

GIST, 대기를 입체(3D)로 읽는 AI 기술로 정확도 크게 높인 기상예보 구현

온도·강수 예측 오차 최대 31% 감소, 한 달 뒤 장기예보까지 정확도 향상

- 환경·에너지공학과 윤진호 교수가 이끈 국제공동연구팀, 기존 수치예보의 넓은 예보 구역(120km)을 23km 수준으로 세분화해 지역별 기상 변화의 정밀도 대폭 향상
- AI가 시간·공간·지형 정보 함께 학습해 기존 모델의 구조적 오차 보정하고, 적은 계산 자원으로도 구동돼 실용성·확장성 확보... 국제학술지 《Geoscientific Model Development》 게재



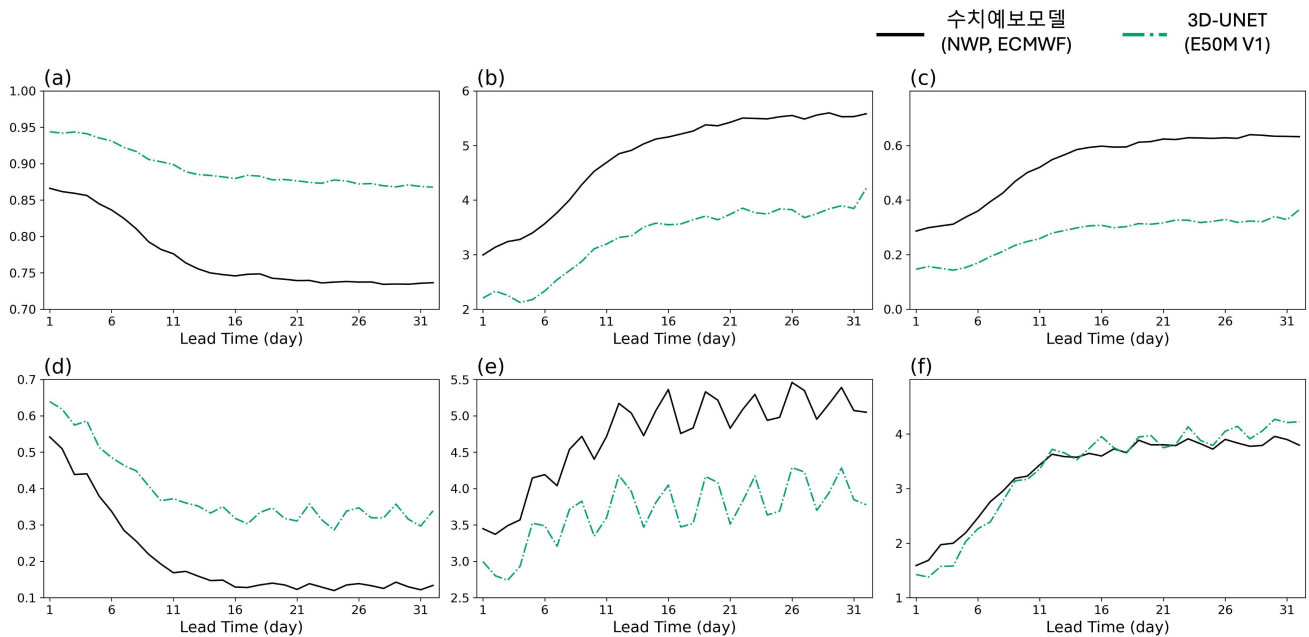
▲ (왼쪽부터) GIST 윤진호 교수, 미국 유타주립대학교 류지훈 박사후연구원, GIST 김희수 석사과정생 광주과학기술원(GIST, 총장 임기철)은 환경·에너지공학과 윤진호 교수가 이끄는 국제 공동연구팀이 인공지능(AI) 기술을 활용해 미국 서부의 기상 상태를 최대 한 달 뒤(32일)까지 기존보다 훨씬 더 세밀하고 정확하게 예측할 수 있는 새로운 방법을 개발했다고 밝혔다.

복잡한 산악·해안·내륙 지형이 얹혀 예측 난도가 높은 미국 서부에서 성능이 검증됐다는 점에서, 기후위기 시대 고해상도 예보 기술의 새로운 가능성을 열었다는 평가다.

연구팀은 기존 기상청·유럽중기예보센터(ECMWF) 등에서 활용하는 수치예보(NWP) 모델*이 약 120km(1.5도) 간격으로 넓게 나뉜 예보 구역 단위로 정보를 제공해 지역 특성을 충분히 반영하지 못하는 문제를 해결하는 데 초점을 맞췄다.

