



Charles Sturt
University



The Data-Schema Bottleneck: Benchmarking Sixteen Deep Learning Architectures for Real-Time Starlink LEO Telemetry Management

Tanmoy Debnath (tdebnath@csu.edu.au)

Sourabhi Debnath (sdebnath@csu.edu.au)

Miroslaw Narbutt (miroslaw.narbutt@tudublin.ie)

Maumita Bhattacharya (mbhattacharya@csu.edu.au)

May 31, 2026



Presentation Outline

- Introduction
- Methodology
- Results
- Limitations and Future Directions
- Conclusion



Introduction: The Rise of LEO Satellite Networks

Non-terrestrial networks require predictive routing at an unprecedented scale.



The Goal:

Real-time data routing, predictive handover triggering, and dynamic beam allocation.

The Requirement:

Accurate forecasting of link-layer quality metrics (RTT, throughput, jitter) from continuous, high-rate historical streams.



Introduction: Orbital Mechanics & Telemetry Dynamics

Where does the predictive bottleneck actually reside?



Algorithmic Capacity

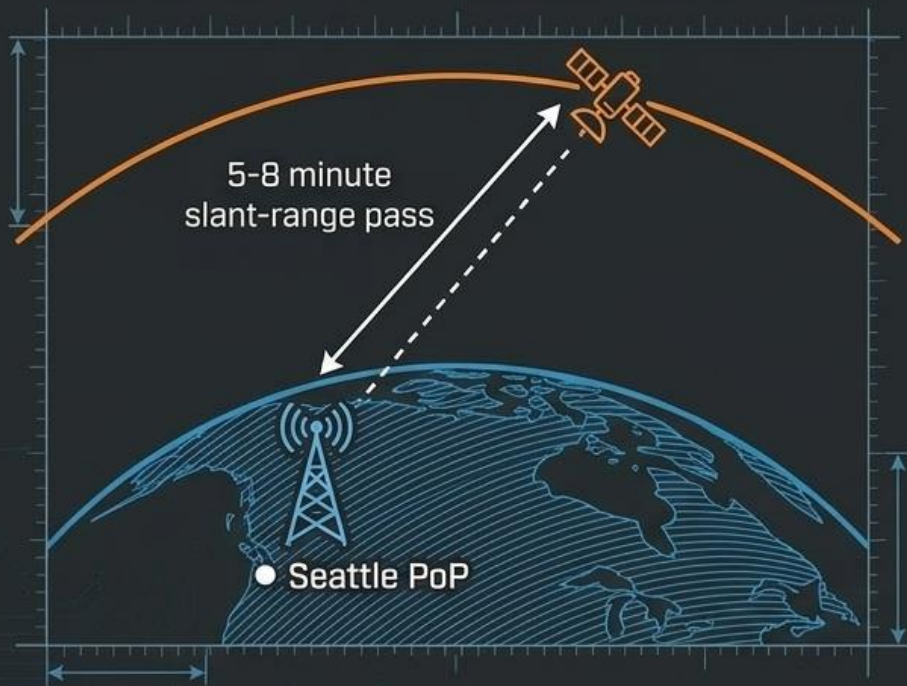
Hypothesis A: We need better algorithms. Current forecasting limits are imposed by the capacity of contemporary deep learning architectures.

