기후 분과 / 기후 2-1

# A Super-Resolution Framework for Downscaling Machine Learning Weather Prediction toward 1-km surface air temperature

Hyebin Park[1], Seonyoung Park[1], Daehyun Kang[2], Jeong-Hwan Kim[2]

[1]Department of Applied Artificial Intelligence, Seoul National University of Science and Technology
[2]Center for Climate and Carbon Cycle Research, Korea Institute of Science and Technology

Deep learning technic has improved the accuracy and efficiency of weather forecasting, outperforming traditional numerical weather prediction models. However, the coarse spatial resolution of global weather forecasting systems limits their ability to resolve fine-scale surface heterogeneity and localized extremes, particularly in regions with complex terrain or urban heat island effects. Here, we introduce a deep learning based super-resolution framework that transforms coarse 0.25° forecasts into 1 km air temperature fields using MODIS-derived temperature targets and high-resolution auxiliary inputs. Our model outperforms existing super-resolution methods by explicitly incorporating spatial context such as topography, impervious surface fraction, and climatological anomalies. When this model was applied to FuXi global weather forecast, the 7-day forecast error in South Korea decreased more than 20%, which is comparable to 1-day forecast error of low-resolution prediction with simple spatial interpolation. Although validated over South Korea, the method is region-agnostic and readily generalizable due to the global availability of MODIS inputs and minimal auxiliary data requirements. This super-resolution framework is a scalable and high-fidelity tool for enhancing ML-based weather forecasts at fine spatial scales.

**Keywords:** Artificial Intelligence, Weather Prediction, Downscaling, Super-Resolution

※ This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2022M3K3A1094114) and Korea Institute of Science and Technology Research Program (2E33621)

기후 분과 / 기후 2-2

# Artificial Intelligence in Climate Science: A Topic Modeling Analysis of Global Research Trends

Tae-Jong Kim, So-Hyung Kim

Meteorological Research Division, National Institute of Meteorological Sciences

Climate change, a data-intensive global challenge, increasingly necessitates the use of Artificial Intelligence (AI) for effective analysis and solution development. Although the application of AI in climate science is rapidly expanding, a systematic and macroscopic map of its intellectual structure remains limited. This study quantitatively analyzes the global research landscape of the climate–AI domain to identify core research areas and detailed thematic topics, thereby offering a comprehensive reference for academics and policymakers. We collected 5,103 articles published between 2018 and 2025 from the Web of Science Core Collection. Latent Dirichlet Allocation (LDA) topic modeling was applied to extract latent topics, with model reliability optimized using the coherence score.

The analysis identified four major research areas. The largest, AI-utilizing Climate Change Mitigation and Policy Analysis (30.41%), emphasizes actionable, solution-oriented strategies, encompassing topics such as AI-integrated climate response technologies and sustainability governance (Topic 2), AI-based energy and industrial optimization for carbon reduction (Topic 5), and quantitative assessments of carbon emission policies and market impacts (Topic 13).

The second, AI-based Climate Change Impact and Risk Analysis (27.96%), highlights the use of AI to quantify diverse consequences of climate change, including agricultural productivity monitoring and hazard detection (Topic 3), ecosystem and species distribution modeling (Topic 9), water resources and runoff scenario analysis (Topic 10), and climate disaster risk and vulnerability assessments (Topic 11).

The third, AI-based Climate System Prediction and Modeling (26.55%), focuses on fundamental challenges in climate science, including meteorological and hydrological variable prediction (Topic 1), deep learning-based climate and carbon emission forecasting (Topic 4), AI-enhanced climate model improvement and physical parameterization (Topic 7), and large-scale variability prediction for precipitation and wind patterns (Topic 8).

Finally, AI Methodology and Data-driven Analysis Techniques (15.07%) provides methodological and technical foundations, including urban heat environment and land cover change analysis (Topic 6), remote sensing–based environmental monitoring (Topic 12), and performance benchmarking of machine learning algorithms in climate research (Topic 14).