특히 미국 서부는 고도 차이가 크고, 바다에서 들어오는 공기와 내륙에서 내려오는 공기가 번갈아 영향을 주는(해양-내륙 간 기단 교환이 활발한) 지역이라 실제 날씨가 지형에 따라 크게 달라지며 예측이 어려운 곳으로 알려져 있다.

* 수치예보(NWP, Numerical Weather Prediction) 모델: 대기의 물리·역학적 과정을 수학 방정식으로 표현하고 이를 초고속 컴퓨터로 계산해 미래의 기상 상태를 예측하는 시스템이다. 온도, 기압, 바람, 습도 등의 관측 자료를 초기값으로 입력해 시간에 따른 대기 변화를 모의하며, 현대 기상예보의 핵심 기반으로 활용된다.



▲ 서부 미국 지역을 대상으로 한 32일 예측 성능 비교 결과. (a) -(c)는 온도, (d)-(f)는 강수를 의미하며, (a), (d)는 상관계수, (b)와 (e)는 RMSE, (C)와 (f)는 분산과 상관계수 종합적 지표를 의미함. 기존 수치예보 (NWP, ECMWF) 대비 3D U-Net 기반 모델(E50M V1)이 온도와 강수에서 전반적으로 더 높은 상관계수와 더 낮은 예측 오차(RMSE)를 보이며, 중기 예측 성능이 크게 개선되었음을 보여줌.

이를 해결하기 위해 연구팀은 날씨가 시간에 따라 어떻게 이어지고 변화하는지를 함께 학습하도록 설계한 '3차원(3D) U-Net' 기반 AI 예보후처리 모델을 개발했다.

이 모델은 오늘부터 예측하려는 날짜까지의 시간 구간(예보 선행시간*)을 하나의 연속된 흐름으로 분석해, 비교적 정확도가 높은 초단기·중기(1~10일) 예보*의 정보를 바탕으로 그 이후 연장중기*(10~32일) 예보까지 정확도를 자연스럽게 확장하도록 설계됐다.

즉, 기존 수치예보가 제공하는 정보를 단순히 보정하는 수준을 넘어, 시간·공간·지형 특성을 동시에 반영해 더 현실적인 결과를 내도록 만든 것이다.

* **U-Net**: 의료 영상 분석을 위해 개발된 딥러닝 기반의 합성곱신경망(CNN) 구조로, 인코더(축소 경로)와 디코더(확장 경로)가 U자 형태로 연결된 것이 특징이다. 영상의 세부 공간 정보를 보존하면서 특징을 추출하고, 이를 기반으로 정밀한 픽셀 단위 분할(segmentation)을 수행하는 데 매우 효과적이다. 현재 의료 영상뿐 아니라 위성영상, 생명과학, 자율주행 등 다양한 분야에서 널리 활용되고 있다.

* **예보 선행시간(Forecast Lead Time)**: 예보가 발행된 시점부터 실제 예측하고자 하는 시점까지의 시간 간격을 의미한다. 예를 들어, '12시간 예보'라면 예보 발표 시각보다 12시간 뒤의 기상 상태를 예측하는 것으로, 예보의 시간적 범위와 정확성을 평가하는 데 중요한 개념이다.

* **초단기 예보**: 앞으로 36시간 이내의 기상 상태를 예측하는 예보를 의미한다. 단기간의 기온, 강수, 풍속 등 기상 변화를 신속하게 파악할 수 있어 재난 대응, 현장 운영, 실시간 의사결정 등에 활용된다.

* **중기 예보**: 36시간 이후부터 일주일 이내의 기상 변화를 예측하는 예보를 의미한다. 단기 예보보다 넓은 시간 범위의 기상 흐름을 파악할 수 있어 생활·산업 활동 계획 수립에 활용된다.

* **연장중기**: 일주일 이후부터 한 달 이내의 기상 예측을 다루는 예보 범위를 말한다. 단기·중기 예보보다 더 긴 시간대의 기상 경향을 파악하는 데 활용되며, 농업, 에너지 수요 관리, 일정 계획 등 중·장기적 의사결정에 도움을 준다.

연구팀이 개발한 AI 모델은 유럽중기예보센터(ECMWF)의 예보 자료를 기반으로 약 23km(0.25도) 수준의 고해상도 정보를 생성하도록 학습됐으며, 동시에 수치예보에 반복적으로 나타나는 구조적 오차까지 보정하는 기능까지 갖췄다.