Informational Boundaries

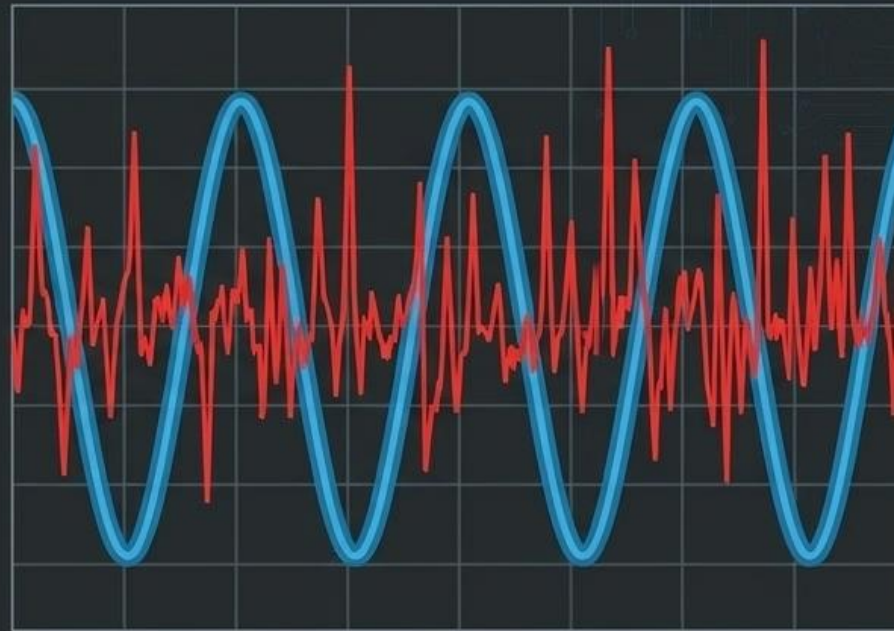
Hypothesis B: We need better data. Forecasting limits are intrinsic to the informational boundaries of end-to-end telemetry itself.

