PANGU-DART: ENABLING ENSEMBLE KALMAN FILTER DATA ASSIMILATION FOR ML-BASED NUMERICAL WEATHER PREDICTION MODELS

Nuo Chen* [1] and Helen Kershaw[2]

[1] *University of Oklahoma, Norman, Oklahoma*
[2] *NSF National Center for Atmospheric Research, Boulder, Colorado*

1. INTRODUCTION

Machine learning (ML)-based numerical weather prediction (NWP) models, such as Pangu-Weather (Bi et al. 2023), provide faster and more computationally efficient alternatives to traditional physics-based models. For example, the Pangu-Weather model generates a global forecast at 0.25° resolution in just 10 seconds on a single GPU, significantly reducing computational cost compared to traditional NWP models.

Ensemble Kalman Filter (EnKF) is a widely used statistical data assimilation (DA) method that improves forecast accuracy by assimilating real-time observations and updating ensemble forecasts. However, effective application of the EnKF requires a sufficient number of ensemble members (typically 30–100) to adequately capture the evolving forecast error covariance. The high computational cost of traditional NWP systems often limits ensemble size, making ML-based models an attractive alternative for cost-effective DA. To address these challenges, we introduce Pangu-DART, an interface between Pangu-Weather and the Data Assimilation Research Testbed (DART, Anderson et al. 2009), which enables ensemble-based DA for ML-based weather forecasting.

Several studies have attempted to accelerate global data assimilation using ML methods that resemble the traditional DA methods. AI-Var (Keller and Potthast 2024) is based on 3-dimensional variational (3D-Var) DA (Lorenc 1986). Other ML approaches (Hatfield et al. 2021; Farchi et al. 2021; Xiao et al. 2024) have been designed for 4D-Var (Rabier and Courtier 1992) and hybrid En4DVar framework (Liu et al. 2008) such as Fuxi-En4DVar (Li et al. 2024), which leverage the computational efficiency of AI models to generate ensemble members. Other recent studies have trained neural networks to perform the data assimilation itself (Chen et al. 2024).

Unlike most of the methods mentioned above, Pangu-DART does not attempt to replace traditional DA algorithms with ML-based ones. Instead, it leverages the DART framework to apply ensemble DA methods directly to any prediction model, even a black-box AI model.

(Slivinski et al. 2024) and (Kotsuki et al. 2024) have tested similar integrations, but their experiments were limited to assimilating only surface pressure or radiosonde data. In this study, we assimilate multiple observation types, including dropsonde profiles, aircraft winds, and global satellite winds.

* *Corresponding author address:* Nuo Chen, 120 David L Boren Blvd, Norman, OK 73072;
E-mail: nuo.chen-1@ou.edu.

2. METHODS

2.1. Pangu-DART Interface

We developed an interface between Pangu-Weather and DART to enable ensemble DA cycles (Fig.1). DART is a modular DA framework with a well-developed forward operator package that converts various observation types to the model space. The default assimilation algorithm is the Ensemble Adjustment Kalman Filter (EAKF; Anderson 2001).

The variables updated during DA are the upper-level model states, including horizontal wind components (U, V), temperature (T) and specific humidity (Q) at 13 pressure levels used in Pangu-Weather Model. The posterior model states serve as the initial condition for the next 6-hour forecast cycle. The first ensemble initial conditions use the ERA5 10-member ensemble dataset(Hersbach et al. 2020).

To support surface data assimilation in the future release, 0.25° global topographical data were generated using the Weather Research and Forecasting Preprocessing System (WPS, Skamarock et al. 2019). However, the current version of the interface only supports vertical interpolation in pressure coordinates.

Pangu-Weather forecasts were run on Google Colab with a single NVIDIA A100 GPU, using up to 34 GB of GPU RAM, while the DA process with DART was performed on the Casper HPC system of the National Center for Atmospheric Research (NCAR).

The observational datasets used in this study are retrieved from NCAR's Research Data Archive ds337.0 (2008) for radiosondes and aircraft observation, and ds351.0 (2004) for satellite-derived winds.