이 기술은 약 120km 간격으로 넓게 구분돼 있던 예보 구역을 23km 수준의 훨씬 촘촘한 구역으로 다시 세분화해, 더 작은 지역 단위까지 예측할 수 있도록 하고 산악·해안·내륙 등 지형에 따른 기상 변화를 더욱 정밀하게 반영할 수 있게 한다. 단순히 해상도를 높이는 데 그치지 않고, 실제 기상 패턴과의 차이를 줄이도록 예보 오차를 정교하게 조정하는 고도화된 방식이다.

성능 평가 결과, 새 모델은 실제 기상 변화와의 일치도가 눈에 띄게 높아졌다.

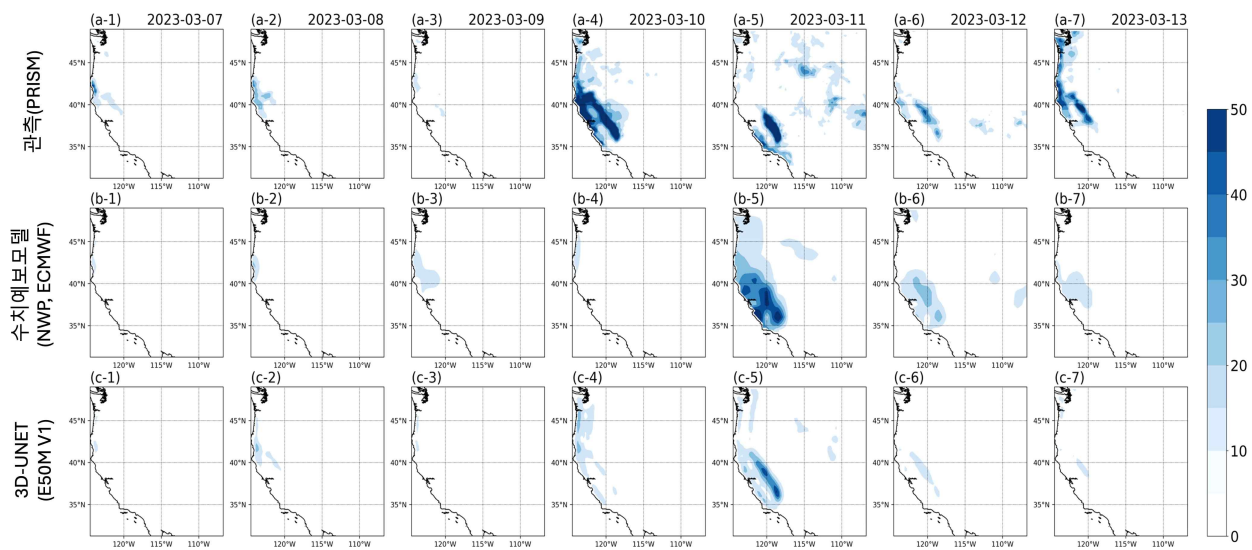
온도의 경우 기상 패턴을 얼마나 잘 맞추는지를 보여주는 상관계수*가 기존보다 0.12 높아졌고, 강수 예측에서도 0.18 상승했다. 상관계수는 값이 1에 가까울수록 실제 기상 변화와 더 일치하는 지표다. 또한 예측 오차(RMSE*)는 온도 기준 약 31%, 강수 기준 약 22% 줄어, 전체적으로 기존 수치예보 대비 큰 폭의 정확도 향상을 보였다.

특히 산악 지형에서 나타나는 미세한 온도 변화나 해안가에 형성되는 강수 집중 구역, 내륙 농경지에서 자주 발생하는 국지적 변화처럼 기존 예보 모델이 포착하기 어려웠던 패턴까지 정밀하게 재현하는 것으로 나타났다. 이는 지형의 영향이 크고 바다와 대기의 상호작용이 복잡하게 얽혀 있는 지역에서, AI 기반 보정 기술이 기존 수치예보가 가진 구조적 한계를 효과적으로 보완할 수 있음을 보여주는 결과다.

* **패턴 상관계수(Pattern Correlation)**: 어떤 변수가 공간적으로 어떻게 분포하는지를 모델이 얼마나 정확하게 재현했는지를 평가하는 지표이다. 관측값과 예측값의 공간 패턴 유사도를 상관계수 형태로 나타내며, 1에 가까울수록 공간 분포를 더 정확하게 모의했음을 의미한다.

* **RMSE(Root Mean Square Error)**: 관측값과 예측값의 차이를 제곱하여 평균한 뒤 제곱근을 취해 계산하는 지표로, 예측 모델의 오차 크기를 나타내는 대표적인 평가 척도이다. 값이 작을수록 모델의 예측 정확도가 높음을 의미한다.

