 연구도 진행되고 있으나(Mardani et al., 2020; Nie et al., 2023), 탄소배출량을 공간적으로 예측하지 않거나 국토

전역의 지역을 다루지 않으며, 무엇보다도 기존 국가 온실가스 인벤토리 및 지역 의사결정체계와의 일관성이 부족하고 모델 결과에 대한 직관적인 해석이 어렵기에 탄소중립 정책 활용에 여전히 한계가 존재한다. 결과에 대한 낮은 해석력을 극복하고 토지이용계획 정책과의 연계성을 높이고자 Luo et al. (2023)은 기계학습 기반의 선형 모델인 릿지 회귀모델(Ridge Regression)과 토지피복도에서 추출한 프록시 데이터를 활용하여 지방정부 수준에서의 현재 탄소배출량을 산출하고 토지피복 변화 시나리오를 구축하여 미래 탄소배출량까지 예측하는 연구를 진행하였다. 해당 연구에서 적용한 기계학습 기반 릿지 회귀모델은 선형모델이기에 입력변수와 배출량 간 인과관계 해석이 용이하여 정책 활용성이 높다는 장점이 있으며, 대리변수를 활용할 수 있기에 기존 계수 기반 모델에 비해 데이터 취득 및 구축 측면에서보다 비용효과적이다.

따라서, 지역 수준으로 세분화된 탄소배출량 분석 필요성은 존재하나, 국내에서 모든 주요 온실가스 배출원을 포괄하는 배출량을 지역 수준에서 예측하고 공간화하는 사례는 부족한 실정이라 할 수 있다. 본 연구는 이러한 문제를 해결하기 위하여 기계학습 기반 릿지 회귀모델과 지역 온실가스 인벤토리를 활용하여 지역 탄소배출량을 시공간적으로 예측하는 공간회귀모델을 개발하고 예측 결과를 적용하여 4단계 행정체계 단위 탄소배출지도도를 구축하고자 한다. 탄소배출지도도는 국내 탄소중립정책 의사결정체계 내 최소 공간 단위인 동리 수준부터 광역지자체 수준까지 구축하고 연구 결과에 대한 활용 및 개선 방안에 대하여 논의하고자 하였다.

2. 연구 재료 및 방법

2.1. 연구 대상지

본 연구는 우리나라 전체를 대상으로 수행하였다. 우리나라는 도/특별시/광역시, 시/군/구, 읍/면/동으로 구성된 3단계 행정체계 따르며, 읍/면의 하부 구역으로 리가 있다. 2022년을 기준으로 도/특별시/광역시는 총 17개, 시/군/구는 총 226개, 읍/면/동은 총 3,532개가 존재한다(MOIS, 2022). 읍면의 하위 행정구역인 리 단위는 전 국토를 포괄하지 못하기에, 본 연구에서는 행정체계상 의사결정의 최소 단위인 동 단위와 리 단위를 병합하여 4단계 행정체계(광역지자체, 기초지자체, 읍면동, 동리 단위)를 설정하였으며, 해당 체계에 따라 탄소배출지도도를 구축하였다(Fig. 1).

국가 온실가스 인벤토리에 따르면, 2021년도의 국가 총배출량은 656.2 MtCO₂ eq, 순배출량은 618.3 MtCO₂ eq이다. 2023년 국가 온실가스 인벤토리에 따르면, 국가 총배출량 중 약 86.8%가 화석연료 연소 과정에서 발생하며 이의 대부분이 공공전기 및 열 생산, 석유정제 등 에너지산업 혹은 철강, 비철금속, 화학 등과 같은 제조업으로부터 비롯된다. 뒤이어 산업공정 부문 배출량(7.4%), 농업 부문 배출량(3.2%), 폐기물 부문 배출량(2.5%) 순으로 국가 총배출량에 기여하는 바가 크다.

2.2. 연구 재료

연구재료는 크게 국가 공간자료와 통계자료가 있으며, 공간자료 중 토지피복도와 전력 생산 및 산업 관련 통계자료에서 배출원의 위치 정보 및 배출량을 추정할 수 있는 대리변수 값을 추출하였다. 생성된 정보들을 탄소배출활동지도 제작에 활용되었으며, 대리변수 값들을 기초지자체와 동리 단위로 집계하여 각각 모델의 학습을 위한 입력변수와 예측을 위한 출력변수로 활용하였다. 이때 GIS (Geographic Information System) 툴인 공간조인(Spatial Join)을 사용하였다.

모델의 입력자료로 활용한 논, 환경기초시설 등 면적 기반 배출원 정보는 환경부 세분류 토지피복지도를 통해 획득하였으며, 발전소, 산업체 등 점 기반 배출원 정보는 통계자료 및 주소 정보를 활용하여 구축하였다. 도로 수송, 철도 수송, 해운, 민간항공과 같은 이동연소 배출원의 정보는 세분류 토지피복지도를 통해 추출하였으며, 하수정화시설, 폐기물처리시설과 같이 활동량 정보 취득이 어려운 경우에도 토지피복도에서 얻은 시설 면적을 대리변수로 활용하였다. 모델 학습 및 검증에 위한 정답자료(Label data)로는 광역지자체 기준 및 기초지자체 기준 지역 온실가스 인벤토리에서 제공하는 직접배출량의 총배출량을 활용하였으며, 모델 입력 시 총배출량을 각 지자체의 산림지를 제외한 총면적으로 나눠 단위면적당 배출량으로 변환하였다. 연구 재료의 출처, 취득시기 및 사용 목적은 Table 1을 통해 확인할 수 있다.

2.3. 연구 방법

본 연구는 탄소배출활동지도 및 모델 입력자료 구축, 기계학습 기반 릿지 회귀모델 및 데이터 증강(Data augmentation) 기법을 적용한 동리 단위 배출량 추정, 행정구역 단위 탄소배출지도 제작 순으로 진행됐다. 탄소배출활동지도 구축을

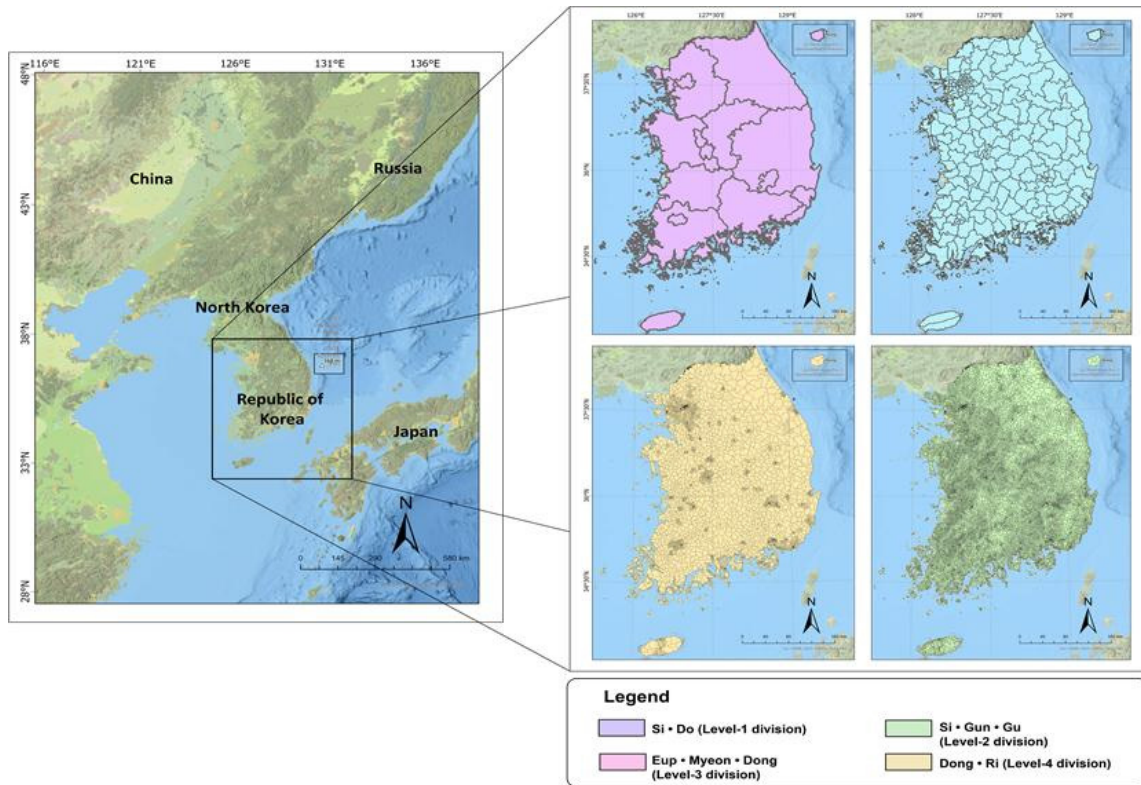


Fig. 1. Republic of Korea and its administrative boundaries in 4 different levels

Table 1. List of the data

	Name	Source	Initial information of the data	Variables produced from the data	Year	Unit / Resolution	Data format for spatialization	Use
Spatial Data	Administrative Boundary (4 levels)	MOLIT	-	-	2023	-	Polygon	Mapp-ing
	Subdivision Land Cover Map	ME	Area of 22 landcover types	Ratio of 7 landcover types to the area excluding forest area	2019 ~ 2020	1 m	Raster	Input data
National Statistics	Electricity Statistics	EPSIS	Yearly electricity production	Yearly energy production per unit area		MWh	Point	Input data
	Handbook for Community Energy Service	KDHCA	Yearly electricity and heat production			MWh	Point	Input data
	National Factory Registration Status	MOTIE	Area of factory site	Area of factory site by industry		m ²	Point	Input data
	Greenhouse Gas Inventory (Level-1 & Level-2 Local Governments)	GIR	Total greenhouse gas emission	Total greenhouse gas emission per unit area	CO ₂ t eq	Polygon	Label data	

Level-1: Administrative divisions of metropolitan cities and provinces (si · do)

Level-2: Administrative divisions of subdivisions in Level-1 (sig · gun · gu)

