Data Assimilation Using Generative Adversarial Networks for Global Models

Otavio Feitosa*[1,2], Haroldo Velho*[2], Saulo Freitas[2], Juliana Anochi[2], and Angel Chovert[1,2]

[1]CEMPA Cerrado – Brazil
[2]Instituto Nacional de Pesquisas Espaciais – Brazil

Abstract

Data assimilation (DA) is a fusion data process for previous forecasting (background) and observations to compute the best initial condition (analysis) for a new prediction cycle. Our approach is different. The calculation of the initial condition for a prediction cycle is executed by using only observations. Spherical convolutions are applied to compute the atmospheric global analyses. Therefore, the developed approach creates initial conditions directly from observational data (satellite measurements, radar data, surface stations, and atmospheric soundings) without requiring a background field. The training and testing sets are data from the European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5), with one-degree resolution. The generative adversarial network (GAN) architecture operates on Hierarchical Equal Area isoLatitude Pixelization (HEALPix) grids and uses index-based convolutions with precomputed geodesic neighbor mappings. The convolution operations maintain the HEALPix grid's equal-area properties, where this approach shows O(N) computational complexity compared to O(N^2) with methods used in latitude-longitude grids. Our study compares a simple regression, with only Mean Squared Error (MSE) loss with a GAN implementation. Initial tests show the GAN model preserves energy conservation in wind fields better than the simple regression approach. The analyses were obtained for five global atmospheric variables (zonal wind, meridional wind, geopotential height, relative humidity, and temperature) on 13 vertical levels, for providing the initial condition for the Model for Prediction Across Scales-National Center for Atmospheric Research (MPAS-NCAR) model. Tests show capability in capturing large-scale and synoptic atmospheric phenomena. Spectral analysis confirms the GAN model produces states that adhere to the Kolmogorov "-5/3" scaling law, indicating proper energy distribution in wind fields. Preliminary tests show the MPAS global model operating with GAN-derived initial condition produces forecasts with error magnitudes comparable with those employing operational analyses used in the Global Forecast System / National Centers for Environmental Prediction (GFS/NCEP) when both are evaluated against ERA5 reanalysis data.

Keywords: Spherical Convolution, HEALPix, Neural Networks, Atmospheric Analysis, Generative Adversarial Networks, Observational Data Processing

*Speaker