연구팀은 2023년 미국 캘리포니아에서 기록적인 폭우가 쏟아졌던 사례를 추가로 분석했다. 그 결과, 새 모델이 비가 내리는 위치와 분포는 더 정확하게 포착하지만 실제 내린 비의 양(절대 강수량)은 다소 적게 예측하는 경향이 있는 것으로 확인됐다.



▲ 2023년 3월 6일에 예보한 3월 7~13일 서부 미국 극한 강수 사례 비교. 관측(PRISM)과 비교했을 때, 3D U-Net 모델(E50M V1)이 강수 위치와 이동 경로를 기존 수치예보(NWP, ECMWF)보다 더 정확하게 재현하며, 극한 강수의 공간 분포를 더 잘 포착하는 모습을 보여줌.

이는 해외에서 개발된 최신 AI 기반 기상예측 모델에서도 공통적으로 나타나는 한계로 알려져 있으며, 강수의 규모(양)를 정확하게 예측하는 것은 여전히 해결되지 않은 과제로 남아 있다는 것을 시사한다.

주목할 점은, 이번 연구가 복잡한 수십 개의 입력 정보를 더하거나 여러 모델의 결과를 모두 세세하게 활용하지 않고, 여러 예보 결과를 묶어 평균낸 값(앙상블 평균*)과 가장 중요한 몇 가지 정보만으로도 높은 예측 성능을 냈다는 것이다.

이런 방식 덕분에 모델이 차지하는 메모리가 줄고 계산 시간도 크게 단축돼 고가 장비가 아닌 일반적인 GPU 환경에서도 안정적으로 운영할 수 있다. 즉, 대규모 AI 기상예보 시스템을 구축하기 전에 예측 능력을 먼저 끌어올릴 수 있는 현실적이고 효율적인 대안으로 평가된다.

또한 추가 학습만 더하면 전 세계 어느 지역에도 적용할 수 있을 만큼 확장성이 높아 산업적·학술적 활용 가능성이 매우 높다.

* 앙상블 평균(Ensemble Mean): 수치예보 모델이 서로 다른 초기 조건으로 생성한 여러 예측 결과(앙상블)를 평균해 얻은 값으로, 단일 예측보다 변동성이 적고 더 안정적이며 신뢰도 높은 예측을 제공하는 방식이다.

윤진호 교수는 "기후변화로 예측의 중요성이 커지는 가운데 기존 모델이 만들어 낸 결과를 AI로 한 번 더 보정해 정확도를 높이는 '후처리 기술'은 수치예보의 한계를 보완하는 유력한 해법이 될 것"이라며 "특히 지형과 지역성이 복잡한 미국 서부 지역 사례에서 보듯, AI는 고해상도 지역 예보를 구현하는 데 결정적인 역할을 할 수 있다"고 강조했다.

이어 “예측 정확도는 높이고 계산 부담은 줄여 운영 효율성까지 끌어올린 기술이기에 산불·홍수·가뭄 등 기후 재난 대응에도 큰 도움이 될 것”이라고 전망했다.

GIST 환경·에너지공학과 윤진호 교수와 미국 유타주립대학교 류지훈 박사후연구원이 주도한 이번 연구에는 GIST 김희수 석사과정생과 태국 까셋삿대학교(Kasetsart University)의 사이먼 왕(Shih-Yu(Simon) Wang) 교수가 참여했으며, 한국연구재단 ‘AI기반미래기후기술 개발 원천연구사업’과 기상청 ‘가뭄특이기상센터’의 지원을 받았다.

연구 결과는 국제학술지 《Geoscientific Model Development》에 2026년 1월 5일 온라인으로 게재됐다.

한편 GIST는 이번 연구 성과가 학술적 의의와 함께 산업적 응용 가능성까지 고려한 것으로, 기술이전 관련 협의는 기술사업화센터(hgmoon@gist.ac.kr)를 통해 진행할 수 있다고 밝혔다.

논문의 주요 정보

1. 논문명, 저자정보

- 저널명 : Geoscientific Model Development (Impact Factor: 4.9) 지구과학-종합 분야 상위 11% 이내(28/258)의 저명한 국제학술지
- 논문명 : Increasing Resolution and Accuracy in Sub-Seasonal Forecasting through 3D U-Net: the Western US
- 저자 정보 : Jihun Ryu (제1저자, GIST 박사과정 졸업생, 유타주립대 박사후연구원), Hisu Kim(참여저자, GIST 석사과정), Shih-Yu Wang, Jin-Ho Yoon (교신저자, GIST 교수)