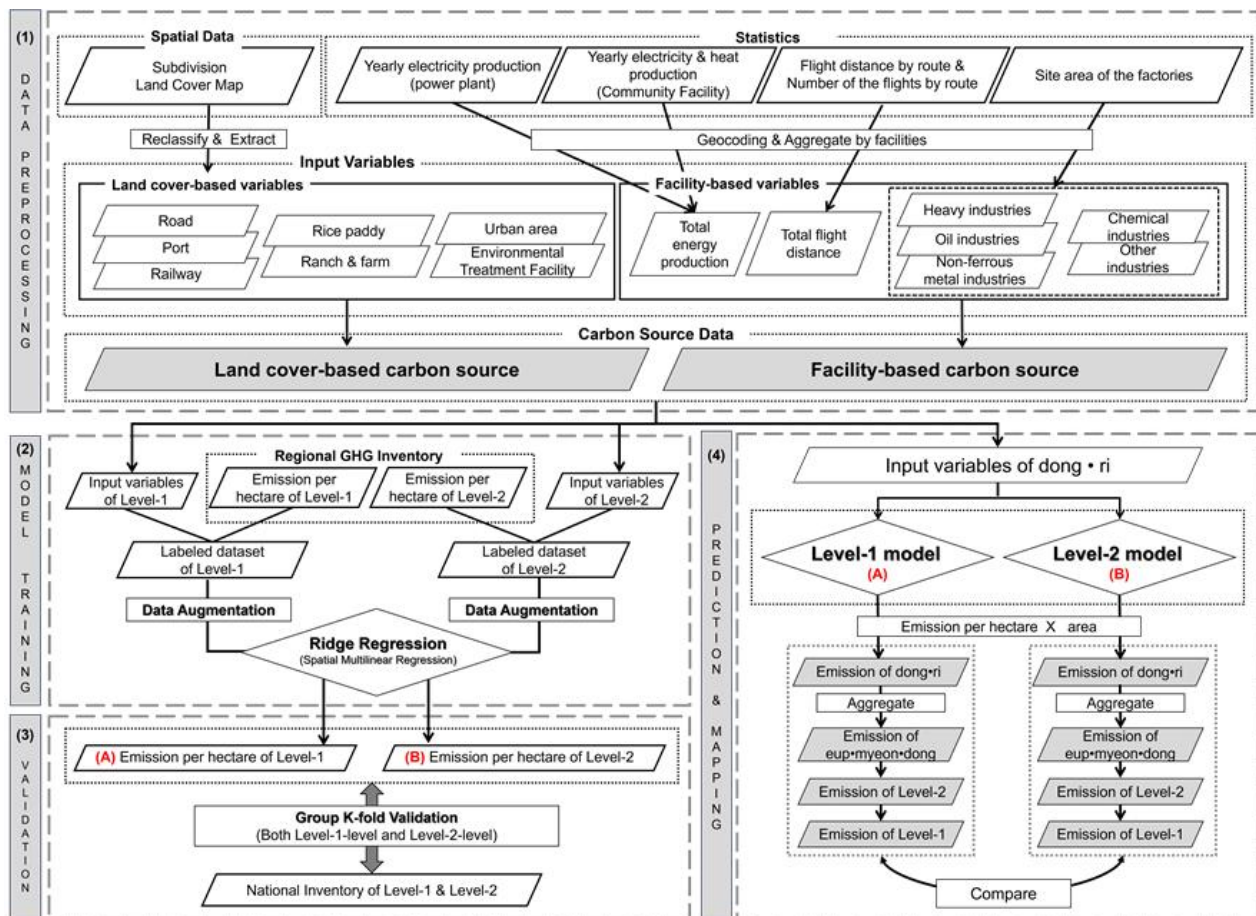
위해 탄소배출량을 설명할 수 있는 15개의 입력변수를 선정하였고 이를 격자 단위 혹은 포인트 형태로 맵핑(Mapping)하였다. 이와 같이 공간화된 변수들을 광역지자체와 기초지자체 및 동·리 단위로 추출 및 집계하여 3가지 행정구역 단위로 입력자료를 구축하였다. 이어 광역지자체와 기초지자체 단위 입력자료를 지역 온실가스 인벤토리 배출량과 결합하여 광역지자체와 기초지자체 수준으로 레이블 된 데이터셋(Labeled dataset)을 구성하였으며 이들 각각을 데이터 증강 기법을 통해 데이터의 크기를 확장했다. 동·리 단위 입력자료의 경우, 학습 및 검증을 거친 모델에 입력되어 동·리 수준으로 탄소배출량이 예측되는 데 활용되었다.

데이터셋의 정답자료로는 광역지자체 기준 온실가스 인벤토리와 기초지자체 기준 온실가스 인벤토리가 사용되었는데, 각각을 기계학습 기반 랫지 회귀모델에 적용하

여 광역지자체 인벤토리 기반 모델(이하 ‘Level-1 모델’)과 기초지자체 인벤토리 기반 모델(이하 ‘Level-2 모델’)을 구축하고 두 모델을 비교 분석하였다. 두 모델의 검증에는 Group K-fold 교차 검증을 적용했다. 이후 구축된 모델들에 최하위 행정구역 단위인 동리 단위 입력자료를 적용하여 동리 단위로 탄소배출량을 추정하고 해당 예측 결과를 상위 행정경계를 기준으로 디졸브(Dissolve)하여 총 4단계 행정구역 단위로 탄소배출지도를 제작하였다. 구축 시기는 연구자료들의 취득 가능 시기를 고려하여 2019년과 2020년으로 선정하였다.

2.3.1. 탄소배출활동지도 및 모델 입력자료 구축

탄소배출활동지도는 모델 입력자료로 활용한 변수들을 격자 단위 혹은 시설 단위로 지도화하여 제작하였다. 모



Level-1: Administrative divisions of metropolitan cities and provinces (Si · Do)

Level-2: Administrative divisions of subdivisions in Level-1 (Sig · Gun · Gu)

Fig. 2. Flow chart of the methodology

델의 입력자료는 15개의 변수로 구성되며 8개는 토지피복 기반 변수이고 7개는 통계 기반 변수로 구성되며, 변수들의 값은 배출원의 활동 규모를 간접적으로 대변할 수 있는 프록시 값이다. 해당 배출원별 변수의 값들을 격자 혹은 포인트 형태로 공간화하여 탄소배출지도를 구축하였고, 탄소배출활동지도에 GIS 기법인 공간조인과 디졸브(Dissolve)을 행정구역 경계 shape 파일을 기준으로 적용하여 광역지자체 단위, 기초지자체 단위, 동-리 단위 입력 자료를 제작하였다. 광역지자체와 기초지자체 단위 입력 자료를 광역지자체와 기초지자체 단위 온실가스 인벤토리와 결합하고 모든 변수를 단위면적당 값으로 변환하여 레이블링 자료를 구축하였다. 해당 입력자료들은 모델 학습과 검증에 사용했고 동-리 단위 입력자료는 단위면적당

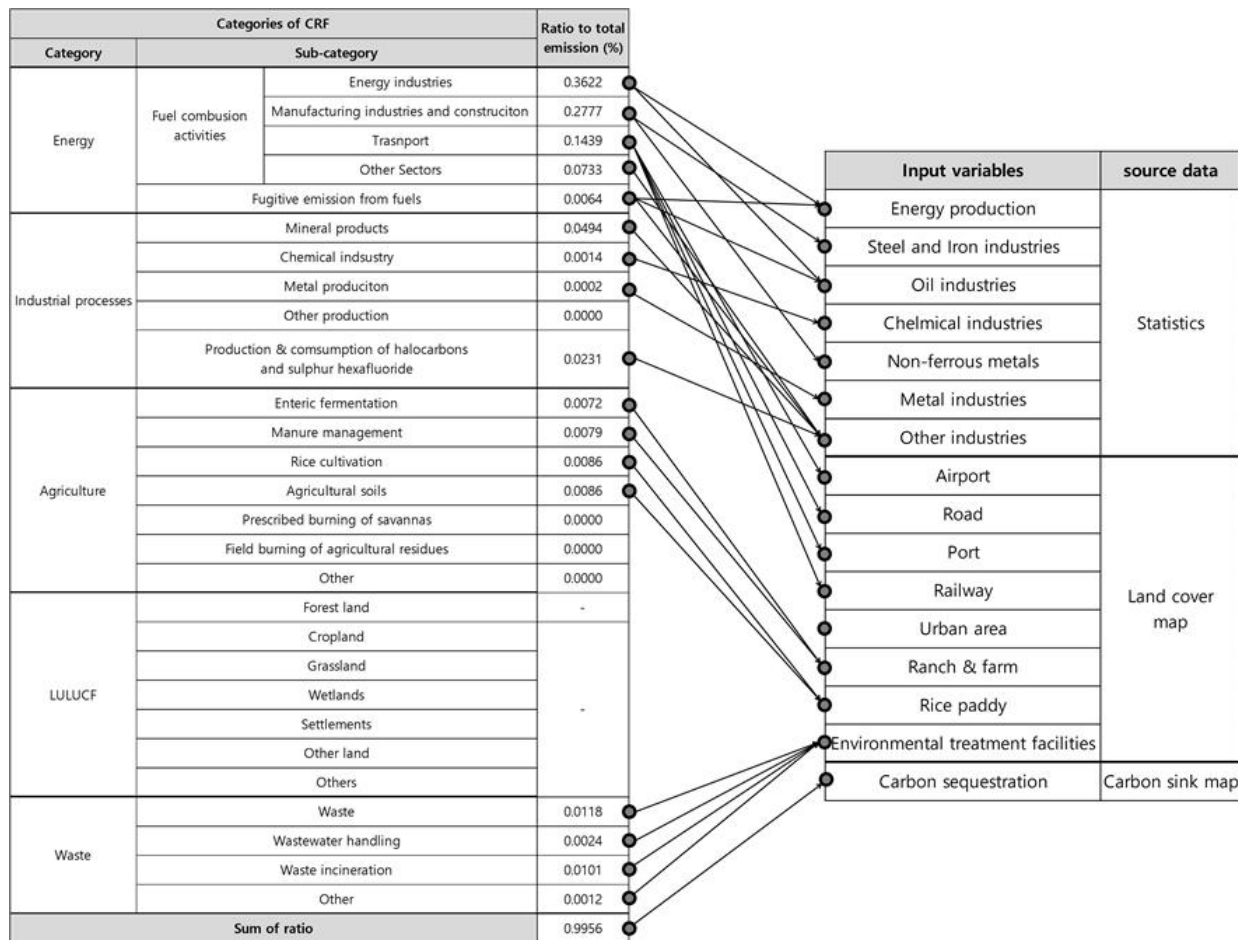
값으로 변환하여 동-리 단위 배출량 예측에 활용했다(Eq. (1)). 단위면적당 값은 산림지를 제외한 지자체의 전체 면적에 대한 항목별 면적 혹은 값의 비율이다.

$$Value\ per\ unit\ area_{ijk}(\%) = \frac{Area\ of\ Land\ Cover_{ijk}}{Total\ Area_{ij}} \quad Eq. (1)$$

i : administrative divisions

j (years): 2019, 2020

k (variables): area of urban, port, railway, road, environmental treatment facilities, rice paddy, ranch·farm, and factory sites, and total energy production



(a) CRF table of national GHG inventory

(b) Input variables

Fig. 3. Comparison between CRF (Common Reporting Format) table of National GHG inventory of ROK and input variables of the model

면적 기반 배출원과 이동연소 배출원의 위치정보 및 면적 값은 세분류 토지피복지도에서 추출하였으며 해당 배출원들은 레스터(Raster) 형식으로 탄소배출활동지도에 표출하였다. 22개의 피복 분류항목 중 국가 온실가스 인벤토리의 부문(Sector)을 참고하여 배출원으로 해석될 수 있는 15개 항목을 선별하였다. 이 중 ‘주거지역(단독주거시설, 공공주거시설)’, ‘상업지역’, ‘공공시설지역(교육·행정시설, 기타 공공시설)’을 ‘시가지’ 항목으로, ‘경지정리가 된 논’과 ‘경지정리가 안된 논’을 합쳐 ‘논’ 항목으로 병합하여 최종적으로 7개의 항목으로 재분류하였다. 탄소배출활동지도. 7개 항목 중 시가지, 논, 목장·양식장, 환경기초시설은 면적 기반 배출원이며 항만, 철도, 도로, 공항은 이동연소 배출원에 해당한다. 토지피복도 기반 변수들은 모델 내에서 연료 연소 내 가정·상업·공공 부문, 수송 부문, 폐기물 부문, 농업 부문에서 발생하는 배출량을 설명할 수 있다(Fig. 3).

시설 기반 배출원에는 발전소와 제조업의 생산시설이 해당하며 이들의 배출활동 규모에 대한 대리변수로는 발전소별 발전량(MWh)과 생산시설별 면적을 사용하였고 탄소배출활동지도에는 포인트 형태로 표출되었다. 제조업 생산 정보의 경우, 전자공시시스템의 감사보고서를 통해 얻을 수 있으나 해당 생산량은 생산시설 단위가 아닌 기업 단위로 보고되고 있고 상장된 기업의 정보만을 제공하기에 배출활동이 실제로 발생하지 않는 지역에 배출량이 할당되거나 상당량의 배출활동이 누락되는 문제가 발생한다. 따라서 산업통상자원부 전국공장등록현황의 생산시설별 제조시설 면적으로 생산량 정보를 대체하였다. 제조시설의 면적은 배출량 규모가 큰 업종을 기준으로 시설별 업종코드를 재분류하여 제조업을 크게 철강산업, 정유 산업, 비철금속 산업, 화학 산업, 비금속 산업, 기타 산업으로 세분화하였다. 발전소와 생산시설의 경우, 집단에너지사업편람 및 전국공장등록현황을 통해 발전량 및 주소 정보를 획득하고 주소가 명시되어 있지 않은 경우, 네이버 지도 서비스를 활용하였다. 수집된 주소 정보들은 공간화를 위하여 위경도 좌표로 변환되었다. 시설 기반 변수들은 모델 내에서 에너지 부문 내 연료 연소 부문과 산업공정 및 제품 생산 부문을 직간접적으로 설명할 수 있다(Fig. 3).