Introduction: Orbital Mechanics & Telemetry Dynamics

The Physics



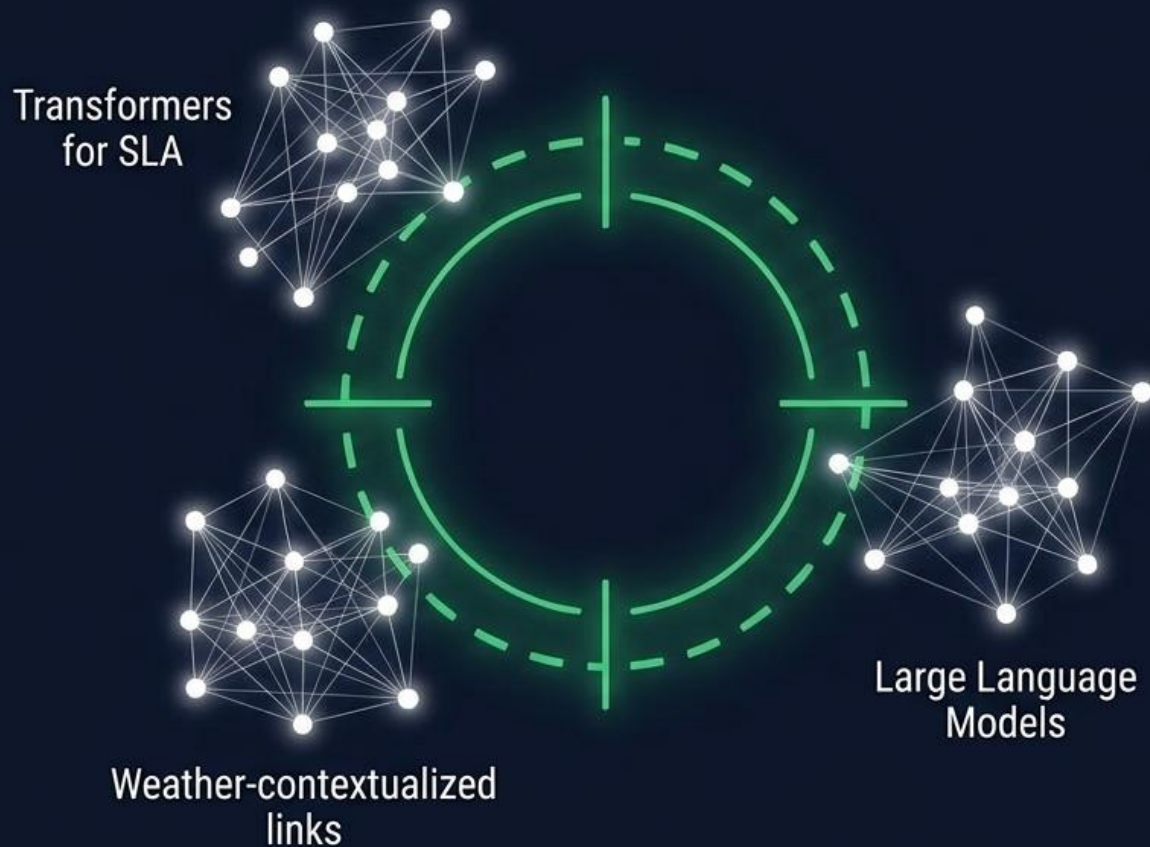
The Data



LEO telemetry has a dual character: physically deterministic orbital waves corrupted by highly stochastic network noise.

Introduction: Gaps in Current Research

Previous deep learning studies lack controlled, architecture-spanning benchmarks.



Fragmented Evaluations

Prior work evaluates narrow subsets of models under heterogeneous dataset splits.

The Missing Link

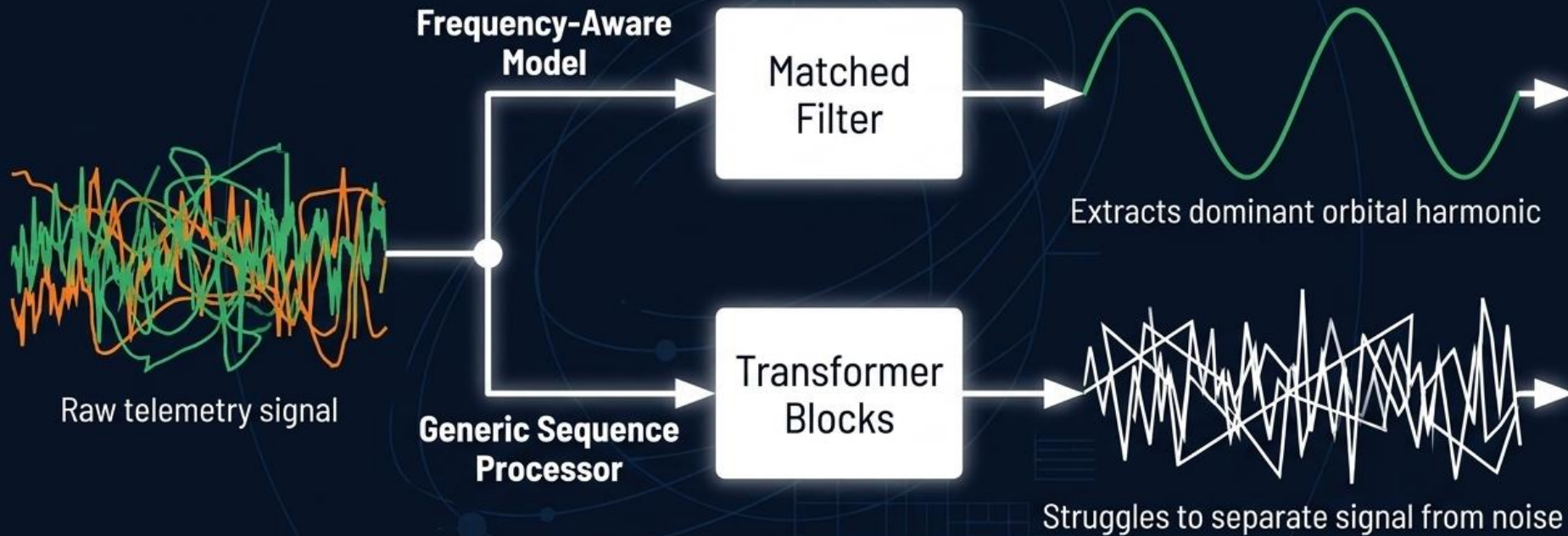
No existing study explicitly defines the empirical predictability ceiling for end-to-end LEO telemetry.