This study provides a data-driven, comprehensive map of the rapidly evolving climate–AI research field, presenting empirical evidence of its thematic pillars and their relative significance. The findings serve as a valuable guide for future research directions and evidence-based R&D strategy development

**Keywords:** Climate Change, Artificial Intelligence, Research Trends, Topic Modeling

※ This work was funded by the Korea Meteorological Administration Research and Development Program "Developing Intelligent Assistant Technology and Its Application for Weather Forecasting Process" under Grant (KMA2021-00123).

기후 분과 / 기후 2-3

# 딥러닝 기반 3차원 다변량 해양 자료동화 기법 개발

이수빈, 함유근

서울대학교 환경대학원 환경관리학과

해양 자료동화 시스템은 수치모형 예측과 해양 관측을 결합하여 현재 해양 상태를 추정함으로써 전지구 기후 변동을 예측 능력을 향상시켜왔다. 그러나 해양 온도와 염분 등 개별 변수를 각각 동화하는 방식은 변수 간 상호 관련성을 반영하지 못해 안정적인 예측에 한계가 있다. 이에 따라 온도와 염분을 함께 고려하는 다변량 자료동화 기법의 적용이 필수적이다. 하지만 기존의 다변량 자료동화 기법은 변수 간의 관측 공분산을 고려하는 과정에서 연산량이 기하급수적으로 증가하고, 변수들 사이의 비선형적 관계를 효과적으로 고려하기 어렵다는 문제가 있다. 본 연구는 이러한 한계를 극복하기 위해 Gated Convolution 기법을 활용한 딥러닝 기반 다변량 자료동화 시스템을 제안한다.

본 연구는 불규칙한 해양 관측을 다루고, 자료동화 과정에서 분석장의 신뢰도를 반영할 수 있도록 한다. Gated Convolution을 통해 산출된 Gating Value는 자료의 신뢰도 정보를 나타내며, 관측 수가 많은 지역일수록 높은 값을 보였다. 또한 Gating Value가 클수록 RMSE(Root Mean Square Error)가 낮아지는 것을 확인함으로써, 제안된 방법이 분석장의 신뢰도를 효과적으로 반영함을 확인하였다. 추가로 온도와 염분 간 상관관계가 높은 지역에서는 온도 자료를 활용하는 것이 상대적으로 관측 밀도가 낮은 염분 자료동화의 성능향상에 기여함을 실험을 통해 확인하였다. 또한, 개발된 딥러닝 기반 분석장(DeepDA)을 다른 재분석장들과 비교한 결과, EN4.2.2 데이터와 더 높은 상관관계를 보였으며 이를 통해 본 연구의 자료동화 시스템이 해양 상태를 보다 정확하게 추정함을 확인하였다. 이러한 결과는 본 연구에서 제안된 자료동화 시스템이 기존 방법의 한계를 극복하고 해양 상태를 보다 정확하게 모의할 수 있음을 시사한다.

**Keywords:** 딥러닝, 해양 자료동화, 다변량 자료동화

기후 분과 / 기후 2-4

# 인공지능 기반 대기-식생 시스템의 우리나라 여름철 폭염 예측 성능 평가

정수연[1], 조동진[1], 함유근[1,2]