2.3.2. 릿지 회귀모델을 적용한 지자체 단위 탄소배출량 추정

본 연구에서는 기계학습 기반 릿지 회귀모델에 광역지자체 단위 및 기초지자체 단위 레이블링 자료를 적용하여

지역 탄소배출량 추정모델(이하 ‘LCE 모델’)을 구축하였다. 구축된 LCE 모델에 동리 단위 입력자료를 투입하여 동리 수준의 탄소배출량을 예측하고 그 값을 행정구역 경계를 적용하여 순차적으로 읍·면·동 단위, 기초지자체 단위, 광역지자체 단위로 병합하여 4단계 행정구역에 대한 탄소배출량을 산출하였다.

LCE 모델에 적용한 릿지 회귀모델은 RSS (Residual Sum of Squares)에 L2 규제항(L2 Regularization)을 추가한 것을 비용함수(Cost function)로 사용하는 다중선형회귀 모델이다(Eq. (2)). L2 규제는 모델이 이전 시점에 예측한 가중치 값의 제곱의 합에 비례하여 가중치의 크기를 축소할 수 있도록 설계되었기에 가중치의 크기가 커질수록 가중치 크기에 대한 억제 효과가 더 커지는 특성이 된다(Hoerl et al., 1985). 이러한 가중치 억제를 통해 독립변수 간 다중공선성으로 가중치 값이 과도하게 커지는 것을 예방하고 이상치(Outlier)에 대한 모델의 민감도를 낮추어 과적합(Overfitting) 문제를 방지하는 효과를 거둘 수 있다는 장점을 지닌다(Duzan and Binti Mohamad Shariff, 2015). 이러한 특성은 변수 간 다중공선성이 높고 양극단의 값이 다수 존재하는 희소행렬(Sparse matrix) 형태를 가진 입력 데이터로 발생할 수 있는 과적합(Overfitting)을 규제할 수 있다.

$$RSS_{ridge}(w, b) = \sum_{i=1}^n (y_i - (w_i x_i + \beta))^2 + \lambda \sum_{j=1}^p w_j^2 \quad \text{Eq. (2)}$$

RSS_{ridge} : cost function of ridge regression

w : coefficient, β : bias

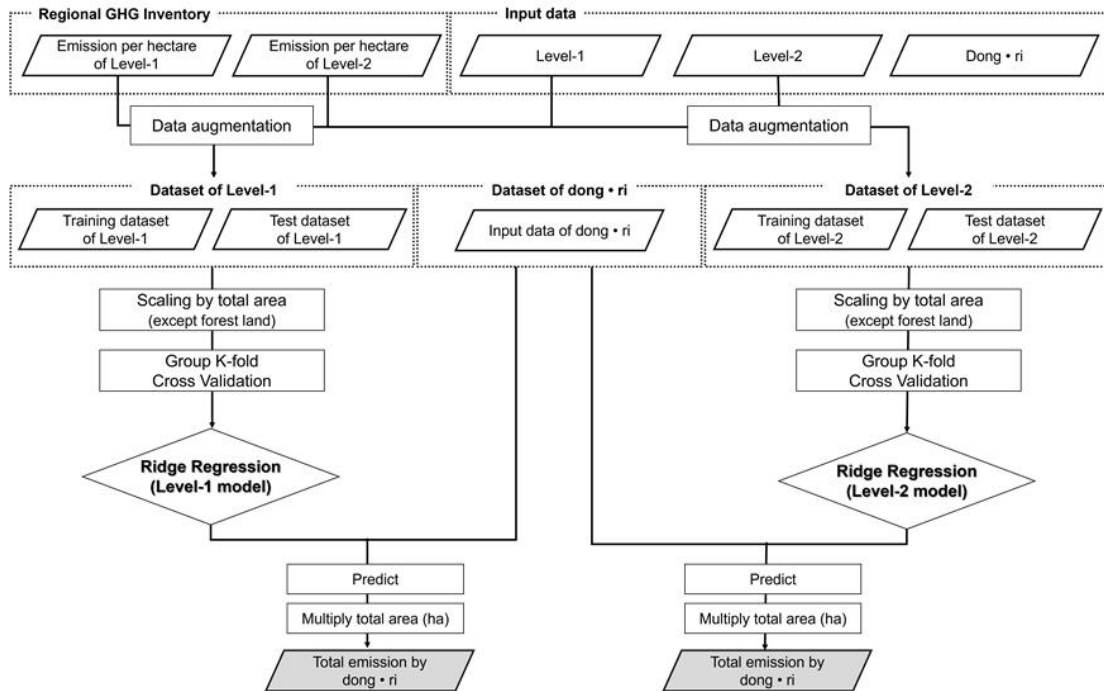
y_i : true values, x_i : predictor variables

λ : lambda (strength of regularization)

i : 1, 2, ..., n

j : 1, 2, ..., p

릿지 회귀모델에 광역지자체 단위 및 기초지자체 단위로 구축된 레이블링 데이터를 각각 투입하여 단위면적당 배출량(tCO₂ eq/ha)을 예측하는 Level-1 모델과 Level-2 모델을 구축하였다(Fig. 4). 학습 및 검증을 마친 Level-1 및 Level-2 모델에 구축된 동리 단위 입력자료를 투입하여 동리 단위 단위면적당 배출량(tCO₂ eq/ha)을 예측하고, 그 결과에 산림지를 제외한 지자체의 총면적(ha)을 곱하여 동리 단위 총배출량(tCO₂ eq)을 산출하였다. 동리 단



Level-1: Administrative divisions of metropolitan cities and provinces (Si · Do)

Level-2: Administrative divisions of subdivisions in Level-1 (Sig · Gun · Gu)

Fig. 4. Flow chart of model training, validation, and prediction using ridge regression

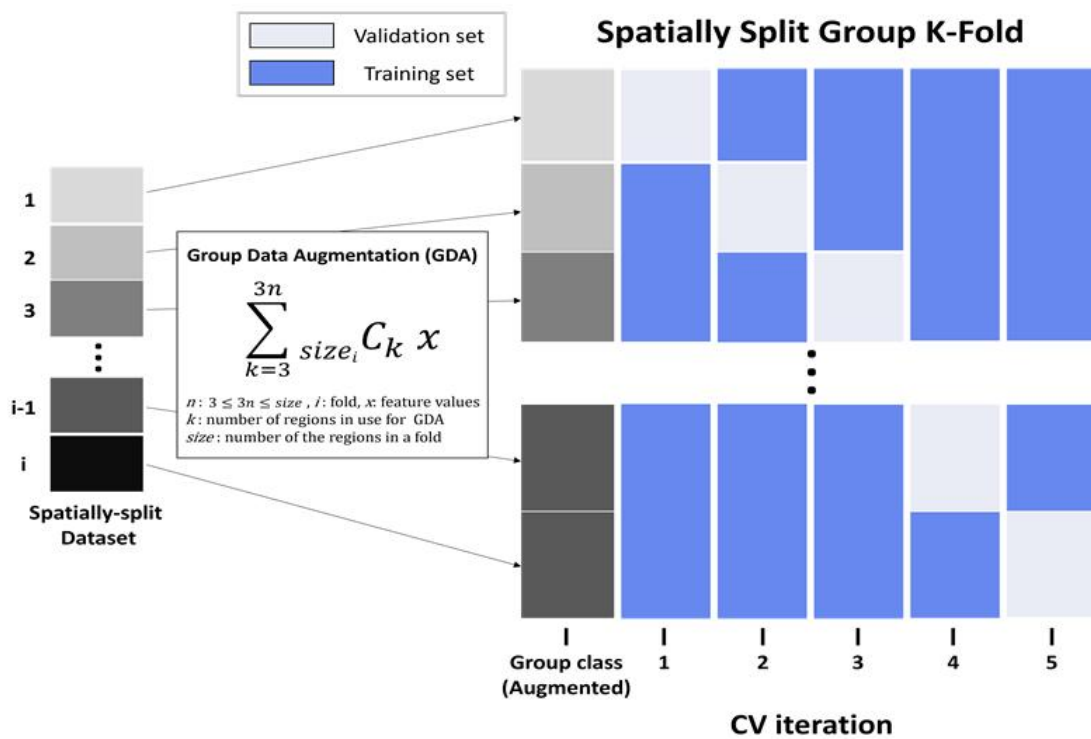


Fig. 5. Group data augmentation and group K-Fold cross validation

위 결과를 읍·면·동, 기초지자체, 광역지자체 단위 행정경계에 따라 공간적으로 병합하여 최종적으로 4단계 행정구역 단위별 지역 총배출량을 추정하였다.

광역지자체 인벤토리의 데이터 수는 2019년과 2020년도 합쳐 34개, 기초지자체는 데이터 수는 총 458개로 기계학습 모델의 학습과 검증 과정을 모두 수행하기에 충분하지 않다고 판단하였다. 따라서 조합(Combination)으로 데이터를 선택 및 합계하는 방식의 데이터 증강(Data augmentation) 알고리즘을 구축하여 광역지자체 단위 자료는 총 262,142개로, 기초지자체 배출량 자료는 총 393,168개로 그 수를 확장하였다. 또한, 과적합 방지 및 모델의 성능 향상을 위해 공간적으로 학습데이터셋(Training dataset)과 검증데이터셋(Validation dataset)을 분리하는 Group K-fold 교차검증(K-fold Cross validation)을 적용하여 모델 학습 및 검증 과정을 수행하였다(Fig. 5).

2.3.3. 탄소배출량 정확도 평가

모델 평가지표(Performance metrics)를 활용한 통계적 방법과 지도 및 순위 비교를 통한 정성적 방법으로 구축된 Level-1 모델과 Level-2 모델의 정확도를 평가하였다. 첫 번째로는 광역지자체 단위 온실가스 인벤토리 및 기초지자체 단위 온실가스 인벤토리의 총배출량과 두 모델이 예측한 결과 간 차이를 분석하여 평가지표 값을 산출하였다. 해당 평가 결과를 통해 모델의 예측 성능 및 국가 온실가스 인벤토리와의 정합성을 파악하였다. 사용된 정확도 지표로는 RMSE (Root Mean Square Error), MAE (Mean Absolute Error), R^2 (Coefficient of determination)가 있다 (Eq. (3)). 두 번째로는 광역지자체 및 기초지자체 인벤토리와 예측된 총배출량을 지도화하여 탄소배출량의 공간적 분포를 정성적으로 비교분석했다. 또한, 기초지자체 인벤토리의 총배출량 및 모델 예측 결과를 기준으로 상위 10위 및 하위 10위 지역을 선정 및 비교하여 고배출 지역과 저배출 지역에 대한 공간적 패턴 유사성을 평가하였다.

$$R^2 = 1 - \left(\frac{\sum (y_i - \bar{y})}{\sum (y_i - \hat{y}_i)} \right)^2 \quad \text{Eq. (3)}$$

R^2 : coefficient of determination

y_i : observation, \bar{y} : mean, \hat{y}_i : prediction

3. 연구 결과

3.1. 탄소배출활동지도

LCE 모델에 활용한 입력자료를 지도화하여 Fig. 6과 같이 탄소배출활동지도(Carbon emission activity map)를 구축하였다. 시가지와 도로는 우리나라 전체 인구의 약 44%가 거주하는 서울특별시와 광역시를 중심으로 집중되어 있고 높은 고도가 낮은 서부 지역에 널리 분포되어 있음을 확인할 수 있다(Fig. 6a). 시설 기반 배출원은 위치 정보와 함께 배출활동 규모에 관한 정보를 저장할 수 있는 포인트 형태로 지도화하였으며 그 크기가 클수록 배출활동량이 더 많음을 의미한다. 집단에너지 발전시설은 인구밀도가 높은 특별시와 광역시를 중심으로 분포하며 발전소, 철강 산업 및 정유 산업 시설은 대체로 화석연료의 수입 및 제품 수출에 유리한 서해안과 남해안을 따라 위치하였다(Fig. 6b,c). 정유 산업 시설과 철강 산업 시설은 소수 지역에 집중적으로 분포하는 반면, 그에 비해 화학 산업 시설은 전역적으로 분포하는 경향이 있었다(Fig. 6(c)). 공항의 경우, 탄소배출활동지도에는 토지피복지도의 공항 대신 공항의 좌표를 이용하여 포인트 형태로 나타냈으며 배출활동의 크기 추정을 위해 국내 및 국외 운항 횟수와 운항 거리를 활용할 수 있었다(Fig. 6(d)). 다만, 지역 온실가스 인벤토리에는 현재 항공 수송을 산정할 때 순항배출량 없이 이착륙 시 배출량만을 반영하기에 모델의 입력변수로는 토지피복지도 상 공항 면적을 활용하는 것이 타당한 것으로 파악되었다(Shin et al., 2017).

