

Machine Learning Global Simulation of Nonlocal Gravity Wave Propagation

Aman Gupta^{*1}, Aditi Sheshadri^{*1}, Sujit Roy^{*2}, Vishal Gaur², Manil Maskey³, Rahul Ramachandran³

¹Department of Earth System Science, Stanford University, Stanford, USA; ²Earth System Science Center, The University of Alabama in Huntsville, Huntsville, AL, USA; ³NASA Marshall Space Flight Center, Huntsville, AL, USA

^{*}Equal contribution, Aman: ag4680@stanford.edu, Sujit: sujit.roy@nasa.gov



Climate models often ‘parameterize’ physical processes they cannot resolve.

Consider, the atmospheric **gravity wave (GW)** shown here. The wave packet exhibits three-dimensional (lateral + vertical) propagation but, for decades, its model parameterizations have assumed strict vertical propagation.

This leads to a **representation bias**, that contributes to climate prediction uncertainty.

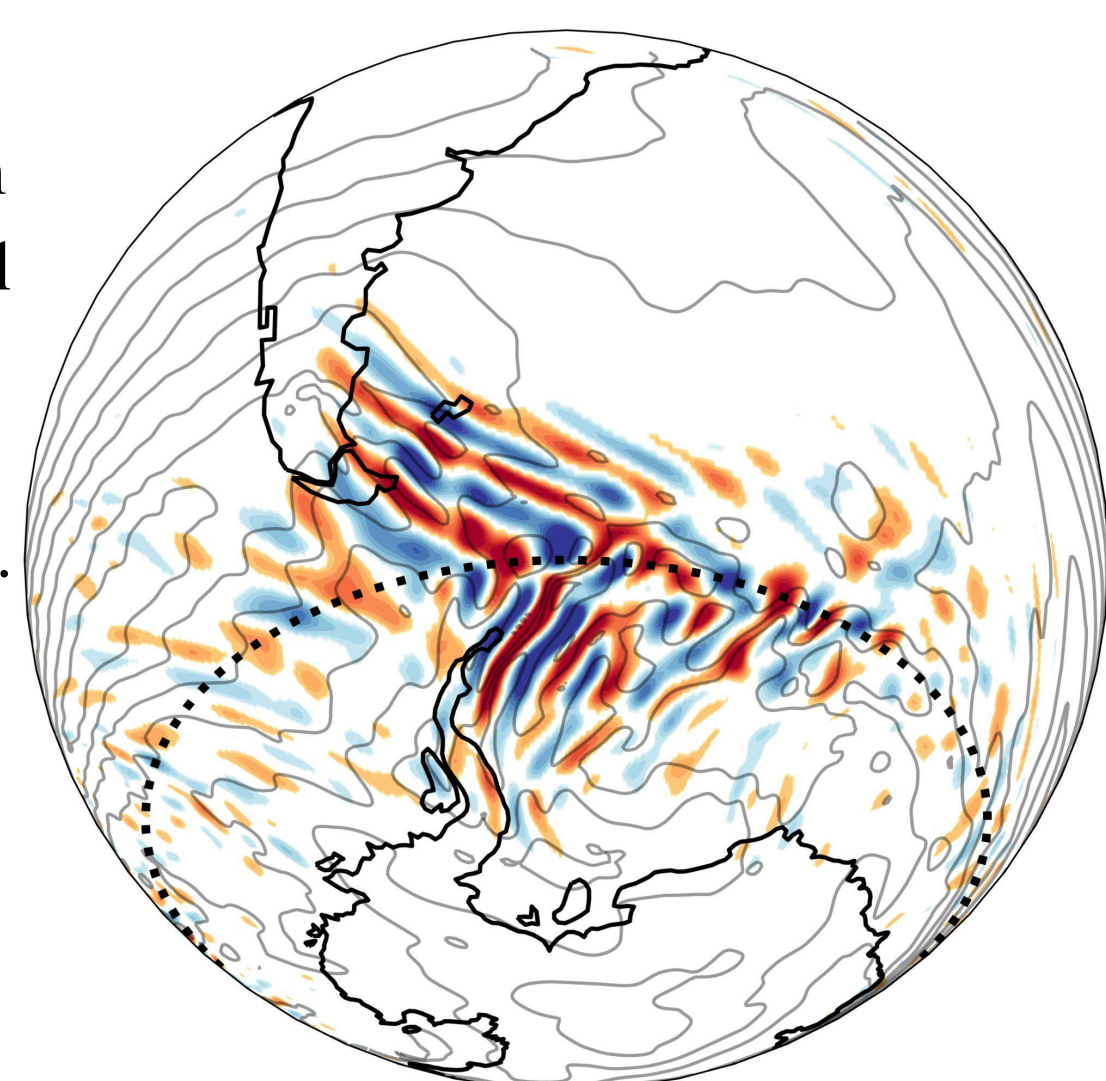


Fig 1: GWs on 17th July 2012 in the stratosphere.

Objectives:

- Develop an ML-based parameterization scheme capable of learning nonlocal GW evolution by training on global resolved fluxes
- Evaluate performance of a hierarchy of nonlocal ML architectures: single column artificial neural network (ANN) to generative encoders

Different methods of embedding spatial nonlocality

Given the background atmospheric conditions (u, v, θ), the models predict the two orthogonal components of the wave momentum flux ($u'\omega', v'\omega'$). We embed nonlocal physical information for different deep learning models with varying degree of complexity.