The Operational Risk

Operations teams are investing in larger AI models without knowing if a hard informational ceiling exists.

Introduction: Two Core Hypotheses

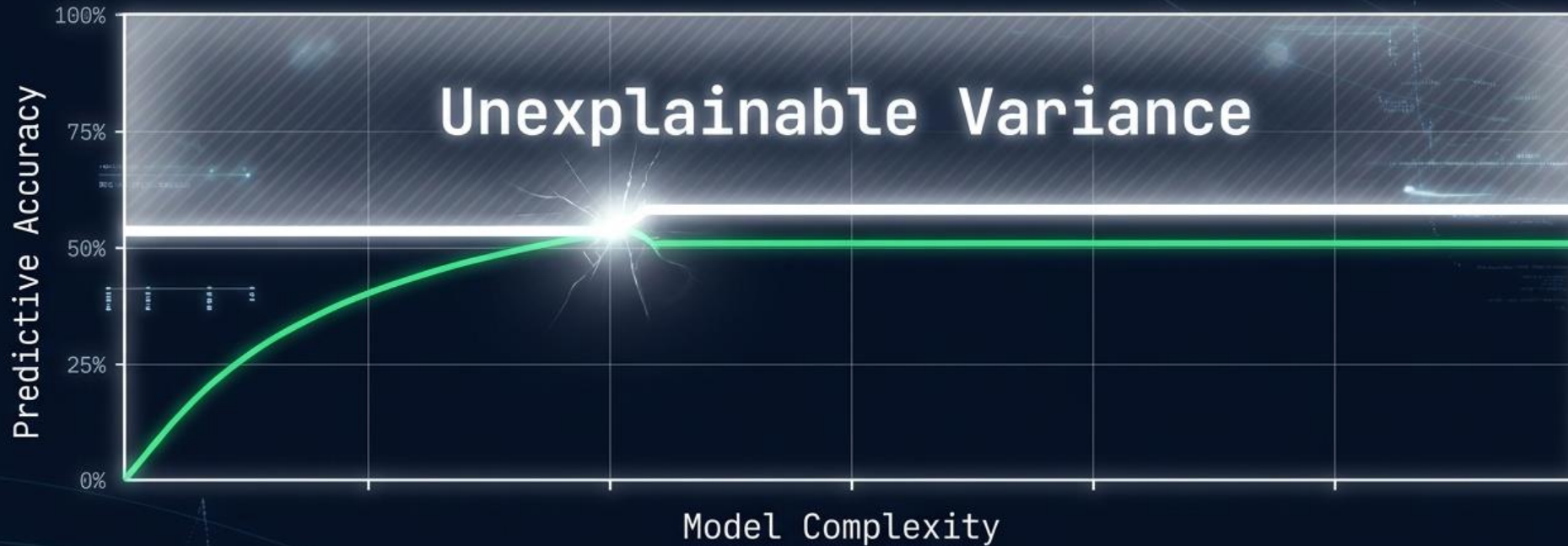
Hypothesis 1: The Spectral Alignment Advantage



Research Question 1 (RQ1): Do architectures with inductive biases matched to the quasi-periodic orbital dynamics of LEO systematically outperform generic sequential processors?

Introduction: Two Core Hypotheses *contd.*

Hypothesis 2: The Predictability Ceiling



Research Question 2 (RQ2): Does an architecture-invariant empirical upper bound on **explainable variance** exist for end-to-end temporal telemetry?

If this ceiling exists, what does it imply for the future design of **AI-assisted edge databases**?

