Deterministic Atmospheric Baselines and Cryptographic Provenance in AI Weather Forecasting: The OD-EARTH2-001 Study for EO-OBSCURA

Zhane Daye
Origin Datum (datum.space)
GitHub: github.com/zrt219
Program on Incomplete-Information Earth Observation (EO-OBSCURA)
Technical Report OD-EARTH2-001-REV2

Abstract—Deep learning atmospheric emulators require rigorous empirical baselines, mathematically sound spatial error metrics, and verifiable cryptographic data provenance. In this paper, we present the architecture and experimental validation of the OD-EARTH2-001 research foundation within NVIDIA's Earth2Studio framework. Operating under strict zero-cost constraints on commodity CPU hardware, we establish an end-to-end meteorological pipeline ingesting real-world 0.25-degree Global Forecast System (GFS) observations from public AWS Open Data. Atmospheric predictions are generated via a deterministic Persistence baseline model, stored in multidimensional Zarr array hierarchies, and evaluated against subsequent GFS analysis states at a 6-hour lead time. We formulate area-weighted spherical error metrics that eliminate polar grid convergence singularities and demonstrate that a 6-hour persistence forecast achieves global latitude-weighted Mean Absolute Errors of $1.2635\,\mathrm{K}$ for 2-metre temperature, $1.4975\,\mathrm{m\,s^{-1}}$ for zonal wind, and $1.6118\,\mathrm{m\,s^{-1}}$ for meridional wind. Furthermore, we develop an isolated, secret-scrubbed client for remote NVIDIA Cloud Functions, verifying authenticated access to the FourCastNet service. All execution parameters, array tensors, and validation reports are cryptographically anchored using SHA-256 manifests. This work establishes the thermodynamic and kinematic baseline for the EO-OBSCURA program, constraining multimodal satellite imagery reconstruction through deterministic meteorological priors.

Index Terms—Atmospheric modeling, Earth2Studio, Global Forecast System, spherical error metrics, cryptographic provenance, EO-OBSCURA, weather forecasting.

1 Introduction

Accurate meteorological modeling forms the cornerstone of modern planetary stewardship, agriculture, hazard mitigation, and Earth observation analytics. In recent years, data-driven machine learning weather models—such as FourCastNet [1], GraphCast [2], and Pangu-Weather [3]—have demonstrated forecast skills rivaling traditional Numerical Weather Prediction (NWP) systems at fractionally lower operational inference costs.

However, the rapid adoption of AI weather modeling has exposed recurring engineering vulnerabilities:

1) **Metric Inconsistency**: Evaluating equiangular latitude-longitude grids without spherical area weighting introduces substantial polar bias, artificially suppressing or exaggerating error metrics.
2) **Data Provenance Gaps**: Incomplete tracking of initial condition cycles, model weight hashes, and execution artifacts impairs scientific reproducibility.
3) **Synthetic vs. Empirical Conflation**: Pipeline smoke tests utilizing synthetic random noise are frequently confused with meteorologically verified forecasts.

The **EO-OBSCURA** program at Origin Datum addresses incomplete, cloud-obscured, and sparse satellite observations by reconstructing multi-spectral surface information using physical atmospheric priors. Before coupling high-capacity neural emulators into the reconstruction pipeline, a deterministic, verifiable, and mathematically sound baseline must be established.

This paper reports the implementation and empirical validation of **OD-EARTH2-001**, an open-source, zero-cost atmospheric forecasting and verification architecture built upon NVIDIA Earth2Studio [4].

2 Mathematical Formulation

2.1 Spherical Coordinate Discretization

Let the continuous atmospheric state over the spherical manifold $\mathcal{S}^2$ be discretized on a regular equiangular latitude-longitude grid:

$$\Omega = \{(\phi_i, \lambda_j) \mid 1 \le i \le N_\phi,\ 1 \le j \le N_\lambda\}, \tag{1}$$

where $\phi_i \in [-\pi/2, \pi/2]$ denotes geodetic latitude and $\lambda_j \in [0, 2\pi)$ denotes geodetic longitude. For a 0.25-degree resolution grid:

$$N_\phi = 721, \quad N_\lambda = 1440, \quad |\Omega| = 1,038,240 \text{ cells}. \tag{2}$$

The infinitesimal surface area element on a unit sphere is given by:

$$dA = \cos(\phi)\, d\phi\, d\lambda. \tag{3}$$

Consequently, discrete grid cells subtend an area proportional to $\cos(\phi_i)$. Cells near the poles ($\phi \to \pm\pi/2$) approach zero physical surface area despite occupying equal representation in Euclidean array representations.