3.2. 지자체 단위 탄소배출량 예측 및 탄소배출지도 구축

3.2.1. 모델 검증 및 입력변수별 기여도 분석

Group K-Fold 교차검증으로 MC 모델을 검증한 결과, MAE는 1306.76, RMSE는 21.12, R^2 는 0.47이었으며 훈련 데이터셋의 R^2 는 0.50로 MC 모델의 탄소배출량에 대한 설명력은 중간 수준 정도이며 과적합은 발생하지 않은 것으로 판단하였다. Level-2 모델의 경우, MAE는 50.97, RMSE는 4773.32, R^2 는 0.93이며 훈련데이터의 R^2 도 약 0.96으로 높은 예측 성능을 보였다. Level-1 모델의 경우, 값이 과소추정되어 지역 온실가스 인벤토리의 총배출량과의 비교를 통하여 보정 계수(Correction Factor)를 산출하고 그 값은 2.3이었으며, 이를 적용한 값을 Level-1 모델의 최종 예측값으로 활용하였다. Level-1 모델과 Level-2 모

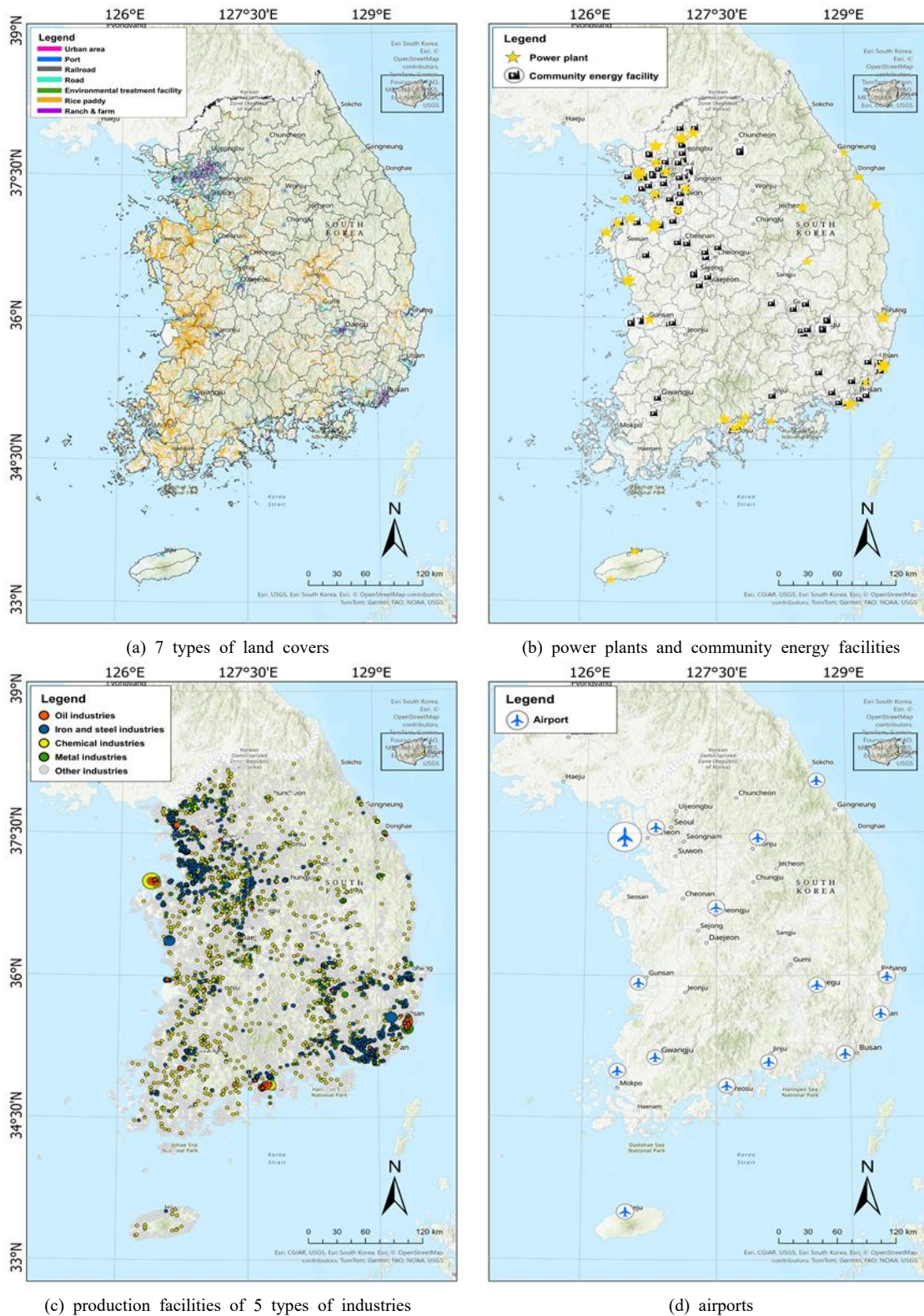


Fig. 6. Carbon emission activity maps

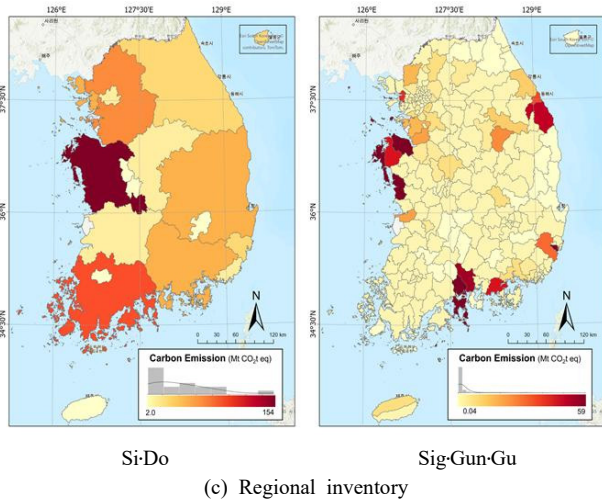
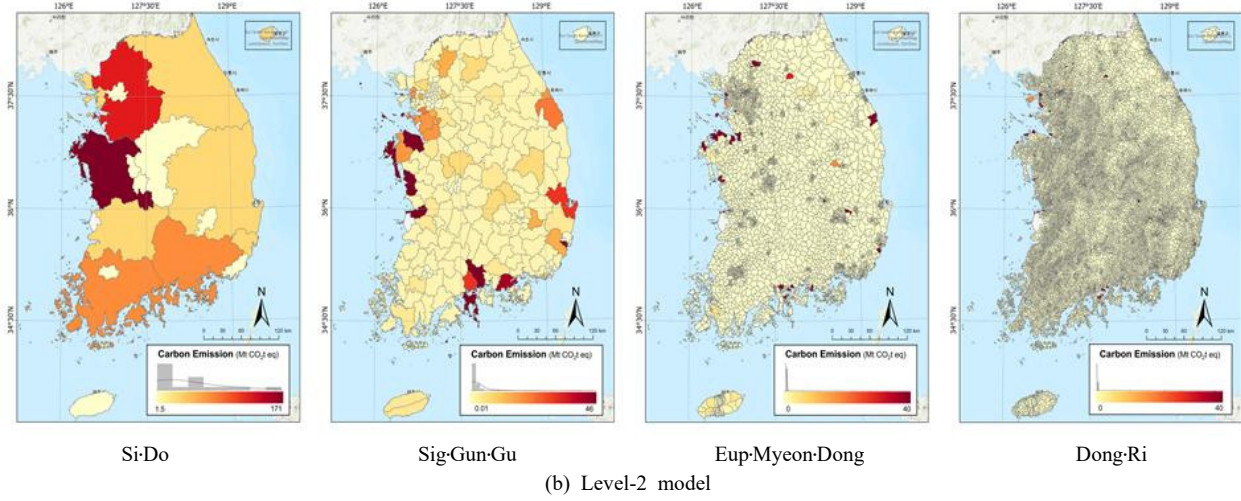
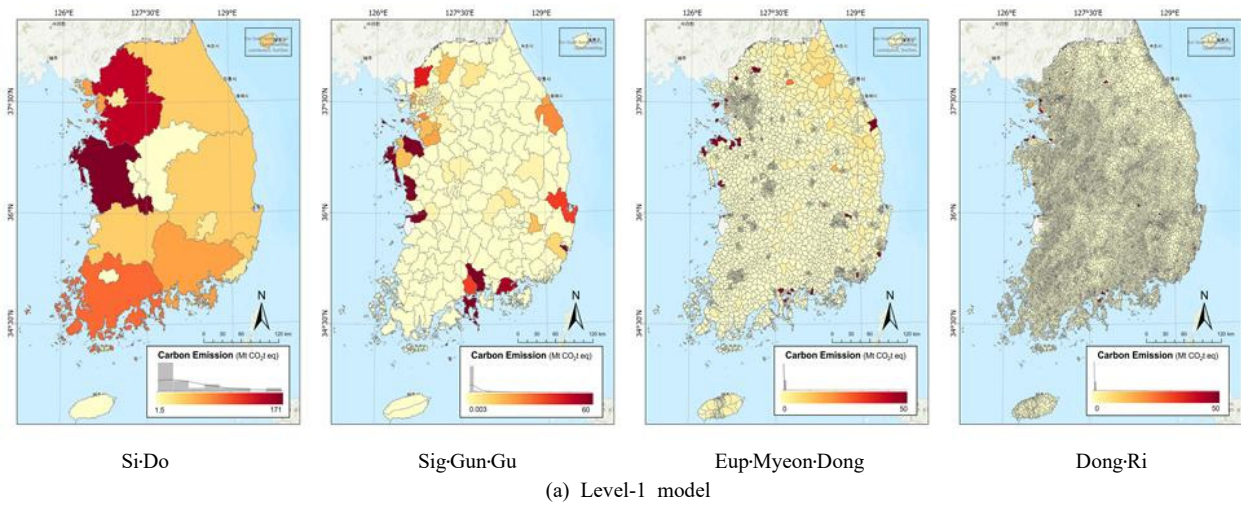