Methodology: The LENS Dataset

The Proving Ground: 30 days of continuous streaming Starlink telemetry

1. Data Source:

LENS Dataset (Starlink rev2_proto3 dish)

2. Raw Scale:

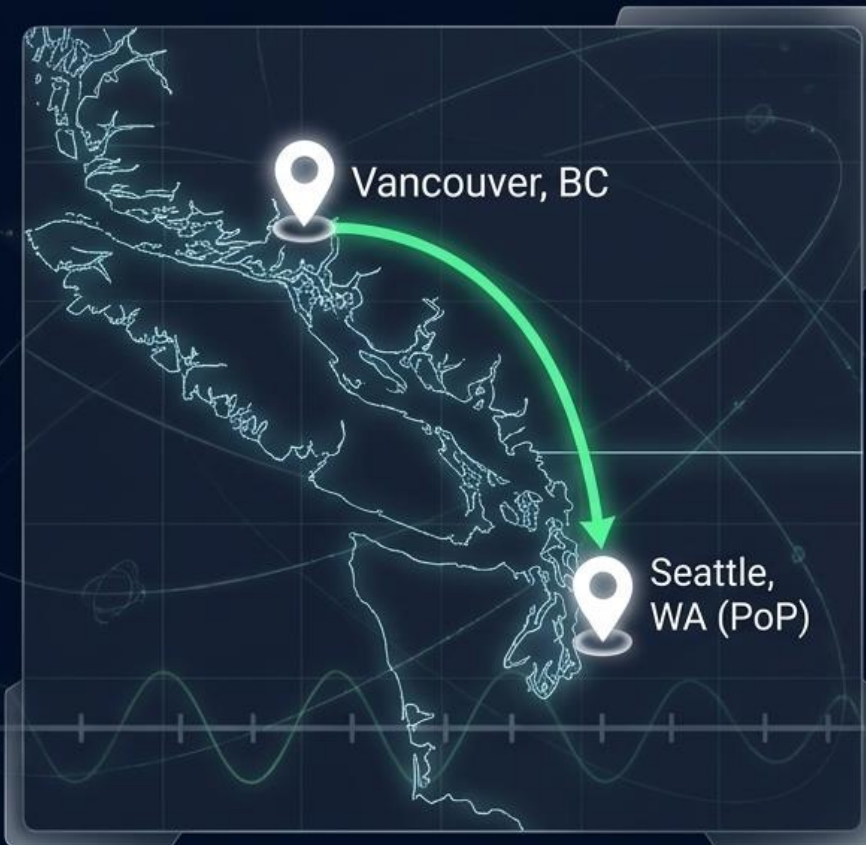
417 million packet-level observations
(~160Hz)

3. Preprocessing:

Deterministic non-overlapping mean
resampling to 5-second intervals

4. Processed Corpus:

518,450 rows capturing true
orbital-period-scale periodicity while
suppressing irreducible micro-jitter



Methodology: Feature Engineering & Preprocessing

Formulating the real-time link quality prediction task



Supervised Targets	Evaluation Metrics
Round-Trip Time (RTT), Uplink/Downlink Throughput, Jitter, Link Availability.	R^2 (predictive fidelity), RMSE (precision against handover spikes), MAE (accuracy).

Methodology: Model Taxonomy (The 16 Competitors)

The LENS Dataset: 30 days of real Vancouver Starlink telemetry,
417 Million observations, strictly chronological split.

The Suspect Board: 16 AI Models

Frequency / Multi-scale	Transformers	Linear / MLP	State-Space / Foundation
A Priori Alignment: Strong	A Priori Alignment: Moderate	A Priori Alignment: Poor	A Priori Alignment: Weak
Explicitly designed to decompose periodic waves.	Heavyweight attention mechanisms looking for sequence patterns.	Basic trend decomposition.	Smooth state evolution and generic pre-training.
MSGNet, TimeMixer, TimesNet, TimeFilter, WPMixer, MultiPatchIF.	iTransformer, TimeXer, PatchTST, PAttn, Autoformer, Informer.	Koopa, DLinear.	Mamba, Sundial.