2.2 Latitude-Weighted Error Metrics

To ensure metric convergence to true continuous L^1 and L^2 norms over the spherical surface, we define non-negative spatial weights:

$$w_i = \max(0, \cos(\phi_i)). \tag{4}$$

Let $\hat{X}(\phi_i, \lambda_j)$ represent the forecasted meteorological field at lead time Δt, and let $X(\phi_i, \lambda_j)$ denote the true verification reference analysis at the identical valid time. The latitude-weighted Mean Absolute Error (MAE_{lat}) is formulated as:

$$\text{MAE}_{\text{lat}} = \frac{\sum_{i=1}^{N_\phi} \sum_{j=1}^{N_\lambda} w_i \cdot \mathbf{1}_{\text{valid}}(i,j) \cdot |\hat{X}_{i,j} - X_{i,j}|}{\sum_{i=1}^{N_\phi} \sum_{j=1}^{N_\lambda} w_i \cdot \mathbf{1}_{\text{valid}}(i,j)}, \tag{5}$$

where $\mathbf{1}_{\text{valid}}(i,j) \in \{0,1\}$ filters missing values and non-finite numbers (NaN/Inf).

Similarly, the latitude-weighted Root Mean Squared Error (RMSE_{lat}) and Mean Bias (MB_{lat}) are defined as:

$$\text{RMSE}_{\text{lat}} = \sqrt{\frac{\sum_{i=1}^{N_\phi} \sum_{j=1}^{N_\lambda} w_i \cdot \mathbf{1}_{\text{valid}}(i,j) \cdot (\hat{X}_{i,j} - X_{i,j})^2}{\sum_{i=1}^{N_\phi} \sum_{j=1}^{N_\lambda} w_i \cdot \mathbf{1}_{\text{valid}}(i,j)}}, \tag{6}$$

$$\text{MB}_{\text{lat}} = \frac{\sum_{i=1}^{N_\phi} \sum_{j=1}^{N_\lambda} w_i \cdot \mathbf{1}_{\text{valid}}(i,j) \cdot (\hat{X}_{i,j} - X_{i,j})}{\sum_{i=1}^{N_\phi} \sum_{j=1}^{N_\lambda} w_i \cdot \mathbf{1}_{\text{valid}}(i,j)}. \tag{7}$$

2.3 Persistence Baseline Dynamics

The Eulerian persistence prognostic operator $\mathcal{P}_{\Delta t}$ freezes the initial atmospheric field:

$$\hat{X}(t_0 + \Delta t) = \mathcal{P}_{\Delta t}[X(t_0)] \equiv X(t_0). \tag{8}$$

For short lead times ($\Delta t = 6\,\text{h}$), the persistence forecast captures stationary synoptic patterns but incurs errors driven by high-frequency diurnal thermal cycles and advective wind displacement. It represents the strictly minimal skill threshold that any neural forecasting model must surpass to demonstrate physical learning.

3 System Architecture

The OD-EARTH2-001 research environment is designed as a standalone, modular package isolated within the Earth2Studio fork:

- **Data Ingestion Layer**: Connects to public NOAA GFS GRIB2 analyses via anonymous S3 transport ('noaa-gfs-bdp-pds'), incurring zero access or ingress fees.
- **Prognostic Execution Layer**: Orchestrates model time-stepping using Earth2Studio's deterministic pipeline, directing outputs into chunked Zarr stores.
- **Array Extraction & Validation Layer**: Decouples model output from storage artifacts, evaluating spherical metrics directly against verification cycles.
- **Security & Scrubber Engine**: Sanitizes API tokens, AWS keys, and confidential environment variables across all diagnostic logs and JSON records.
- **Remote NVIDIA API Client**: Implements an asynchronous-ready, retry-protected client with exponential backoff for remote FourCastNet NIM inference.

4 Empirical Validation & Results

4.1 Experimental Configuration

The system was initialized using the GFS operational cycle at $t_0 =$ 2024-01-01 00:00:00 UTC. Forecast arrays were generated for a $\Delta t = 6\,\text{h}$ horizon ($t_v =$ 2024-01-01 06:00:00 UTC). Three fundamental surface prognostic variables were extracted:

- $T_{2\text{m}}$: Surface air temperature at 2 metres (K).
- $u_{10\text{m}}$: Zonal wind velocity at 10 metres ($\text{m}\,\text{s}^{-1}$).
- $v_{10\text{m}}$: Meridional wind velocity at 10 metres ($\text{m}\,\text{s}^{-1}$).

4.2 Quantitative Results

Table 1 reports the resulting latitude-weighted error statistics computed over all $1,038,240$ global grid nodes.