[1]서울대학교 환경계획연구소
[2]서울대학교 환경대학원 환경관리학과

기후변화로 인해 우리나라 여름철 폭염과 열대야의 발생 빈도가 증가하고 있으며, 이에 따른 인명 피해도 지속적으로 늘어나고 있다. 이러한 피해를 최소화하기 위해서는 높은 정확도의 폭염 예측 시스템 구축이 필수적이다. 특히 선행연구들은 폭염과 식생 간 상호작용의 중요성을 제시해왔으나, 인공지능 기반 기온 예측 연구에서는 식생 정보를 충분히 반영한 사례가 여전히 부족하다. 이에 본 연구에서는 머신러닝 기반의 식생 정보를 효과적으로 활용하여 한반도 여름철 기온 예측의 정확도의 향상 가능성을 평가하고자 하였다. 본 연구는 한반도 전역에 분포한 54개 종관기상관측소를 대상으로 하였으며, 2000년부터 2019년까지 20년간의 여름철(6–8월) 자료를 활용하였다. 폭염 예측에서 식생 정보의 효과를 검증하기 위해 두 가지 실험 체계를 적용하였다. 첫 번째 실험(Scheme 1)은 대기장 변수만을 입력으로 사용하였고, 두 번째 실험(Scheme 2)은 동일한 대기장 변수에 더해 식생 상태를 나타내는 일별 LAI(Leaf Area Index) 이상(anomaly) 값을 추가하여 학습에 반영하였다. 예측 모델로는 기상·환경 분야에서 널리 사용되는 머신러닝 기법인 LightGBM 알고리즘을 적용하였으며, Leave-One-Year-Out Cross Validation을 통해 교차검증을 수행하였다. 이를 바탕으로 Scheme별 Skill Score를 산출하여 모델 예측 성능 차이를 정량적으로 평가하였다. 그 결과, 한반도 중부 내륙 지역에서 Scheme 2가 Scheme 1 대비 유의미한 성능 향상을 보였으며, 이는 식생의 동적 변화를 반영하는 것이 폭염 예측 정확도 향상에 기여함을 보여준다. 또한 민감도 실험을 통해 인공지능 모델이 대기-식생 상호작용을 학습함을 확인하였고, 식생 감소가 폭염의 발달과 강도와 관련됨을 검증하였다. 따라서 본 연구는 식생 정보의 활용이 한반도 여름철 폭염 예측 성능 향상에 실질적으로 기여할 수 있음을 시사한다.

**Keywords:** 폭염, 대기-식생 상호작용, 머신러닝

기후 분과 / 기후 2-5

# 딥러닝 기반 4DVAR 자료동화 기법 개발

이현정[1], 함유근[1,2]

[1]서울대학교 환경대학원 환경관리학과
[2]서울대학교 환경계획연구소

4차원 변분(4DVAR) 자료동화 기법은 대기·해양 모델의 초기 조건을 최적화함으로써 예측 성능을 향상시킬 수 있는 방법론으로, 이를 구현하기 위해서는 배경 오차 공분산 행렬(B)과 Adjoint 모델이 필수적이다. B는 배경장의 오차 크기와 공간적 상관구조를 반영해 관측 정보가 동화 과정에서 어떤 공간적 범위와 변수 간 관계 속에서 적용될지를 규정하며, Adjoint 모델은 목적함수의 기울기를 계산하고 민감도를 전파함으로써 관측 정보가 초기장으로 전달되도록 한다. 이 두 요소의 품질은 4DVAR의 수렴 효율과 분석장의 신뢰도를 좌우하므로, 안정적이고 효과적인 자료동화를 위해 반드시 필요하다. 그러나 기존 연구들은 전통적 수치모델 기반 접근에 의존해왔으며, B는 통계적 방법으로 근사되었지만 다수의 앙상블 실행을 요구해 계산 비용이 크고, 통계적 가정에 따라 복잡한 공간·시간적 상관구조를 충분히 반영하지 못하는 한계가 있다. 특히 고차원 상태벡터에 대해 대규모 행렬 연산으로 인해 계산비용이 기하급수적으로 증가한다. 또한, Adjoint 모델은 수치모델의 선형화에 기반해 구현되므로 비선형 과정을 충분히 고려하기 어렵고, 모델의 복잡성이 커질수록 개발 및 계산 비용이 크게 증가하는 문제가 있다.

