MIT ARCLab Prize for AI Innovation in Space 2025 Phase 1 – Model Description DRAFT

Atilla Saadat

April 2025

Abstract

Accurate thermospheric density forecasts are essential for orbit-determination tasks such as collision avoidance and re-entry prediction. In this paper, we present a lightweight, physics-aware deep-learning model that attained a public weighted score of 0.5675 on the STORM-AI Challenge – Phase 1 leaderboard, surpassing empirical baselines while sustaining real-time inference on off-the-shelf GPUs. The approach combines (a) rigorous data curation from OMNI-2 and precise-orbit-determination datasets, (b) residual learning against an Orekit-driven NRLMSISE prior, and (c) an attention-pooled bidirectional GRU equipped with a time-weighted loss that mirrors the competition metric.

1 Introduction

Empirical models such as NRLMSISEprovide first-order estimates of atmospheric drag but underperform during geomagnetic storms. Recent studies demonstrate that machine learning (ML) trained on precise orbit densities can close this gap [1]. Here, we advance those findings with an ML architecture that learns only the *residual* to NRLMSISE, yielding robustness and faster convergence.

2 Data Pipeline

Table 1 summarises the end-to-end preprocessing (60-day history, 1h cadence). Hourly OMNI-2 indices are cleaned, gap-filled, and extended with short-term means and lags. Static satellite descriptors—Keplerian elements, geodetic state, and six cyclical encodings—augment each record. Files with fewer than 432 valid truth samples (3 days at 10 minute intervals) are discarded; longer series are truncated.

Table 1: End-to-end preprocessing workflow.

Stage	**Action**
Ingestion	Stream OMNI-2, Orekit-propagated NRLMSISE densities, and POD truth.
Cleaning	Replace sentinels, resample to 1h, bidirectional interpolate gaps.
Feature engineering	11 base indices, running means (1 h/3 h), lags {1,2,3,4,6,12} h; six orbital and six cyclical static terms; persistence NRLMSISE prior.
Length filter	Keep files with exactly 432 valid truth points.
Scaling	`QuantileTransformer` (inputs) and `StandardScaler` (log-residual targets).

2.1 Physics Prior

The NRLMSISE baseline used for residual normalization was generated using a high-fidelity numerical orbit propagator built on the Orekit astrodynamics library. For each file, a 3-day trajectory was propagated at 10-minute intervals using initial Keplerian elements, aligning timestamps with space weather data. The densities along this trajectory were evaluated using a custom `MSISPersistenceAtmosphere` class that combines the physics of atmospheric drag with neural inference of thermospheric density.

The MSIS driver includes a key preprocessing step for solar activity: for each timestamp, the F10.7 index and a 7-element `Ap` vector were computed. This vector includes the current Ap, the previous 4 values at 3-hour intervals, and two 24-hour means for the prior 12–33h and 36–57h intervals:

$$\texttt{Ap} = [A_p(t), A_p(t{-}3h), \ldots, \overline{A_p(12{:}33h)}, \overline{A_p(36{:}57h)}]$$

This construction mimics the original inputs expected by MSIS-type models and enhances stability during geomagnetic disturbances.

Propagation was performed using Orekit's high-fidelity propagator `DormandPrince853Integrator`, configured with tolerances of 10^{-3} m, time steps from 10^{-6} to 1000 seconds, and a fixed initial step size of 10 seconds. The force model includes Earth gravity (Holmes–Featherstone), atmospheric drag, and optional slots for third-body and SRP effects. Geodetic coordinates (lat/lon/alt) are derived from ITRF transformations of each propagated state to compute the appropriate density.

Each resulting dataset—containing synchronized (lat, lon, alt, density) vectors—was saved to directories and used to compute the training targets via:

$$y = \log_{10}\left(\max\left[\rho_{\text{true}}/\rho_{\text{MSIS}},\, 10^{-14}\right]\right)$$

This baseline is critical for aligning the model's objective with physical priors, improving both convergence and generalization across geomagnetic conditions.

3 Model Architecture

A three-layer bidirectional GRU ($h = 384$) encodes the $T = 1440$-step sequence. Scalar soft-attention pools the hidden states to a context vector (Table 2), which feeds a two-layer MLP to predict the full 432-step horizon in a single forward pass.

Table 2: Neural-network configuration.

Block	**Specification**
Encoder	3×Bi-GRU, h=384, dropout 0.3
Pooling	Soft-attention over 1440 timesteps
Head	LayerNorm → FC(384) → GELU → Dropout 0.3 → FC(432)
Output	432 scaled log-residuals (Δt = 10min, 3days)

4 Training Methodology

Loss. A time-weighted MSE $\mathcal{L} = N^{-1}\sum_t w_t(\hat{y}_t - y_t)^2$ uses weights $w_t = e^{-\gamma t}$ (γ chosen s.t. $w_T = 10^{-5}$) to mirror the OD-RMSE leaderboard metric.

Optimisation. AdamW was employed ($\eta = 3 \times 10^{-4}$, weight-decay 1×10^{-4}), linearly clipped gradients ($\|g\|_2 \leq 1$), `ReduceLROnPlateau` (patience 4, factor 0.5, floor 3×10^{-6}), mixed precision (AMP), and early stopping with patience 16.

Hyper-parameters. An 80/20 stratified split, batch size 64, and an epoch cap 120 yield convergence within ≤ 100 epochs on an RTX 3090 Ti.

5 Results

Table 3 lists the official Phase-1 public-leaderboard outcomes. Inference for a single 3-day forecast takes ≈10 ms.

Table 3: Public leaderboard scores (Phase 1).

	P1.1 Medium	**P1.2 Hard**	**Runtime P1.1 (s)**	**Runtime P1.2 (s)**
OD-RMSE ↑	0.6784	0.5398	677.95	322.56
Weighted score			**0.5675**	

6 Discussion

Table 4: Key innovations and their empirical benefit.

Technique	**Observed impact**
Physics-aware residual learning	Removes global bias; sharpens storm-time skill.
Attention pooling	Models multiday dependencies without quadratic Transformer cost.
432-sample filter	Eliminates label sparsity; stabilises training loss.
Metric-matched loss	Ensures monotonic link between training loss and OD-RMSE.
Geodetic fallback	Prevents NaNs from malformed LLA records.
Orekit + NRLMSISE prior	Supplies physically consistent normalisation.

7 Conclusion

The model was able to demonstrate that coupling a lightweight attention-pooled RNN with a physics-based prior delivers state-of-the-art OD-RMSE skill within strict runtime limits. Future work will explore transformer-free long-context encoders and domain-adaptation to higher-altitude regimes.

Reproducibility

All code is contained in a single, platform-agnostic file `submission.py`. The random seed is fixed (`random_state=42`), and data splits are deterministic.

References

[1] Giacomo Acciarini, Edward Brown, Tom Berger, Madhulika Guhathakurta, James Parr, Christopher Bridges, and Atılım Güneş Baydin. Improving thermospheric density predictions in low-earth orbit with machine learning. *Space Weather*, 22(2):e2023SW003652, 2024.