3. RESULTS

3.1. Case Study: Hurricane Milton

Hurricane Milton formed on October 5th, 2024, in the western Caribbean Sea and became extratropical on October 10th. It rapidly deepened by more than 100 hPa in two days over the Gulf of Mexico, becoming the second most intense Atlantic hurricane with a peak minimum sea level pressure of 897 hPa, second only to Hurricane Rita (2005).

We evulate the performance of Pangu-Weather model forecast and Pangu-DART analysis result in predicting Milton's track and intensity. The experiments start at 0000 UTC on October 6, 2024 with ERA5 10-member ensemble as the initial ensemble conditions and lasted for four days.

3.2. Pangu-DART Performance

We compare the 4-day Pangu-Weather forecast and the 4-day Pangu-DART analysis against the International Best Track Archive (IBTrACS). Figure 2a shows that assimilating upper-level observations corrected track errors, pulling the analysis closer to the observed best track

Figure 2b shows IBTrACS's minimum sea level pressure (SLP) and the maximum wind speed (V_{max}). Neither the Pangu-Weather forecast nor the Pangu-DART analysis fully captured Milton's intensity in terms of minimum SLP (e.g. Fig.2c) and V_{max} (e.g. Fig.2c). The mean SLPs of the forecast and analysis ensemble during the storm peak are greater than 990 hPa and the V_{max} values remain below 38 kts, far from the 897 hPa and 145 kts recorded in Best Track. The

The 10th-90th percentile range of the analysis ensemble (red shading in Fig.2c,d) suggests that the spread of the ensemble rapidly collapses after two DA cycles, highlighting the need for implementing appropriate prior inflation.

The 500 hPa temperature increment calculated as the difference before and after the assimilation increases with more DA cycles performed (e.g. Fig.3), with the largest increment emerging in the southern ocean. This pattern agrees with the surface increment analysis from Slivinski et al. (2024, their supplementary figure 1).

4. DISCUSSION AND FUTURE WORK

In this study, we implemented the Pangu-DART interface to perform data assimilation with a ML-based model. Aside from the success to improve the Hurricane track, the intensity forecast and analysis remain a known challenge for ML-based NWP as previously pointed out by Bonavita (2024). The study confirms that ML models often exhibit regression toward the mean, leading to limited ensemble spread that hampers their performance in ensemble-based DA frameworks. This limitation impairs their ability to accurately represent forecast uncertainty, which is critical for EnKF approaches.

While Slivinski et al. 2024 observed model divergence after 10 to 30 days when only assimilating the surface-level pressure values with the similar approach used in this study, other studies (Xu et al. 2024; Xiao et al. 2024; Farchi et al. 2021; Kotsuki et al. 2024) have demonstrated stable year-long DA cycle is achievable using ML-trained data assimilation methods or with Local Ensemble Transform Kalman Filter (LETKF) on a coarser grid ML-based forecast model.

For future development, we plan to implement prior inflation techniques to mitigate ensemble spread collapse and extend assimilation to surface variables.

5. CODE AVAILABILITY

GitHub repository used for ongoing code development is available at `https://github.com/NCAR/DART/tree/main/models/pangu`

Acknowledgments. This study was supported by NSF award 2141537 and the NCAR-DART team. Computational resources were provided by NCAR's Casper HPC and Google Colab.

REFERENCES

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The data assimilation research testbed: A community facility. *Bulletin of the American Meteorological Society*, **90 (9)**, 1283 – 1296, doi:10.1175/2009BAMS2618.1.

Anderson, J. L., 2001: An ensemble adjustment kalman filter for data assimilation. *Monthly Weather Review*, **129 (12)**, 2884 – 2903, doi:10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

Bi, K., L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian, 2023: Accurate medium-range global weather forecasting with 3d neural networks. *Nature*, **619 (7970)**, 533–538, doi:10.1038/s41586-023-06185-3, URL https://doi.org/10.1038/s41586-023-06185-3.

Bonavita, M., 2024: On some limitations of current machine learning weather prediction models. *Geophysical Research Letters*, **51 (12)**, e2023GL107 377, doi:https://doi.org/10.1029/2023GL107377, e2023GL107377 2023GL107377.