Fig. 7. Carbon emission maps by 4 different administrative levels in 2019

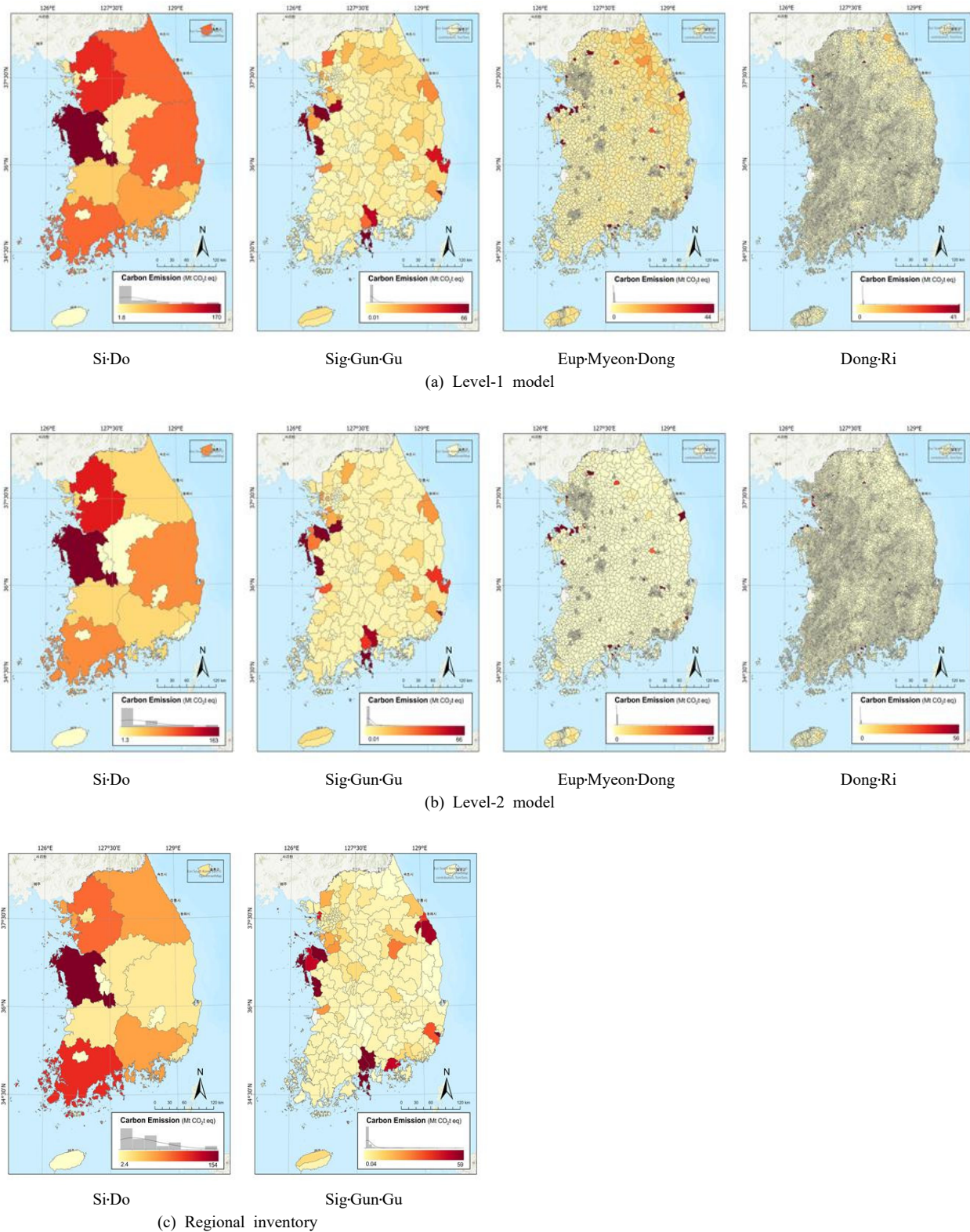


Fig. 8. Carbon emission maps by 4 different administrative levels in 2020

델의 검증결과를 비교하였을 때 광역지자체 단위 온실가스 인벤토리보다 기초지자체 단위 온실가스 인벤토리를 활용하는 것이 탄소배출량 예측 측면에서 유리함을 확인하였다. 이는 기초지자체 인벤토리 내 지역의 수가 더 많기에 보다 다양한 지역별 배출 특성 정보를 모델 내 반영할 수 있었기 때문이다.

Table 2는 Level-1 모델과 Level-2 모델의 입력변수별 가중치와 절편값을 보여준다. Level-1 모델의 가중치 크기가 전체적으로 Level-2 모델의 크기보다 상대적으로 크게 산출됨을 알 수 있는데, Level-2 모델의 학습데이터로 사용된 지역의 산림지를 제외한 최소 면적 평균 19,293 ha이고 Level-1 모델은 평균 240,130 ha로 Level-1 모델이 추정하는 지역의 규모가 상대적으로 크기 때문이다.

Level-1 모델과 Level-2 모델 모두 릿지 회귀 기반 선형 회귀분석을 적용하였기에, 가중치의 크기가 클수록 해당 변수의 배출량에 대한 기여도 및 단위면적당 배출 강도가 높다고 해석할 수 있다. 따라서, 두 모델 내 입력변수 간 크기를 비교하였을 때, 두 모델 모두 발전소의 에너지 발전량의 배출량에 대한 기여도가 높았으며, 그 뒤로 Level-1 모델은 기타 산업, 철강업, 금속산업, 정유업, 목장·양식장 순으로 높았고, Level-2 모델에서는 도로, 시가지, 기타 산업군, 정유업, 항만 순으로 높았다. 두 모델의 가중치 패턴이 이처럼 다른 이유는 도로 및 시가지와 산업시설 사이의 높은 상관성으로 인해 발생한 다중공선성이 Level-1 모델에서 더 강하게 작용했기 때문이다. 따라서 Level-1 모델의 입력자료에는 도로 및 시가지와 산업시설이 혼재된 지역이 많았기에 두 변수 간 다중공선성이 산업시설과 관련된 입력변수에 가중치가 집중하는 효과가 발생했다. 반면, Level-2 모델의 입력자료는 비교적 공간적 규모가 작아짐에 따라 도로 및 시가지와 산업시설이 분리되면서 상대적으로 다중공선성이 감소하였고 산업시설로 집중되었던 가중치가 도시 및 시가지로 분산된 것으로 분석되었다.

Level-1 모델에서 가중치가 0으로 해석된 변수로는 환경기초시설, 논, 화학산업, 비철금속 산업, 공항면적이 있으며 Level-2 모델에서는 목장·양식장, 공항면적이 있다. 릿지 회귀는 규제항을 통해 다중공선성이 높거나 모델의 설명력을 낮추는 변수의 가중치에 0에 가까운 값을 부여한다. 따라서, 이처럼 가중치가 0인 변수는 해당 배출원에서 실제로 배출이 발생함에도 불구하고 논과 목장·양식장, 정유산업과 화학산업처럼 변수 간 다중공선성이 높거나 공항면적과 같이 배출량에 미치는 영향력이 대부분의 지역에서 매우 작거나 없음을 시사하였다.

3.2.2. 탄소배출지도 구축

Level-1 모델과 Level-2 모델로 예측된 탄소배출량을 적용하여 2019년과 2020년 탄소배출지도를 구축하였다 (Fig. 7, Fig. 8). Level-1 모델과 Level-2 모델의 광역지자체 단위 탄소배출지도는 국가 인벤토리의 광역지자체 단위 탄소배출지도와 비교하였을 때, 배출량의 공간적 패턴이 높은 수준으로 유사함을 확인할 수 있다. 기초지자체 단위 지도 또한, Level-1 모델과 Level-2 모델 모두 기초지자체 단위 인벤토리로 구축한 지도와 공간적 패턴이 상당히 일치하였고 세 지도 모두 탄소배출량이 1000 MtCO₂ eq 이상인 고배출 지역은 대부분 서해안과 남해안을 따라 분포함을 확인하였다. 반면, 1000 MtCO₂ eq 이하인 저배출 지역의 경우, Level-1 모델보다 Level-2 모델로 구축한 지도에서 배출량의 공간적 분포가 국가 인벤토리 지도와 더 유사한 것으로 보인다. 읍·면·동 단위 지도를 비교하였을 때는 2019년과 2020년 모두 Level-1 모델의 지도에서 강원도 북부 지역의 색이 진하지만 Level-2 모델 지도는 그렇지 않음을 알 수 있는데, 이는 토지피복 기반 배출원이 지배적인 지역에서는 모델의 결과가 지역 면적이 커짐에 따라 배출량이 커지는 경향이 있으며 강원도 북부의 지역의 면적이 타지역보다 그 면적이 넓어 배출량이 높게 추정된 것으로 사료된다. 따라서, 기초지자체 자료를 적용한 Level-2 모델의 탄소배출지도가 우리나라 지역 탄소배출량의 공간적 패턴을 Level-1 모델보다 잘 모사하며 예측 결과의 면적 크기에 대한 종속성이 비교적 낮음을 알 수 있다.

광역지자체 단위 이하의 지도부터는 배출량이 극히 일부 지역에 집중되어 있으며 행정체계 수준이 낮아질수록 고배출 지역(붉은색)과 저배출 지역(옅은 노란색) 간 배출량 격차가 점점 커졌으며, 이러한 경향은 Level-1 모델과 Level-2 모델에서 모두 나타났다. 특히, Level-1 모델은 산업시설과 관련된 변수에 가중치가 집중되어 기초지자체 단위 배출량의 최솟값은 Level-2 모델보다 작고 최댓값은 더 큼을 확인하였는데, 이러한 배출량 분포의 차이가 Level-1 모델의 지도에서 산업시설이 위치한 지역과 산업시설이 위치하지 않은 지역 간 배출량 격차가 더 두드러지게 하였다. 특히, Level-1 모델에서는 토지피복 기반 배출원이 우세한 지역의 배출량을 Level-2 모델보다 더 과소추정했기에, Level-2 모델의 지도가 기초지자체 인벤토리의 지도에 비해 내륙 지역이 전반적으로 더 옅은 색을 띠고 있음을 알 수 있었다.

Table 4. Coefficients and intercepts of the Level-1 and Level-2 model

Variable		Model	Level-1 model	Level-2 model
Urban area			0.00004034	0.00007561
Port			0.00000606	0.00000507
Railroad			0.00000248	0.00000170
Road			0.00006731	0.00014843
Environmental treatment facility			0.00000000	0.00000020
Rice paddy			0.00000000	0.00000031
Ranch & farm			0.00012016	0.00000000
Airport			0.00000000	0.00000000
Energy production			0.58215733	1.0018809
Industrial Area	Iron and steel		0.00812046	0.00000050
	Oil refining		0.00482160	0.00000823
	Chemical		0.00000000	0.00000474
	Non-ferrous metals		0.00000000	0.00000048
	Metal		0.00549156	0.00000099
	Others		0.37629577	0.00002736
Intercept			5.2086931	49.259828

Table 5. MAE, RMSE, and R2 of Level-1 and Level-2 model

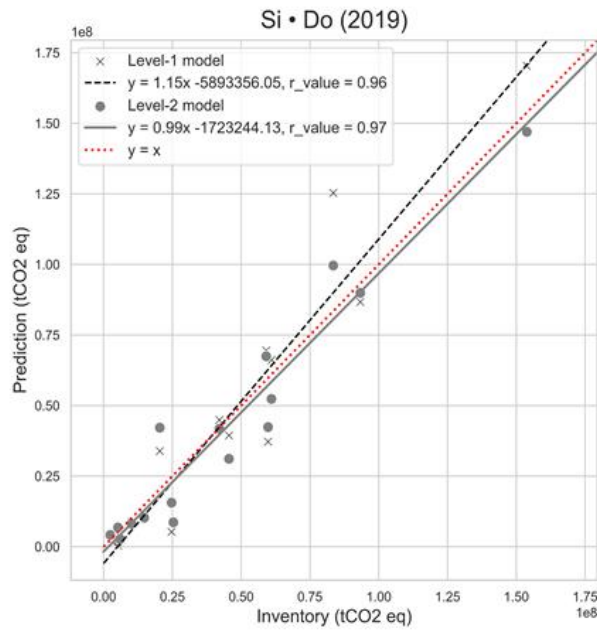
Model	Test dataset	year	MAE	RMSE	R ²
Level-1 LCE model	Level-1 Regional GHG inventory	2019	10,523,878.46	14,647,314.37	0.86
		2020	16,503,420.96	22,439,346.27	0.62
		Total	13,513,649.71	18,543,330.32	0.74
	Level-2 Regional GHG inventory	2019	1,690,099.73	3,702,103.01	0.75
		2020	2,121,275.37	5,345,091.41	0.37
		Total	1,905,687.55	4,523,597.21	0.56
Level-2 LCE model	Level-1 Regional GHG inventory	2019	8,195,088.16	10,455,690.08	0.93
		2020	11,940,595.15	18,010,304.72	0.76
		Total	10,067,841.66	14,232,997.40	0.84
	Level-2 Regional GHG inventory	2019	1,568,853.28	2,975,327.91	0.84
		2020	1,845,196.38	3,952,726.57	0.66
		Total	1,707,024.83	3,464,027.24	0.75

Level-1: Administrative divisions of metropolitan cities and provinces (si · do)

