

## Title: Machine Learning for Weather and Climate Modeling: Nonlocal Gravity Wave Parameterizations and Monsoon-time Benchmarking of AI Weather Models

### Abstract:

Global climate models run at 100–200 km resolution and rely on simplified parameterizations for unresolved processes such as atmospheric gravity waves (GWs). These approximations are a leading source of uncertainty in climate projections and contribute to persistent biases such as the cold-pole bias. We develop three neural networks, trained on multi-year global reanalysis, that predict GW momentum fluxes with increasing horizontal nonlocality: a single-column  $1\times 1$  ANN (M1), a  $3\times 3$  ANN-CNN (M2), and a global Attention UNet (M3). By encoding lateral wave propagation, which current single-column schemes omit, the fully nonlocal M3 best captures the transient evolution and seasonal distribution of fluxes over major GW hotspots. Because reanalysis under-resolves GWs, we refine the networks through transfer learning on a 1.4 km global forecast model, which corrects flux magnitudes while preserving what was learned from reanalysis. Ongoing efforts to couple these Python-based networks to Fortran climate models will be briefly discussed.

In a second line of work, MAUSAM, we evaluate seven leading AI weather models, including GraphCast, Pangu-Weather, Aurora, AIFS, and GenCast, against ground stations, rain gauges, and geostationary satellites during the South Asian Monsoon, rather than against reanalysis. The models are skillful at large scales, but their errors against observations are 15–45% larger than reanalysis-based scores suggest. They also show clear gaps in extreme rainfall, cyclone tracks, and the mesoscale energy spectrum, which limits their use for operational forecasting, monsoon precipitation prediction, and subseasonal-to-seasonal forecasting.