

WRF-CMAQ 모델의 PM_{2.5} 구성성분 농도 모의 개선을 위한 기계학습 기법 적용

김지민¹, 최민서², 전예지¹, 김태희², 광경환¹

¹강원대학교 환경학과

²강원대학교 미세먼지통합관리학과

PM_{2.5} 구성성분 농도는 오염물질의 배출원을 파악하고 대기오염 관리에 필요한 정보를 제공하는 중요한 지표이다. 현재 환경부에서 권역별 대기환경연구소를 운영하여 PM_{2.5} 구성성분 농도를 측정 및 관리하고 있다. 하지만 실측만으로 대기질을 관리하고 정책을 평가하는 것은 시공간적 한계가 있으며, 중규모 화학 수송 모델 결과를 이용한 해석이 필요하다. WRF-CMAQ 모델은 국내에서 대기질 정책 효과를 평가할 때 주로 사용되는 모델로, 지속적인 시공간적 모의가 가능한 장점이 있다. 하지만, 배출량, 기상 과정 및 화학 메커니즘의 불확실성으로 인해 정확성이 떨어진다는 단점이 있다. 최근 수치 모델과 기계학습(Machine Learning)을 결합하여 예측 성능을 높이는 연구가 활발히 진행되고 있다. 기계학습 모델을 활용하면 기상 및 지형 조건을 고려하여 수치 모델의 모의 결과를 사후 보정할 수 있다. 따라서 본 연구에서는 기계학습 기법을 활용하여 WRF-CMAQ 모델의 PM_{2.5} 농도 및 구성 성분 모의 성능을 개선하는 방법을 제안하고자 한다. 기계학습 입력자료로 WRF-CMAQ 모델의 모의 결과를 사용하였으며, PM_{2.5} 및 구성성분(SO₄²⁻, NO₃⁻, Na⁺, NH₄⁺, K⁺, Cl⁻, Mg₂⁺, Ca²⁺, OC) 농도를 사용하였다. 기상 변수는 u, v 바람 성분, 기온, 습도, 해면기압을 사용하였다. 지형자료 관련 변수는 Terrain Height, Land mask, Albedo를 사용하였으며, PM₁₀ 배출량 자료를 사용하였다. 학습자료로 대기환경연구소 PM_{2.5} 성분 측정 자료를 사용하여 기계학습 모델을 구축하였다. 연구 대상 기간은 2021년과 2022년의 1월 1일부터 3월 31일이다. 연구 대상 지역은 대기환경연구소가 위치한 10개 권역(백령도, 수도권, 호남권, 중부권, 제주도, 영남권, 경기권, 충청권, 전북권, 강원권)이다. 기계학습 모델 검증 결과 R²는 0.85 이상으로 나타났으며, 기계학습 예측 성능이 양호한 것을 확인하였다. 이를 통해 본 연구에서 제안한 대기환경연구소 자료를 학습하는 기계학습 기법이 CMAQ 모델의 PM_{2.5} 구성성분 농도 모의 성능을 개선할 수 있음을 확인하였다.

Key words: WRF-CMAQ, PM_{2.5} Components, Machine Learning

※ 대기환경연구소 측정 자료를 제공해주신 국립환경과학원 대기환경연구과에 감사드립니다. 본 논문은 환경부를 지원기관으로 하는 강원권 미세먼지연구·관리센터 사업의 지원을 받았습니다. 이 성과는 정부(환경부)의 재원으로 한국환경산업기술원의 미세먼지관리특성화대학원 사업의 지원을 받아 수행된 연구입니다.