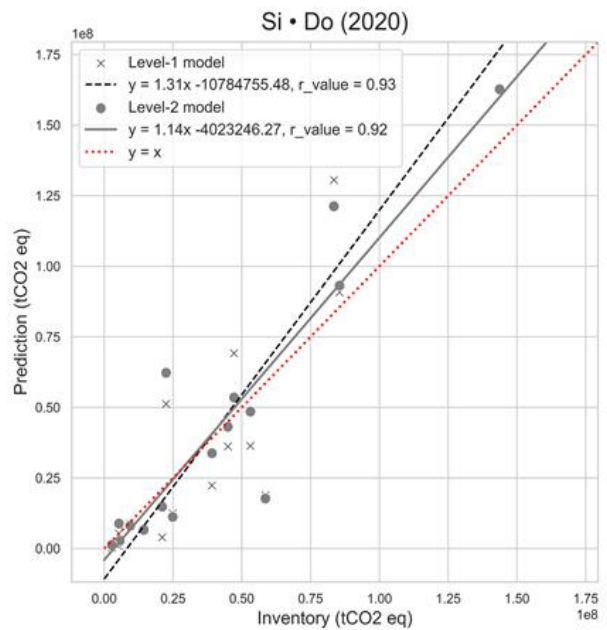
Level-2: Administrative divisions of subdivisions in Level-1 (sig · gun · gu)

Level-1 model: LCE model based on Level-1 National GHG inventory

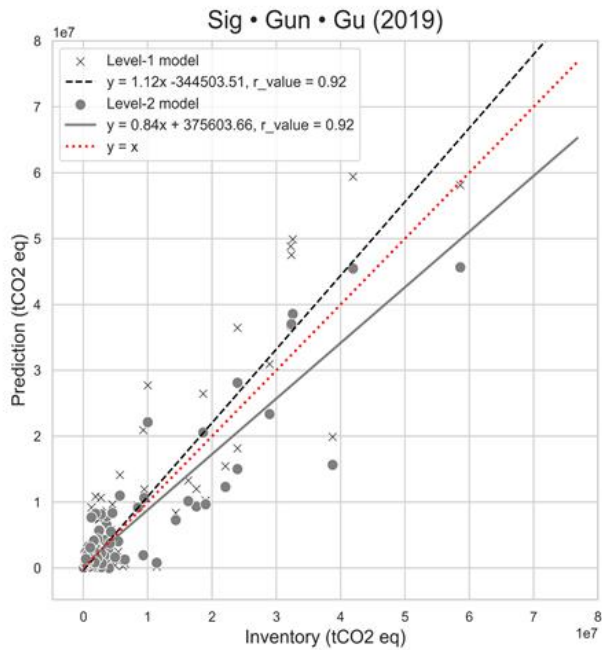
Level-2 model: LCE model based on Level-2 National GHG inventory



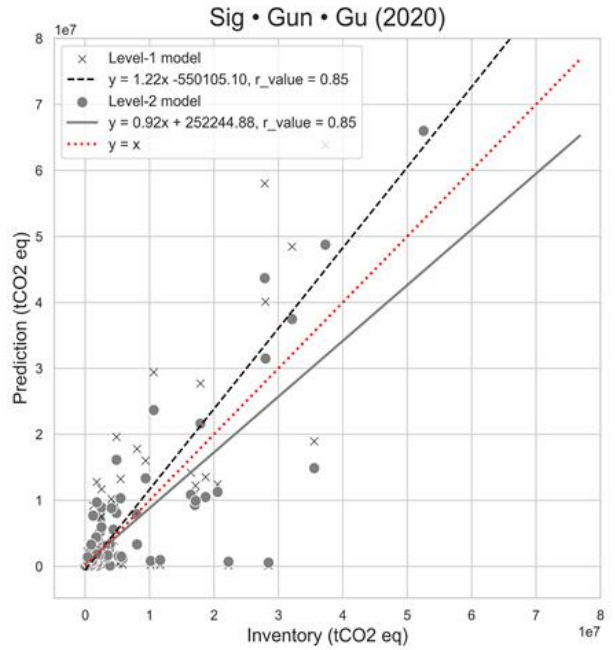
(a) Si · Do-level in 2019



(b) Si · Do-level in 2020



(c) Sig · Gun · Gu-level in 2019



(d) Sig · Gun · Gu-level in 2019

※ 'r-value' on the graph is Pearson correlation coefficient between prediction and inventory, different from the 'R²' used as model's performance metric

Fig. 9. Correlation between carbon emission of national inventory and model prediction (both Level-1 and Level-2 model)

Table 6. Emissions in the level of Si · Do of the Regional GHG inventory, Level-1 model and Level-2 model in 2019

Unit: tCO₂ eq

Name	Level-2 regional GHG inventory	Level-1 model		Level-2 model	
		Predicted value	Residual	Predicted value	Residual
Seoul	25,853,156.61 (3.78%)	8,978,431.89 (1.25%)	-16,389,714.32	8,644,198.78 (1.29%)	-16,723,947.43
Busan	14,736,669.8 (2.15%)	11,453,428.50 (1.60%)	-3,283,241.31	10,333,370.08 (1.54%)	-4,403,299.73
Deagu	10,132,888.88 (1.48%)	8,775,605.75 (1.22%)	-1,357,283.12	8,213,397.22 (1.22%)	-1,919,491.66
Incheon	60,978,851.89 (8.91%)	66,164,761.27 (9.22%)	+5,185,909.38	52,419,505.22 (7.79%)	-8,559,346.67
Gwangju	5,252,960.67 (0.77%)	427,726.55 (0.06%)	-4,825,234.12	1,642,148.42 (0.24%)	-3,610,812.25
Deajeon	5,962,672.20 (0.87%)	2,237,987.05 (0.31%)	-3,724,685.15	2,759,309.07 (0.41%)	-3,203,363.13
Ulsan	45,537,012.44 (6.65%)	39,444,895.68 (5.50%)	-6,092,116.76	31,198,891.98 (4.64%)	-14,338,120.46
Sejong	2,397,493.59 (0.35%)	4,201,781.49 (0.59%)	+1,804,287.90	4,208,926.75 (0.63%)	+1,811,433.16
Gyeonggi	79,468,570.15 (11.61%)	125,347,147.82 (17.47%)	+41,905,518.65	99,666,535.36 (14.82%)	+16,224,906.19
Gangwon	59,730,504.46 (8.73%)	37,210,750.75 (5.19%)	-22,519,753.71	42,424,293.84 (6.31%)	-17,306,210.62
Chungbuk	24,663,728.79 (3.60%)	5,286,156.93 (0.74%)	-19,377,571.86	15,631,590.08 (2.32%)	-9,032,138.71
Chungnam	153,801,771.80 (22.47%)	170,371,067.58 (23.74%)	+16,569,295.77	147,023,007.50 (21.86%)	-6,778,764.30
Jeonbuk	20,410,151.23 (2.98%)	33,915,242.38 (4.73%)	+13,505,091.15	42,211,277.55 (6.28%)	+21,801,126.33
Jeonnam	93,280,112.06 (13.63%)	86,737,385.26 (12.09%)	-6,542,726.80	89,988,593.03 (13.38%)	-3,291,519.03
Gyeongbuk	18,111,723.39 (2.65%)	44,933,592.55 (6.26%)	+2,873,822.96	41,836,374.24 (6.22%)	-223,395.36
Gyeongnam	59,093,293.54 (8.63%)	69,507,469.80 (9.69%)	+10,414,176.25	67,523,529.12 (10.04%)	+8,430,235.58
Jeju	5,173,074.78 (0.76%)	2,637,570.13 (0.37%)	-2,535,504.65	6,831,462.91 (1.02%)	+1,658,388.13
Total	712,020,731.12 (100%)	717,631,001.39 (100%)	+5,610,270.27	672,556,411.17 (100%)	-39,464,319.95

Level-1: Administrative divisions of metropolitan cities and provinces (Si · Do)

Level-1 model: LCE model based on Level-1 National GHG inventory

Level-2 model: LCE model based on Level-2 National GHG inventory

3.3. 탄소배출량 정확도 평가

3.3.1. 지역 온실가스 인벤토리와의 통계적 검증

Level-1 모델과 Level-2 모델로 예측한 동리 단위 배출량을 기초지자체 온실가스 인벤토리의 총배출량과 비교하여 평가지표를 계산하였으며, 두 모델의 시계열 예측 성능을 파악하고자 2019년과 2020년의 결과를 각각 산출하였다. Level-1 모델은 광역지자체 수준 R^2 가 2019년도, 2020년도 각각 0.86, 0.62, 기초지자체 수준에서는 각각 0.75, 0.37이고, Level-2 모델의 광역지자체 수준 R^2 는 2019년도와 2020년도 각각 0.93, 0.76이고 기초지자체 수준은 각각 0.84, 0.66으로 계산되었다(Table 4). Level-2 모델의 2019년도와 2020년도 정확도가 모두 Level-1 모델에 비해 높았으므로 탄소배출량의 시공간적 예측 측면에서 Level-2 모델이 Level-1 모델보다 적합함을 알 수 있었다. 다만, 두 모델 모두 2020년도에 대한 예측력이 2019년도 대비 낮으며 광역지자체 수준보다 기초지자체 수준에서 비교적 낮은 정확도를 보였기에, Level-1 모델과 Level-2 모델 모두 2019년도에 과적합 된 경향이 있으며 공간해상도가 높아짐에 따라 정확도가 낮아짐을 확인하였다.

모델의 정확도는 지역의 위치와 주요 배출원에 따라서도 상이하였다. Level-1 모델의 예측 결과와 기초지자체 단위 인벤토리의 총배출량을 광역지자체 단위로 비교하였을 때, 경기도, 충청남도, 전라북도, 경상남도, 인천광역시, 경상북도, 세종특별시는 과대추정했고 그 외 광역지자체는 과소추정했다. Level-2 모델의 경우, 전라북도, 경기도, 경상남도, 세종특별시, 제주특별시를 과대추정했으며 나머지 지역은 과소추정했다(Table 5). 광역지자체별로 예측 경향이 상이한 원인으로는 교차검증으로 인해 모델의 가중치가 학습데이터로 사용된 지역의 특성에 편중되어 최적화되었을 경향과 선형회귀식이기에 배출원 간 비선형적 상호작용과 지역에 따라 상이한 배출 특성을 충분히 반영하지 못하였음이 있다고 사료된다. 주요 배출원에 따른 정확도 차이를 보기 위해 에너지 발전소 혹은 산업시설인 위치한 지역(189개)과 그 외 지역(39개)로 나누어 R^2 을 계산한 결과, 각각 0.87, 0.24로 두 지역 간 정확도도 크게 차이가 남을 발견하였다. 예측 결과와 인벤토리 값 사이의 상관관계를 그린 그래프에서도 배출량이 커질수록 회귀선 근처에 값들이 분포하는 반면, 그 값이 작아질수록 분산이 커짐을 알 수 있다(Fig. 9). 이는 발전소 및 대규모 산업시설이 분포한 지자체는 배출 요인이 특정 배출원에 집중되어 있어 그 패턴이 비교적 단순하지만, 온

실가스 배출원이 다양하고 배출 특성이 복잡한 지자체는 외부 영향 및 자연적 요인에 의한 배출량의 변동성이 크기에 정확한 예측이 더 어려운 것으로 분석된다(Kim, 2022). 특히, 지역 탄소배출량이 토지피복 기반 배출원에 의존적인 도시일수록 값의 잔차가 크게 나타났다.