Chen, K., and Coauthors, 2024: Towards an end-to-end artificial intelligence driven global weather forecasting system. URL https://arxiv.org/abs/2312.12462, 2312.12462.

ds337.0, 2008: Ncep adp global upper air and surface weather observations (prepbufr format). Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder CO, URL https://rda.ucar.edu/datasets/dsd337000/.

ds351.0, 2004: Ncep adp eta / nam upper air observation subsets. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder CO, URL https://rda.ucar.edu/datasets/dsd351001/.

Farchi, A., P. Laloyaux, M. Bonavita, and M. Bocquet, 2021: Using machine learning to correct model error in data assimilation and forecast applications. *Quarterly Journal of the Royal Meteorological Society*, **147 (739)**, 3067–3084, doi:10.1002/qj.4116.

Hatfield, S., M. Chantry, P. Dueben, P. Lopez, A. Geer, and T. Palmer, 2021: Building tangent-linear and adjoint models for data assimilation with neural networks. *Journal of Advances in Modeling Earth Systems*, **13 (9)**, e2021MS002 521, doi:https://doi.org/10.1029/2021MS002521.

Hersbach, H., and Coauthors, 2020: The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 1999–2049, doi:https://doi.org/10.1002/qj.3803.

Keller, J. D., and R. Potthast, 2024: Ai-based data assimilation: Learning the functional of analysis estimation. URL https://arxiv.org/abs/2406.00390, 2406.00390.

Kotsuki, S., K. Shiraishi, and A. Okazaki, 2024: Ensemble data assimilation to diagnose ai-based weather prediction model: A case with climax version 0.3.1. URL https://arxiv.org/abs/2407.17781, 2407.17781.

Li, Y., and Coauthors, 2024: Fuxi-en4dvar: An assimilation system based on machine learning weather forecasting model ensuring physical constraints. *Geophysical Research Letters*, **51 (22)**, e2024GL111 136, doi:https://doi.org/10.1029/2024GL111136.

Liu, C., Q. Xiao, and B. Wang, 2008: An ensemble-based four-dimensional variational data assimilation scheme. part i: Technical formulation and preliminary test. *Monthly Weather Review*, **136 (9)**, 3363 – 3373, doi:10.1175/2008MWR2312.1.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, **112 (474)**, 1177–1194, doi:https://doi.org/10.1002/qj.49711247414.

Rabier, F., and P. Courtier, 1992: Four-dimensional assimilation in the presence of baroclinic instability. *Quarterly Journal of the Royal Meteorological Society*, **118**, 649–672, doi:https://doi.org/10.1002/qj.49711850604.

Skamarock, W. C., and Coauthors, 2019: A description of the advanced research wrf version 4. NCAR Tech. Note NCAR/TN-556+STR, NCAR, 145 pp. doi:10.5065/1dfh-6p97.

Slivinski, L. C., J. S. Whitaker, S. Frolov, T. A. Smith, and N. Agarwal, 2024: Assimilating observed surface pressure into ml weather prediction models. URL https://arxiv.org/abs/2412.18016, 2412.18016.

Xiao, Y., L. Bai, W. Xue, K. Chen, T. Han, and W. Ouyang, 2024: Fengwu-4dvar: Coupling the data-driven weather forecasting model with 4d variational assimilation. URL https://arxiv.org/abs/2312.12455, 2312.12455.

Xu, X., X. Sun, W. Han, X. Zhong, L. Chen, and H. Li, 2024: Fuxi-da: A generalized deep learning data assimilation framework for assimilating satellite observations. URL https://arxiv.org/abs/2404.08522, 2404.08522.