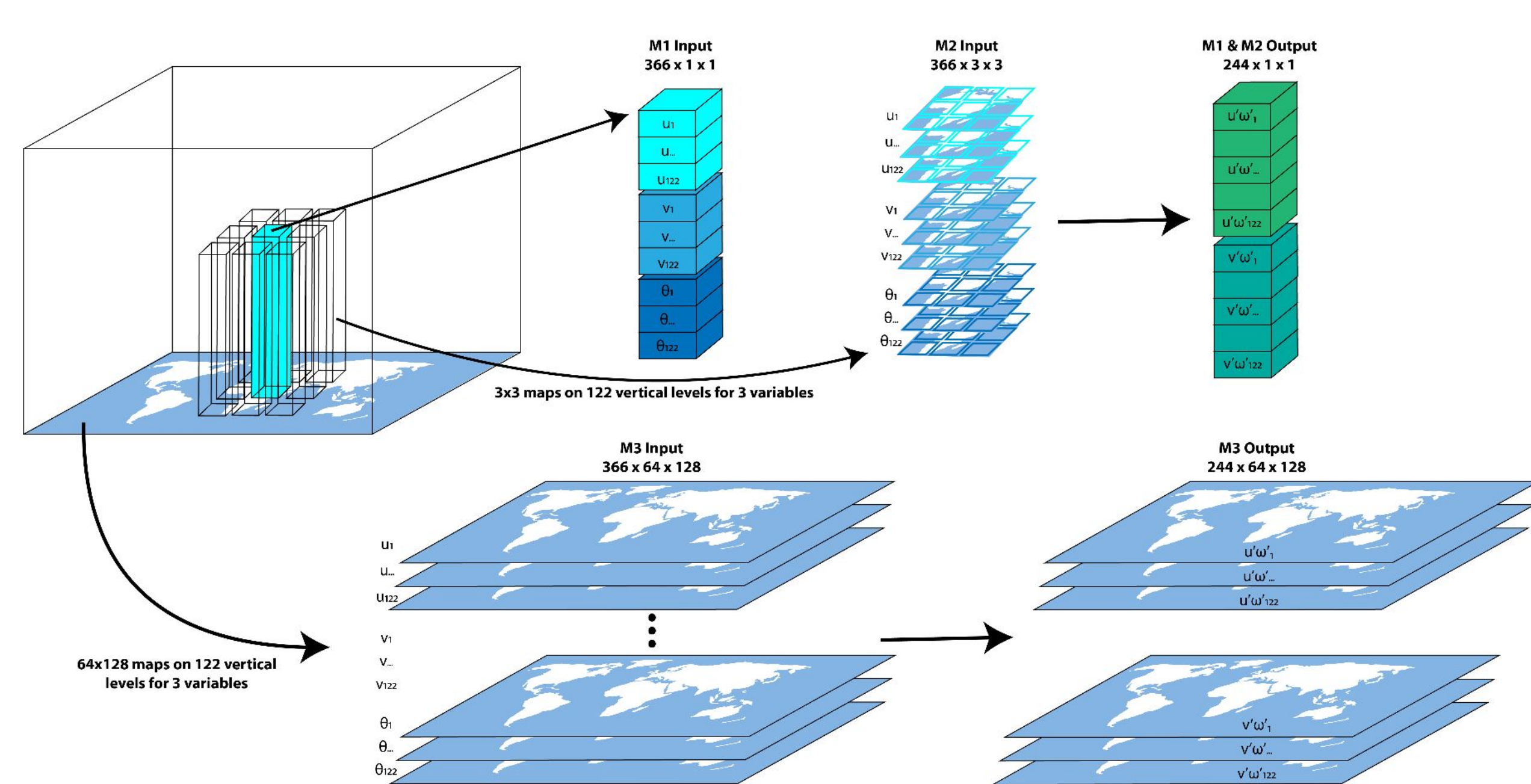


Fig 2: Model architectures. The three architectures used for global GW resolved momentum flux simulation. The three architectures employ three different degrees of nonlocality. On one end, M1 uses single-column background data to predict the fluxes within that column. A timeslice is therefore a single vector of length 366. Intermediately, M2 uses background information in a 3×3 stencil to predict the fluxes within the single-column at the center of the stencil. A timeslice for M2 has dimensions $3 \times 3 \times 366$. On the other end, M3 uses global maps of the background field to predict global maps of fluxes. A timeslice for M3, thus, has dimensions $366 \times 64 \times 128$.