3.3.2. 배출량의 공간적 분포 비교 분석

탄소배출지도를 통한 배출량의 공간적 분포 비교는 3.2.2에서 수행하였으며, 총배출량을 기준으로 상위 10개 기초지자체와 하위 10개 기초지자체를 나열하여 예측 탄소배출량과 기초지자체 온실가스 인벤토리 총배출량 간 공간적 분포를 추가적으로 비교하였다(Table 6). 상위 지역의 경우, 순위에 있어 다소 차이가 존재하였으나 순서와 상관없이 Level-1 모델은 7개의 지역이, Level-2 모델은 8개의 지역이 기초지자체 온실가스 인벤토리의 상위 10개 지역과 일치하였다. 하위 지역의 경우, Level-1 모델과 Level-2 모델 모두 기초지자체 인벤토리와 일치하는 지역이 존재하지 않았다. 이는 탄소배출량의 공간적 분포 예측에 있어 시설 기반 배출원이 집중된 고배출지에서는 높은 정확도를 보였지만 토지피복 기반 배출원이 우세한 저배출지에서의 정확도가 낮음을 보여준다. 특히, 기초지자체 온실가스 인벤토리와 달리 모델 예측값의 하위권 지역에는 모두 서울특별시, 부산광역시, 인천광역시, 대구광역시에 위치하였다. 특별시와 광역시에는 토지피복 기반 배출원이 대부분의 배출량을 차지하고 있어 지자체 면적에 따라 탄소배출량이 크게 변동하며, 그 면적의 크기도(道) 내 지역보다 작아 Level-1 모델과 Level-2 모델에서 모두 특별시와 광역시의 탄소배출량을 실제보다 과소 추정하는 경향이 뚜렷하게 나타났다.

4. 고찰

본 연구는 지역 맞춤형 탄소중립계획 수립 및 이행 지원을 위한 우리나라의 지역 수준 탄소배출 공간 정보를 보완하고자 기계학습 기반 릿지 회귀모델 모델로 광역지자체부터 동리 단위까지의 지역 탄소배출량을 예측하였다. 구축한 LCE 모델에는 Level-1 모델과 Level-2 모델이 있으며, 특히, Level-2 모델은 광역지자체 수준과 기초지자체 수준에서 2019년 및 2020년도 모두 R^2 가 0.65 이상으로 높은 수준의 예측 성능을 보였다. 또한, 기초지자체 온실가스 인벤토리 값을 적용한 탄소배출지도와의 비교

에서도 탄소배출량의 공간적 분포가 유사함을 정성적 및 정량적으로 확인하였다. 따라서, Level-2 모델로 구축한 탄소배출지도는 기초지자체 내 탄소배출량의 분포와 지역별 주요 배출원에 대한 정보를 제공하며 국가 온실가스 인벤토리라도 높은 정합성을 보였다. 게다가 탄소배출모델에 기계학습 및 대리변수를 적용하였기에 부문별 배출계수 및 활동자료 기반 산정 방법에 비해 데이터 수집에 투입되는 자원을 절감하고 배출량 추정 프로세스를 자동화할 수 있었다. 이는 기존의 통계자료 기반 방법론보다 비용효과적인 방식으로 총량적 개념에서의 온실가스 배출량의 공간적 추정이 가능함을 시사한다(Bhatt et al., 2023; Luo et al., 2023; Zhong et al., 2021).

탄소배출활동지도의 경우, 시설 기반과 토지 기반 배출원 정보를 장소 기반(Location-based)으로 제공하기에 이를 동·리 단위 탄소배출지도와 중첩하여 Scope 1 접근법을 기반으로 각 지자체의 배출원 구성 및 특성에 부합하는 탄소 감축 전략을 세우는 것을 지원할 수 있다. 그 예로, 시설 기반 배출원은 산업 부문 중심 및 국가 혹은 민간 주도적 관리 대상이기에 해당 배출원이 우점하는 지자체는 국가와의 협력으로 지자체의 탄소 감축 역량을 강화할 수 있으며, 비산업 부문이자 지자체 관리 부문인 토지 피복 기반 배출원이 우점하는 지자체는 지자체가 직접 토지이용계획을 수립하고 관련 탄소 감축 사업을 발굴 및 실행할 수 있다. 더 나아가 향후 탄소배출지도를 인구, 지역산업연관표와 같은 사회·경제적 자료, 시설별 관리정보와 결합하여 지역 간 생산 및 소비의 흐름과 시설의 관리주체를 파악할 수 있다면, Scope 2 혹은 더 나아가 Scope 3 접근법으로 확장하여 지자체 탄소배출량을 파악 및 관리할 수 있을 것으로 기대된다.

LCE 모델의 예측값으로 구축한 4단계 행정체계 탄소배출지도는 국가 의사결정체계 내 가장 최소 단위인 동리 단위부터 광역지자체 단위까지의 탄소배출량 정보를 공간적으로 제공한다. 이러한 공간자료를 통해 배출량에 대한 핫스팟(Hotspot)과 콜드스팟(Coldspot)을 발견하고 여러 공간해상도에서 탄소배출량을 다층적으로 모니터링할 수 있다(Olawade et al., 2024). 탄소배출지도의 최소 해상도인 동리 단위에 해당하는 지역의 평균 면적은 5.35 km²으로 현재 구축된 탄소배출량 공간자료의 해상도가 100 m²에서 250 km² 사이인 것을 고려하였을 때, 탄소배출량 추적을 위해 충분히 정밀한 해상도인 것으로 사료된다(Gurney, Liang et al., 2019; Gurney, Patarasuk et al., 2019; MOLIT, 2023; Patarasuk et al., 2016; Roest et al.,

2020). 더불어 동리는 일종의 지역 커뮤니티로서 사회·경제적 및 자연적 환경의 유사성이 높은 지역들의 집합체이며 기초지자체와 민간 부문의 매개체 역할을 수행하고 있기에, 동리 지역의 배출량 및 배출 특성에 대한 정보는 향후 지자체 탄소중립계획의 실효성 및 활용성 증진과 지역 토지이용계획과의 연계에 기여할 수 있다(Kallushi et al., 2012; Lee, 2023; Seto et al., 2021). 더 나아가 탄소배출지도를 현재 국내에서 개발 중인 산림생장식 기반 탄소흡수 모델로 제작한 격자 단위 탄소흡수지도와 연계한다면, 전국 탄소수지도 구축이 가능하며 이를 탄소중립마을 지정, 탄소 배출·흡수 특성을 반영한 탄소중립전략 수립과 같은 지자체 탄소저감전략의 이행을 촉진시킬 수 있는 과학적 자료로 활용할 수 있을 것이다(Kim, 2022; Shin et al., 2017).

다만, 현재 Level-2 모델의 가중치가 일부 지역과 시기에만 최적화되어 있으며 토지피복 기반 배출원에서의 배출량 예측에서 낮은 성능을 보인다. 그 원인에는 대표성이 낮거나 불완전한 대리변수의 적용, 온실가스 인벤토리 내 불확도, 일부 탄소배출원의 누락 등이 있다(Gurney et al., 2017; Hogue et al., 2016). 이러한 한계점을 보완하기 위해 향후 데이터셋의 시계열을 확장하거나 비선형 회귀 모델 혹은 지역별 모형을 구축할 필요가 있다. 부문별 배출활동량 추정에 있어 피복 및 생산 시설 면적보다 대표성 있는 대리변수를 적용한다면 기계학습을 통한 예측 정확도를 높이고 더불어 보다 상세한 온실가스 배출활동 현황을 파악할 수 있을 것이다. 또한, 현재 탄소배출지도는 격자 단위가 아니기에, 유역 단위 등 다양한 공간 규모 데이터를 유연하게 활용하는 데 한계가 존재한다. 이는 향후 원격탐사 자료 등과 같은 격자 단위 자료를 통해 온실가스 공간자료를 구축하고 통계자료의 픽셀 단위 지도화 등을 통해 탄소배출 공간자료의 활용성을 향상하는 방향이 필요함을 의미한다.

또한, 현재 탄소중립 달성의 필요성과 온실가스 배출원에 대한 시공간적 분석의 필요성이 함께 증가하고 있다. 이런 측면에서 모든 배출원에 대한 시공간 정보 확보가 가능하고 이에 기반하여 정확한 배출량 수치가 지역 내에서 산정될 수 있는 것이 바람직하다. 그러나, 온실가스 배출원 추적의 시급성과 이에 대한 공간정보가 부재한 상황에서 본 연구에서 제시하는 방식을 통해 시의적절한 온실가스 배출량의 공간적 분석이 가능해지고, 지역 내의 탄소중립 계획 수립과 지원이 더 쉽게 이루어질 수 있을 것으로 생각한다. 따라서 추후 모형의 고도화와 더불어 지

역 온실가스 배출량에 대한 지속적인 시공간적 추적 및 관리를 할 수 있는 지역 주도형 체계를 마련해야 할 것이다(Bistline, 2021; Feng et al., 2016; Fong et al., 2014; Gurney, Liang et al., 2019; Gurney, Patarasuk et al., 2019; Kongboon et al., 2022).

5. 결론

본 연구에서는 지역 탄소중립계획 및 탄소감축전략 수립 지원을 위한 공간데이터를 구축하고자 기계학습 기반 탄소배출량 예측 모델 및 4단계 행정체제 단위 탄소배출 지도를 구축하였다. 지역 온실가스 인벤토리, 토지피복지도, 시설 기반 통계자료를 활용하여 모델을 학습 및 검증하였으며, 광역지자체 및 기초지자체 온실가스 인벤토리로 각각 Level-1 모델과 Level-2 모델을 구축 및 비교 분석하였다. 모델의 일반화 성능을 높이기 위하여 데이터 증강 기법과 Group K-Fold 교차 검증 기법을 적용하였다. 또한, 입력자료를 면적으로 스케일링함으로써 동리 수준까지 탄소배출량을 추정할 수 있는 공간회귀모델을 설계할 수 있었다. 최종적으로 동리 단위 입력자료를 모델에 적용하여 예측된 탄소배출량을 지도화하여 광역지자체, 기초지자체, 읍면동, 동리 단위의 4단계 행정구역별 탄소배출지도를 도출하였다. 그 결과, Level-1 모델에 비해 Level-2 모델의 정확도가 전반적으로 높았으며, Level-2 모델의 기초지자체 수준 R^2 는 2019년, 2020년도 각각 0.84, 0.66이었으며 광역지자체 수준은 각각 0.93, 0.76이었다. 또한, Level-2 모델 결과를 반영한 탄소배출지도가 Level-1 모델에 비해 기초지자체 온실가스 인벤토리와 보다 유사한 공간적 패턴을 보였으며, 특히 산업 기반 배출원이 존재하는 지역에 대한 배출량 예측이 토지 기반 배출원이 우점하는 지역 보다 높은 성능을 보였다. 다만, 해당 결과는 연료 기반의 산업 활동량 정보가 아닌 제조업 시설 면적이라는 대리변수를 사용하였기에 전체 배출량에서 각 산업이 기여하는 정도를 분석하기 위해서는 향후 실제 산업별 혹은 시설별 활동량과의 추가 검증 절차가 필요할 것으로 사료된다.

Level-2 모델과 탄소배출지도를 통해 기계학습 기반 릿지 회귀모델이 IPCC 온실가스 산정식 혹은 배출계수 기반 모형보다 비용효과적인 방식으로 탄소배출량을 공간화할 수 있음을 증명하였다. 이는 향후 보다 활용도 높은 기계학습 기반 탄소배출량 예측 모델 개발을 위한 지속적인 연구가 필요함을 시사한다. 또한, 본 연구에서 구축한

탄소배출지도는 최소 의사결정 단위인 동리부터 국가 전체까지 탄소배출 특성을 시공간적으로 분석할 수 있기에 지역 탄소중립계획 수립 및 평가를 지원할 수 있을 뿐만 아니라 하향식(Top-down) 및 상향식(Bottom-up) 정책 수립 과정의 유기적인 연계에 기여할 수 있을 것으로 기대된다. 또한, 온실가스 인벤토리 내 반영된 주요 배출원과의 연결성 및 토지피복정보를 고려하였기에 기초지자체 탄소중립계획 및 토지이용계획과의 연계성을 보다 높일 수 있었다. 따라서, 향후 기초지자체와의 협력을 통해 탄소배출지도의 읍면동리별 배출 특성뿐만 아니라 사회경제적 여건도 함께 반영한다면, 보다 실질적인 지역 맞춤형 탄소감축전략을 수립 및 실행을 위한 과학적 기초자료가 생산될 수 있을 것이다. 다만, 현재 모델이 일부 지역과 2019년에 편향되어 최적화되었기에, 특정 지역의 탄소중립계획에 탄소배출지도를 활용할 경우, 해당 지역의 통계자료 및 공간자료를 추가적으로 취득하여 모델을 보완 및 검증을 해야할 것으로 사료된다.