TABLE 1
Global Latitude-Weighted Metric Verification ($+6$h Forecast)

Variable	MAE_{lat}	RMSE_{lat}	MB_{lat}	Physical Scale
$T_{2\text{m}}$	**1.2635**	2.3813	-0.3313	K
$u_{10\text{m}}$	**1.4975**	2.4188	-0.0341	$\text{m}\,\text{s}^{-1}$
$v_{10\text{m}}$	**1.6118**	2.5368	+0.0719	$\text{m}\,\text{s}^{-1}$

4.3 Meteorological Analysis

The computed values align precisely with physical expectations for 6-hour atmospheric changes:

1) **Diurnal Thermal Cycling**: The $T_{2\text{m}}$ mean bias of $-0.3313\,\text{K}$ captures the nocturnal cooling phase across large continental landmasses in the Northern Hemisphere winter.
2) **Momentum Conservation**: Zonal wind bias remains negligible ($-0.0341\,\text{m}\,\text{s}^{-1}$), demonstrating global angular momentum balance over short lead times.
3) **Baroclinic Wave Advection**: Wind velocity RMSE values ($\approx 2.4\,\text{m}\,\text{s}^{-1}$) reflect the propagation speed of mid-latitude jet streaks and cyclone frontal boundaries.

4.4 Remote NVIDIA Cloud API Verification

To establish cloud-accelerated inference capability for models exceeding local hardware capacity (e.g. FourCastNet requiring $\geq 16\,\text{GB}$ VRAM), the remote client was tested against NVIDIA Cloud Functions (NVCF). Minimal, non-destructive health probes demonstrated:

- Build Catalog Probe: HTTP 200 OK (175.07 ms latency, 85 models available).
- NVCF Service Probe: HTTP 200 OK (292.38 ms latency, 237 functions detected).
- FourCastNet Function Status: Active ('id: 5bdd622c-14f1-4783-854c-d9d310e4fab0').

4.5 Automated Regression Suite

An automated test suite comprising **55 comprehensive tests** was executed in the WSL2 Ubuntu virtual environment. All 55 tests passed cleanly (100% pass rate), verifying metric broadcasting, NaN handling, secret redaction, and API mock resilience.

5 Cryptographic Provenance

Every forecast run emits a machine-readable JSON manifest containing SHA-256 digests for all generated artifacts. Table 2 details the cryptographic manifest generated for the primary validation run.

TABLE 2
Cryptographic Artifact Attestation

Artifact	SHA-256 Checksum	Size
`forecast_t2m.npy`	`e63a97e30e71a98a3510...`	4.15 MB
`forecast_t2m.png`	`019fa67b79ff4ceb126e...`	534 KB
`reference_t2m.npy`	`deddf7f7ee0d13afbcb5...`	8.31 MB
`validation_report.json`	`d667443282ed358901ba...`	451 B

6 Conclusion and Future Work

The OD-EARTH2-001 project has successfully established a scientifically sound, reproducible research environment for AI weather forecasting within NVIDIA Earth2Studio. By enforcing spherical area-weighted error calculations, isolating credentials, and executing genuine public data workflows, we have eliminated synthetic verification artifacts and proven operational stability on zero-cost commodity hardware.

This foundation enables immediate progression toward the primary objectives of the **EO-OBSCURA** program:

1) Deploying local DLWP inference on CPU to benchmark neural error suppression against the persistence baseline.
2) Integrating authenticated FourCastNet remote inference via the verified NVCF client.
3) Conditioning Sentinel-1 SAR-to-optical translation models on Earth-2 boundary layer wind and moisture fields to reconstruct cloud-obscured surface features.

References

[1] J. Pathak, S. Subramanian, P. Harrington, S. Raja, A. Chattopadhyay, M. Mardani, M. Kurth, D. Hall, Z. Li, K. Azizzadenesheli *et al.*, "FourCastNet: A global data-driven high-resolution weather model using adaptive Fourier neural operators," *arXiv preprint arXiv:2202.11214*, 2022.

[2] R. Lam, A. Sanchez-Gonzalez, M. Willson, P. Wirnsberger, M. Fortunato, F. Alet, S. Ravuri, T. Ewalds, F. Eaton, G. Broomhead *et al.*, "Learning skillful medium-range global weather forecasting," *Science*, vol. 382, no. 6677, pp. 1416–1421, 2023.

[3] K. Bi, L. Xie, H. Zhang, X. Chen, X. Gu, and Q. Tian, "Accurate medium-range global weather forecasting with 3D neural networks," *Nature*, vol. 619, no. 7970, pp. 533–538, 2023.

[4] NVIDIA PhysicsNeMo Team, "Earth2Studio: Open-source deep-learning framework for AI weather and climate workflows," *NVIDIA Corporation*, 2024. [Online]. Available: https://github.com/NVIDIA/earth2studio