‘Nonlocal’ architectures outperform ‘local’ architectures

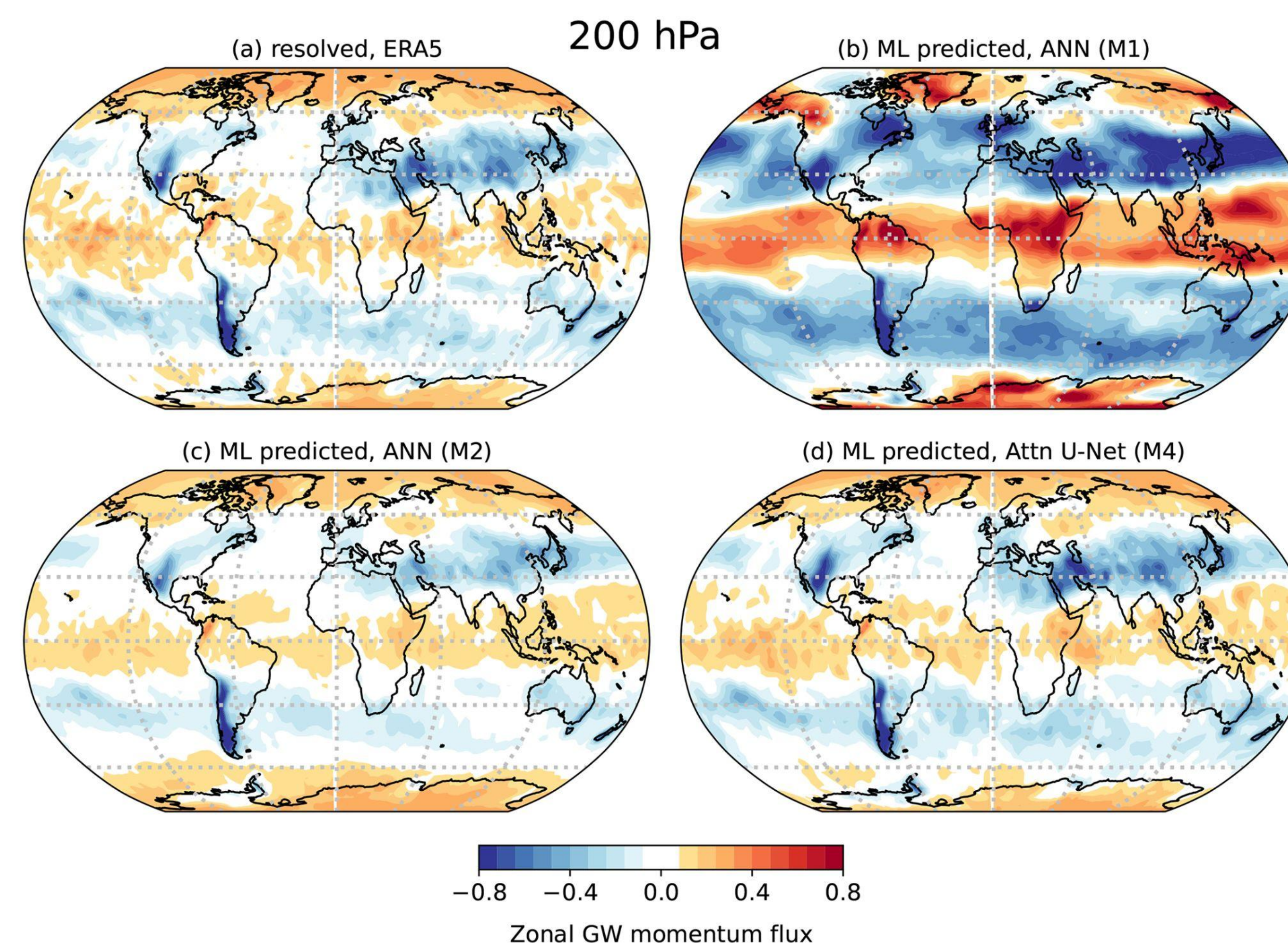


Fig 3: Mean prediction test. The model prediction improves with the degree of embedded nonlocality. The figure above shows the global distribution of predicted momentum fluxes at 200 hPa (~15 km height), averaged over the month of May 2015. The true, normalized momentum fluxes, computed from ERA5 reanalysis, are shown in (a). The 1x1 model (M1) notably overestimated the fluxes, as shown in (b). The predictions from the 3x3 model (M2), as shown in (c), are more accurate. Thus even a small degree of nonlocality appears to be sufficient to learn the nonlocal evolution of GW fluxes. Lastly, the global Attention U-net (M4) outperforms both the 1x1 ANN and the 3x3 CNN + ANN, as shown in (d).