사사

본 연구는 2022년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원(NRF-2021K1A3A1A78097879)과 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업 자율운영형 중점연구소 고려대학교 오정리질리언스연구원(NRF-2021R16A1A10045235)의 지원을 받아 수행되었습니다.

References

- Allam Z, Bibri SE, Chabaud D, Moreno C. 2022. The '15-minute city' concept can shape a net-zero urban future. *Humanit Soc Sci Commun* 9(1): 126. doi: 10.1057/s41599-022-01145-0
- Bhatt H, Davawala M, Joshi T, Shah M, Unmarkat A. 2023. Forecasting and mitigation of global environmental carbon dioxide emission using machine learning techniques. *Clean Chem Eng* 5: 100095. doi: 10.1016/j.clce.2023.100095
- Bistline JET. 2021. Roadmaps to net-zero emissions systems: Emerging insights and modeling challenges. *Joule* 5(10): 2551-2563. doi: 10.1016/j.joule.2021.09.012

- Castán Broto V, Westman L, Ruiz Campillo X. 2022. The role of local governments in international climate policy. *Oxford research encyclopedia of climate science*. doi: 10.1093/acrefore/9780190228620.013.842
- Duzan H, Binti Mohamad Shariff NS. 2015. Ridge regression for solving the multicollinearity problem: Review of methods and models. *J Appl Sci* 15(3): 392-404. doi: 10.3923/jas.2015.392.404
- EPA (United States Environmental Protection Agency). 2022. 2022 greenhouse gas emissions from large facilities. [accessed 2024 Mar 5]. <https://ghgdata.epa.gov/ghgp/main.do>
- Feng S, Lauvaux T, Newman S, Rao P, Ahmadov R, Deng A, Díaz-Isaac LI, Duren RM, Fischer ML, Gerbig C, et al. 2016. Los Angeles megacity: A high-resolution land-atmosphere modelling system for urban CO₂ emissions. *Atmos Chem Phys* 16(14): 9019-9045. doi: 10.5194/acp-16-9019-2016
- Fong WK, Sotos M, Doust M, Schultz S, Marques A, Deng-Beck C. 2014. *Global protocol for community-scale greenhouse gas emission inventories: An accounting and reporting standard for cities*. Washington, DC: World Resources Institute.
- Gurney KR, Liang J, O’Keeffe D, Patarasuk R, Hutchins M, Huang J, Rao P, Song Y. 2019. Comparison of global downscaled versus bottom-up fossil fuel CO₂ emissions at the urban scale in four U.S. urban areas. *J Geophys Res Atmos* 124(5): 2823-2840. doi: 10.1029/2018JD028859
- Gurney KR, Liang J, Patarasuk R, O’Keeffe D, Huang J, Hutchins M, Lauvaux T, Turnbull JC, Shepson PB. 2017. Reconciling the differences between a bottom-up and inverse-estimated FFCO₂ emissions estimate in a large US urban area. *Elem Sci Anth* 5: 44. doi: 10.1525/elementa.137
- Gurney KR, Patarasuk R, Liang J, Song Y, O’keeffe D, Rao P, Whetstone JR, Duren RM, Eldering A, Miller C. 2019. The Hestia fossil fuel CO₂ emissions data product for the Los Angeles megacity (Hestia-LA). *Earth Syst Sci Data* 11(3): 1309-1335. doi: 10.5194/essd-11-1309-2019
- Gurney KR, Razlivanov I, Song Y, Zhou Y, Benes B, Abdul-Massih M. 2012. Quantification of fossil fuel CO₂ emissions on the building/street scale for a large U.S. city. *Environ Sci Technol* 46(21): 12194-12202. doi: 10.1021/es3011282
- Han J, Meng X, Liang H, Cao Z, Dong L, Huang C. 2018. An improved nightlight-based method for modeling urban CO₂ emissions. *Environ Model Softw* 107: 307-320. doi: 10.1016/j.envsoft.2018.05.008
- Hoerl AE, Kennard RW, Hoerl RW. 1985. Practical use of ridge regression: A challenge met. *Appl Statist* 34(2): 114-120. doi: 10.2307/2347363
- Hogue S, Marland E, Andres RJ, Marland G, Woodard D. 2016. Uncertainty in gridded CO₂ emissions estimates. *Earth’s Future* 4(5): 225-239. doi: 10.1002/2015EF000343
- Huang C, Zhuang Q, Meng X, Zhu P, Han J, Huang L. 2022. A fine spatial resolution modeling of urban carbon emissions: A case study of Shanghai, China. *Sci Rep* 12(1): 9255. doi: 10.1038/s41598-022-13487-5
- Huovila A, Siikavirta H, Antuña Rozado C, Rökman J, Tuominen P, Paiho S, Hedman Å, Ylén P. 2022. Carbon-neutral cities: Critical review of theory and practice. *J Clean Prod*. 341. doi: 10.1016/j.jclepro.2022.130912.
- ICLEI (International Council for Local Environmental Initiatives) East Asia. 2022. 2022 annual report. Seoul, Korea: Author.
- Kallushi A, Harris J, Miller J, Johnston M, Ream A. 2012. Think bigger: Net-zero communities. *Proc ACEEE* 11: 115-127.
- Kang HS, Park JS, Im ES, Ahn YH, Jo MS, Yoon EJ, Lee YJ, Park JH, Hong NE. 2021. Greenhouse gas mitigation strategy in the unit of local government and in spatial unit. Sejong, Korea: Korea Research Institute for Human Settlements. KRIHS Issue Report No. 51.
- Khazini L, Kalajahi MJ, Rashidi Y, Ghomi SMMM. 2022. Real-world and bottom-up methodology for emission inventory development and scenario design in medium-sized cities. *J Environ Sci* 127: 114-132. doi: 10.1016/

- j.jes.2022.02.035
- Kim JS. 2022. Engagement of local governments into carbon neutrality and their action plans. *Climate Change and Carbon Neutrality* 23: 26-35.
- Kim TH, Lee JY, Song SK, Song JK, Kim G, Kim T, Kwon Y, Hong S, Neubert M, Wolfgang W. 2024. Integrated management plan for local government land-environment planning considering carbon neutrality. *KEI Res Briefs* 12(1): 3-9.
- Kongboon R, Gheewala SH, Sampattagul S. 2022. Greenhouse gas emissions inventory data acquisition and analytics for low carbon cities. *J Clean Prod* 343: 130711. doi: 10.1016/j.jclepro.2022.130711
- Kuenen J, Dellaert S, Visschedijk A, Jalkanen JP, Super I, van der Gon HD. 2022. CAMS-REG-v4: A state-of-the-art high-resolution European emission inventory for air quality modelling. *Earth Syst Sci Data* 14(2): 491-515. doi: 10.5194/essd-14-491-2022
- Lee WK. 2023. Climate smart forestry in the era of carbon neutrality. Seongnam, Korea: The Korean Academy of Science and Technology.
- Li Y, Chen M, Montero D, Abraham B, Jin Z, Akagi J, Kang S. 2021. Comparative study on low carbon city development in China, Japan, and the Republic of Korea. Incheon, Korea: UNESCAP-ENEA (United Nations Economic and Social Commission for Asia and the Pacific East and North-East Asia Office). NEASPEC North-East Asia Low Carbon City Platform.
- Li Y, Qian X, Zhang L, Dong L. 2017. Exploring spatial explicit greenhouse gas inventories: Location-based accounting approach and implications in Japan. *J Clean Prod* 167: 702-712. doi: 10.1016/j.jclepro.2017.08.219
- Luo H, Li Y, Gao X, Meng X, Yang X, Yan J. 2023. Carbon emission prediction model of prefecture-level administrative region: A land-use-based case study of Xi'an city, China. *Appl Energy* 348: 121488. doi: 10.1016/j.apenergy.2023.121488
- Mardani A, Liao H, Nilashi M, Alrasheedi M, Cavallaro F. 2020. A multi-stage method to predict carbon dioxide emissions using dimensionality reduction, clustering, and machine learning techniques. *J Clean Prod* 275: 122942. doi: 10.1016/j.jclepro.2020.122942
- MOIS (Ministry of the Interior and Safety). 2022. Administrative division and population of local governments.
- MOLIT (Ministry of Land, Infrastructure and Transport). 2023. Emission/sequestration estimation methodology. Carbon Spatial Map System. [accessed 2024 Mar 4]. <https://www.carbonmap.kr/methodology.do>
- Nie W, Ao O, Duan H. 2023. A novel grey prediction model with a feedforward neural network based on a carbon emission dynamic evolution system and its application. *Environ Sci Pollut Res* 30(8): 20704-20720. doi: 10.1007/s11356-022-23541-4
- Olawade DB, Wada OZ, David-Olawade AC, Fapohunda O, Ige AO, Ling J. 2024. Artificial intelligence potential for net zero sustainability: Current evidence and prospects. *Next Sustain* 4: 100041. doi: 10.1016/j.nxsust.2024.100041
- Park JK, Kim HH, Lee WD. 2022. Measures to strengthen the implementation of local governments' policy towards carbon neutrality. Technical Report 2022-08.
- Patarasuk R, Gurney KR, O'Keeffe D, Song Y, Huang J, Rao P, Buchert M, Lin JC, Mendoza D, Ehleringer JR. 2016. Urban high-resolution fossil fuel CO₂ emissions quantification and exploration of emission drivers for potential policy applications. *Urban Ecosyst* 19(3): 1013-1039. doi: 10.1007/s11252-016-0553-1
- Rankl, F, Collins A, Tyers R, Carver D. 2023. The role of local government in delivering Net Zero. London, UK: The House of Commons Library. Debate Pack No. CDP-2023-0122.
- Rhoden I, Vögele S, Ball C, Kuckshinrichs W, Simon S, Mengis N, Baetcke L, Yeates C, Steuri B, Manske D, Thrän D. 2021. Spatial heterogeneity - Challenge and opportunity for net-zero Germany. <https://www.helmholtz-klima.de>
- Roest GS, Gurney KR, Miller SM, Liang J. 2020. Informing urban climate planning with high resolution data: The Hestia fossil fuel CO₂ emissions for Baltimore, Maryland. *Carbon Balance Manag* 15: 22. doi: 10.1186/

s13021-020-00157-0

- Seto KC, Churkina G, Hsu A, Keller M, Newman PWG, Qin B, Ramaswami A. 2021. From low- to net-zero carbon cities: The next global agenda. *Annu Rev Environ Resour* 46: 377-415. doi: 10.1146/annurev-environ-050120-113117
- Shin DS, Jung DJ, Kim HS, Hwang JW, Park JH, Shin YK, Bin EM, Choi JH, Kim JR, Lee MN. 2017. Guidelines for local government greenhouse gas inventories (Ver.4.1). Incheon, Korea: Korea Environment Corporation. K-eco(2017)-RF10-21.
- Shin SK, Son JS, Lee DY. 2023. Improvement strategies of urban space policy for the fulfillment of carbon neutrality. Sejong, Korea: Korea Research Institute for Human Settlements. Land Policy Brief No. 919.
- Tsagatakis I, Richardson J, Evangelides C, Pizzolato M, Ben R, Hows S-M, Pearson B, Passant N, Pommier M, Otto A. 2023. UK spatial emissions methodology - A report of the national atmospheric emission inventory 2021. https://naei.beis.gov.uk/reports/reports?report_id=1112
- UNFCCC (United Nations Framework Convention on Climate Change). 2018. Paris Agreement to the United Nations Framework Convention on Climate Change.
- Wu X, Tian Z, Guo J. 2022. A review of the theoretical research and practical progress of carbon neutrality. *Sustain Oper Comput* 3: 54-66. doi: 10.1016/j.susoc.2021.10.001
- Zhong S, Zhang K, Bagheri M, Burken JG, Gu A, Li B, Ma X, Marrone BL, Ren ZJ, Schrier J, et al. 2021. Machine learning: New ideas and tools in environmental science and engineering. *Environ Sci Technol* 55(19): 12741-12754. doi: 10.1021/acs.est.1c01339