따라서 본 연구에서는 딥러닝 기법을 적용하여 4DVAR 구현에 필요한 핵심 구성요소를 모델링하였다. B는 기존 연구에서 통계적 특성을 고려하기 위해 많은 학습 데이터를 요구하는 지도학습(supervised learning)에 의존해 왔으나, 본 연구에서는 가 배경 오차의 상관성을 제거하고 분산을 표준화하는 성질에 착안하여, 이를 학습 목표로 하는 비지도학습(unsupervised learning) 방식을 적용하였다. 이 접근은 별도의 정답 데이터 없이도 B의 구조적 특성을 학습할 수 있다는 장점을 가지며, 실제 실험에서 일정 수준의 작동 가능성을 확인하였다. 또한, Adjoint 모델은 딥러닝의 역전파(backpropagation) 과정과 구조적으로 유사하다는 점에 착안하여 구현되었다. 다수의 기상 변수들을 입력장으로 처방한 U-net 모형은 목적 변수의 일반적인 상류(upstream) 지역에서의 민감도 강화뿐 아니라, 순환장, 압력장 등에 대한 민감도 변화를 함께 반영함으로써, 실제 대기 과정의 특성을 현실적으로 고려할 수 있음을 보였다. 본 연구는 딥러닝 기반 4DVAR 체계 구현의 가능성을 제시하며, 향후 대기·해양 자료동화에 새로운 접근을 마련할 수 있을 것으로 기대된다.

**Keywords:** 4DVAR, 자료동화, 딥러닝

기후 분과 / 기후 2-6

# 기상기후데이터와 트랜스포머 기반 딥러닝 모델을 활용한 하천 수위 예측 성능 분석

박지훈[1], 윤다은[2], 안성욱[3], 김병식[4]

[1]강원대학교 방재전문대학원 석사과정
[2]강원대학교 AI기후재난기술융합연구소 주임연구원
[3]강원대학교 방재전문대학원 박사과정
[4]강원대학교 전자·AI시스템공학과 / 방재전문대학원 정교수

본 연구는 기후변화로 인한 극한 강우와 그로 인한 국지적 홍수 위험의 증가에 대응하기 위해, 하천 수위 예측의 정밀도와 신뢰성을 높이고자 최신 딥러닝 기반 시계열 모델들을 적용·비교하였다. 연구 대상은 한탄강과 금강 인근 삽교천 유역으로, 기상관측소 강우량, 하천 수문 자료, 댐 방류량을 통합한 데이터셋을 구축하여 예측 실험에 활용하였다. 적용된 모델은 전통적인 LSTM, GRU 계열 모델을 보완하는 최신 구조로서, 장·단기 패턴 학습과 외생변수 반영에 강점을 가진 TFT(Temporal Fusion Transformer), 긴 시계열 처리에 효율적인 Informer, 시공간 패턴을 동시에 학습할 수 있는 ConvLSTM을 중심으로 분석하였다. 성능 평가 결과, 세 모델은 MAE, RMSE, NSE 등에서 기존 모델들보다 우수한 결과를 보였으며, 특히 앙상블 기법 적용 시 검증기간(2024.06 ~ 2024.08) 동안 실제 수위와 매우 유사한 시계열 변화를 나타냈다. 더불어 Meta-Gaussian 기법을 도입하여 예측값의 신뢰구간을 산출하고 불확실성을 정량적으로 평가함으로써, 모델 결과의 실질적 활용 가능성을 높였다. 본 연구는 최신 딥러닝 기법의 적용을 통해 하천 수위 예측 정확도를 개선했을 뿐만 아니라, 예측 신뢰성 확보 측면에서도 의미 있는 기여를 했다는 점에서 홍수 조기경보 체계와 수자원 관리의 고도화에 기여할 수 있다. 다만, 연구 범위가 특정 유역과 강우 중심 입력에 제한되어 있어 댐 운영, 토양수분, 지형학적 요인 등 다양한 변수의 반영이 필요하며, 향후에는 더 많은 유역과 극한 기상사상을 고려한 확장 연구가 요구된다.

**Keywords:** 하천 수위 예측, 기상데이터, TFT, Informer, ConvLSTM

※ 이 논문은 기상청 재원으로 한국기상산업기술원의 기상기후데이터 융합분석 특성화대학원 사업의 지원을 받아 수행되었습니다.