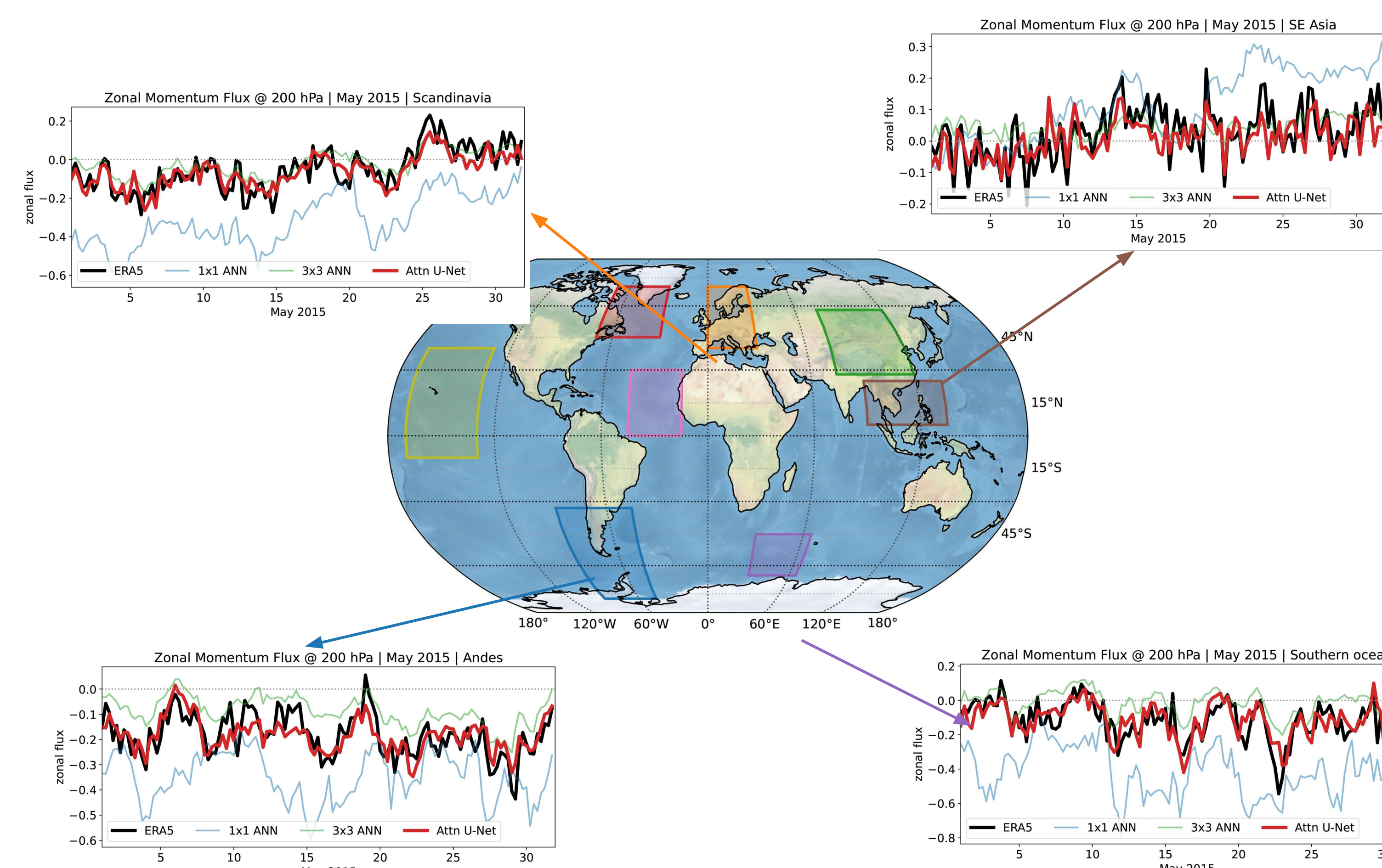


Fig 4: Instantaneous prediction test. The Attention U-Net model also successfully predicts spikes or intensification of fluxes over prominent orographic and nonorographic GW hotspots. The prediction skills, quantified by the coefficient of determination (R^2), indicates higher prediction skill in the midlatitudes than in the tropics. More generally, the prediction is better over stationary features (e.g., mountains) than over non-stationary features (e.g., strong convective systems).