6. FIGURES

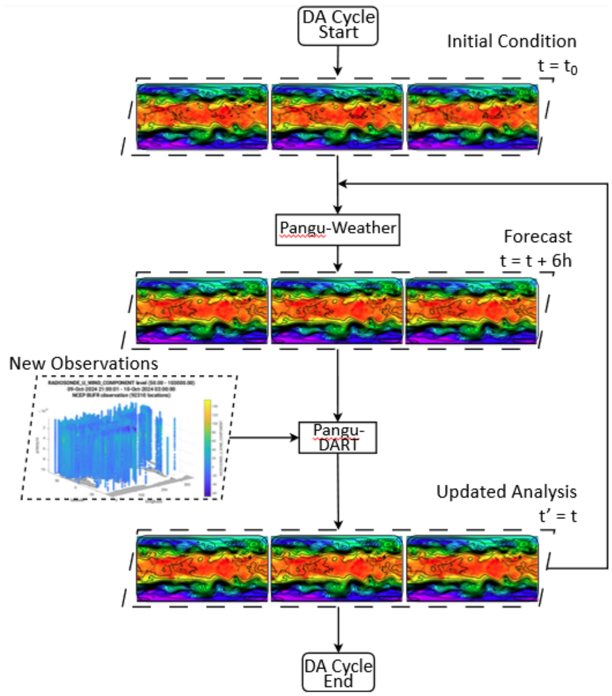

FIG. 1. Pangu-DART flowchart.

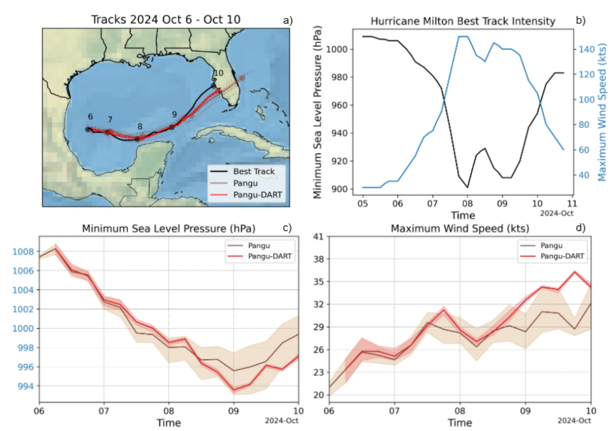

FIG. 2. a) Tracks of Hurricane Milton (0000 UTC 6 Oct 2024 to 0000 UTC 10 Oct 2024) from IBTrACS (International Best Track Archive for Climate Stewardship, black), Pangu-Weather forecast ensemble mean (brown), and Pangu-DART analysis ensemble mean (red). Individual ensemble members are shown with dotted lines; b) Milton's minimum sea level pressure (black) and maximum wind speed (blue) from IBTrACs; c) Milton's minimum sea level pressure from Pangu-Weather forecast ensemble mean (brown), and Pangu-DART analysis ensemble mean (red). Shading area shows range from 10th to 90th ensemble percentiles; d) Same as c) but for maximum wind speed.

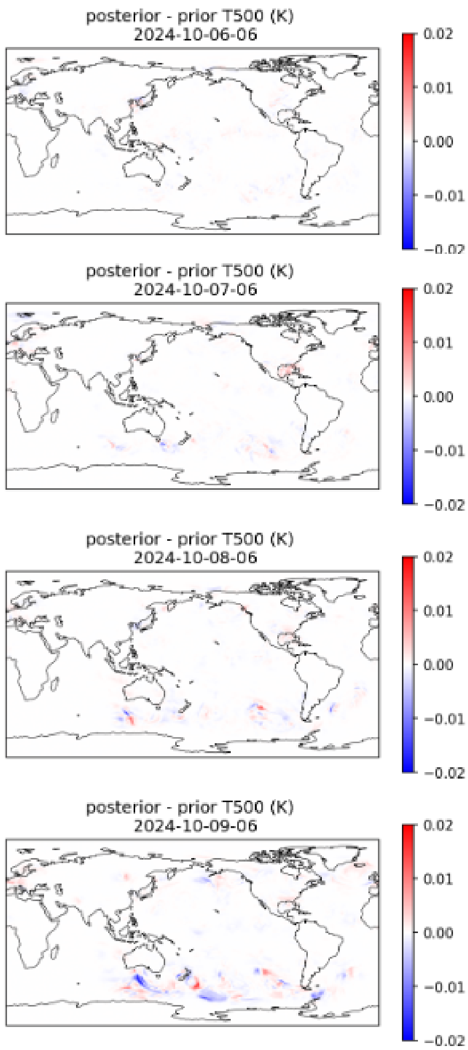

FIG. 3. Temperature increment (K) at 500 hPa added by data assimilation from 0600 UTC 6 Oct 2024 to 0600 UTC 9 Oct 2024.