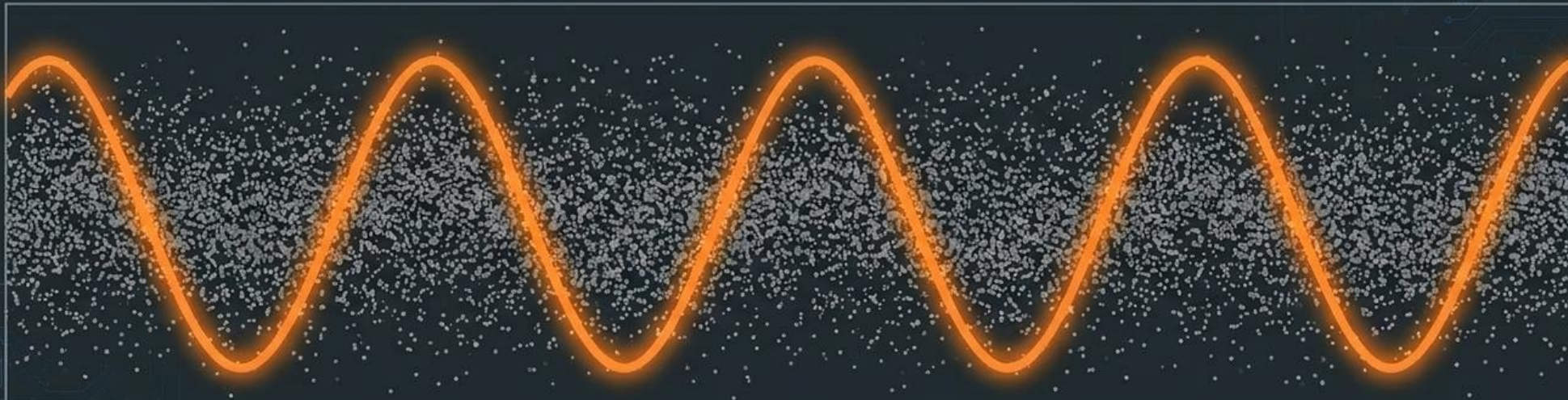
Methodology: Experimental Protocol

Suppressing confounds to isolate pure architectural performance.



Results : Research Question 1

Research Question 1: Do physics-aware models systematically win?