Nonlocal models better capture the distribution tail

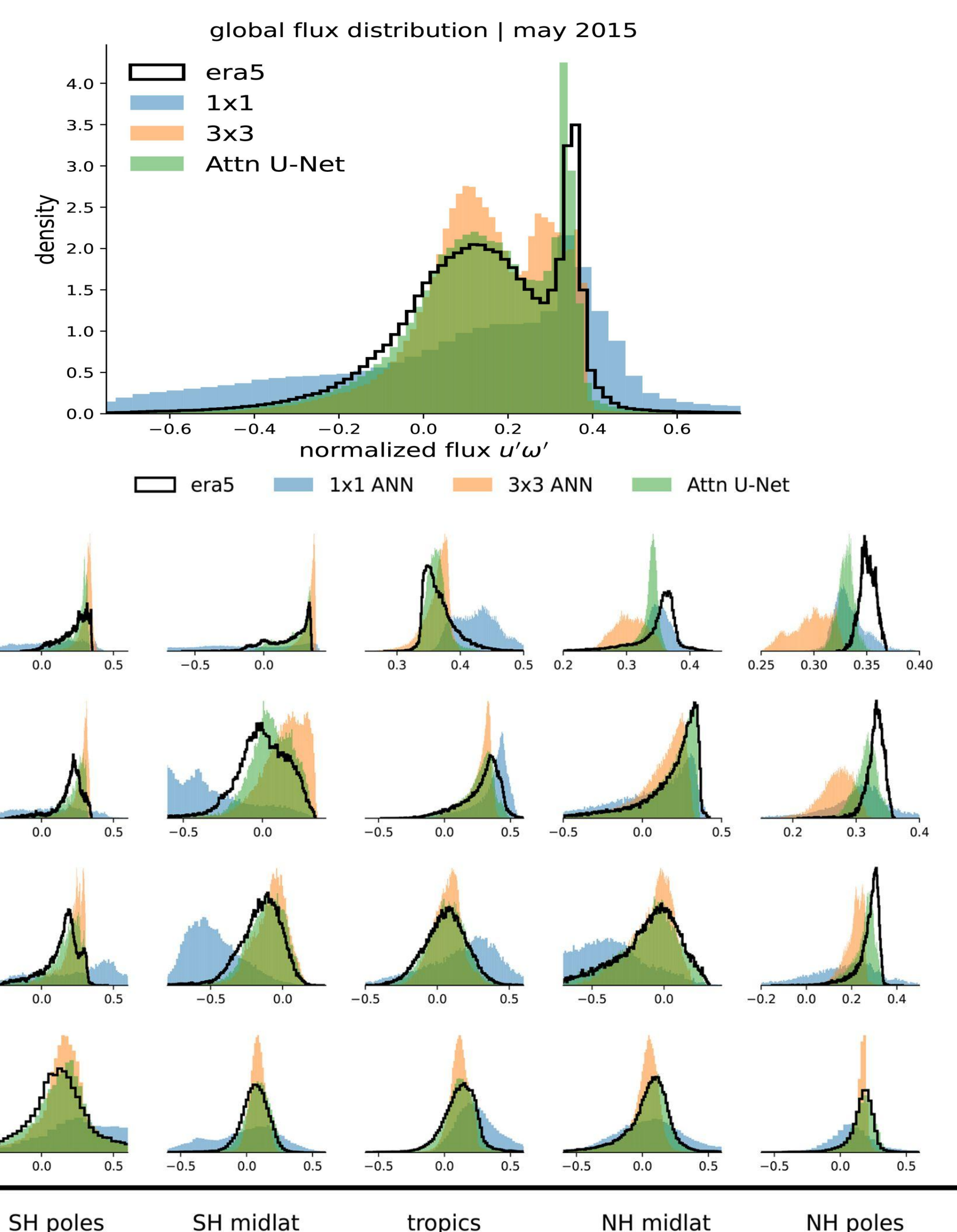


Fig 5: Distribution test. The Attention U-Net model captures the extremes in the gravity wave momentum flux better than other ANNs. The predicted distribution closely agrees with the true distribution for all latitudes and altitudes in the troposphere, for the Attention U-Net model. The 3x3 model performs reasonably well in the troposphere but has a stronger bias towards the mean and fails to capture the tail of the distribution effectively.

Key findings

- We show that data-driven, nonlocal ML models have the potential to transform climate model parameterization design that often ignore horizontal evolution of mesoscale processes.
- Prediction of resolved momentum flux greatly improves with the degree of nonlocality embedded in the ML architecture.
- The nonlocal models, especially the globally nonlocal Attention U-Net, generates accurate mean climatology instantaneous predictions, and commendably captures the global flux distribution.