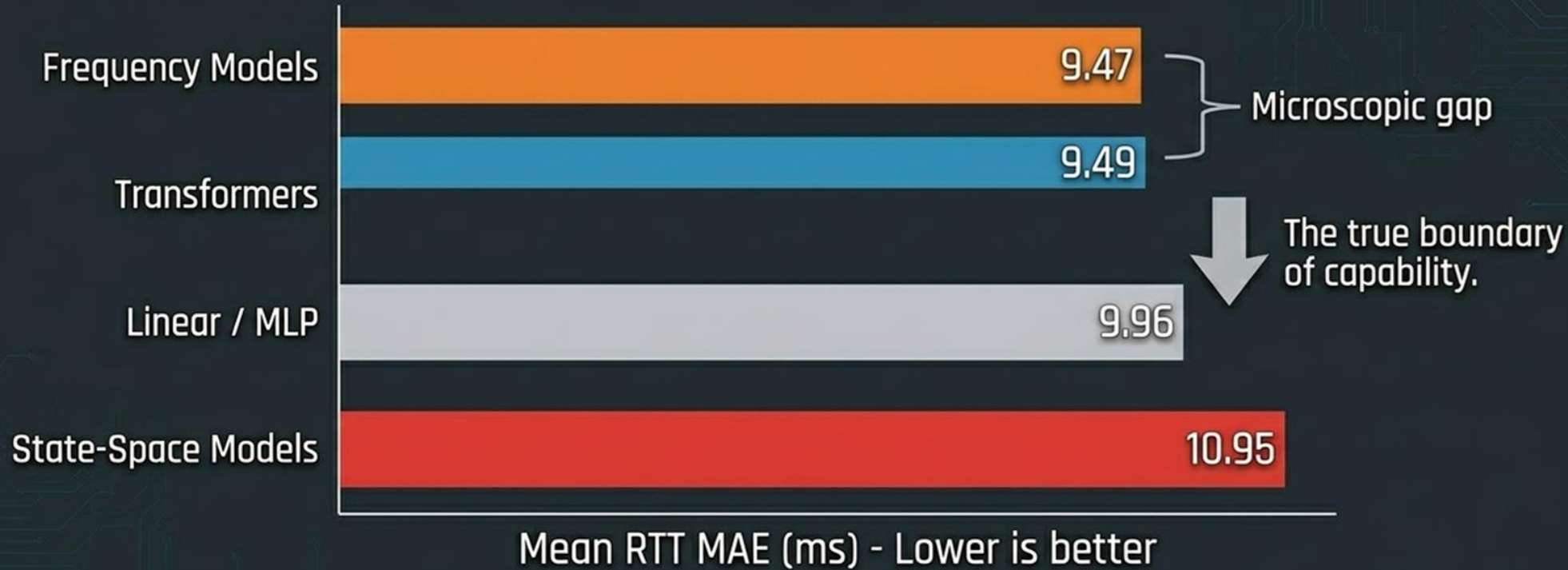


Hypothesis 1 assumes that architectures matching the quasi-periodic orbital dynamics of LEO satellites (Frequency models) will effortlessly outperform architectures treating telemetry as generic sequential data (Transformers).



Results : Research Question 1

The Paradigm Showdown: RTT Prediction Accuracy



Spectral alignment provides a victory, but not an absolute dominance in predictive accuracy.

Results : Research Question 1

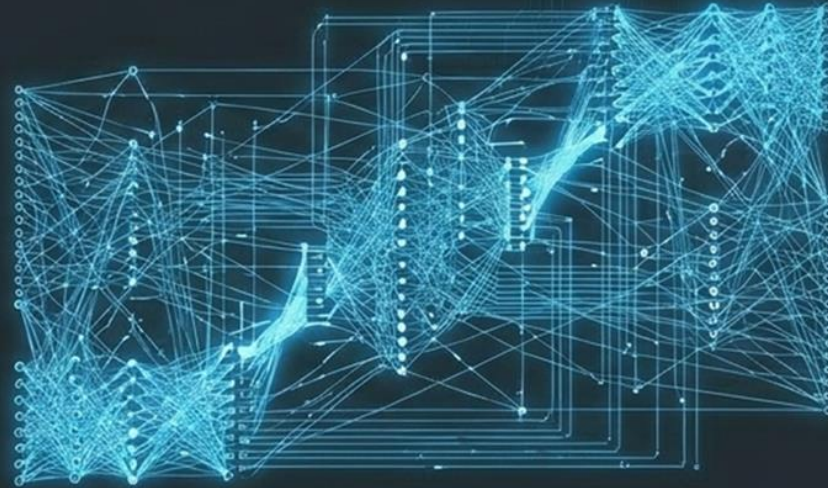
The Real Advantage is Efficiency, Not Accuracy



TimeMixer

Size: 93.9K Parameters
Accuracy: 9.38 ms RTT MAE

=



iTransformer

Size: 6.36M Parameters
Accuracy: 9.38 ms RTT MAE

The real advantage of spectral alignment is achieving state-of-the-art accuracy at 1/67th the computational size.

Results : Research Question 1

The **Catastrophic Failure** of Pre-training

Zero
(Naive Mean Baseline)

Sundial
R-squared = -0.471

SYSTEM LOG ALERT:

Sundial was pre-trained on one trillion generic time points.

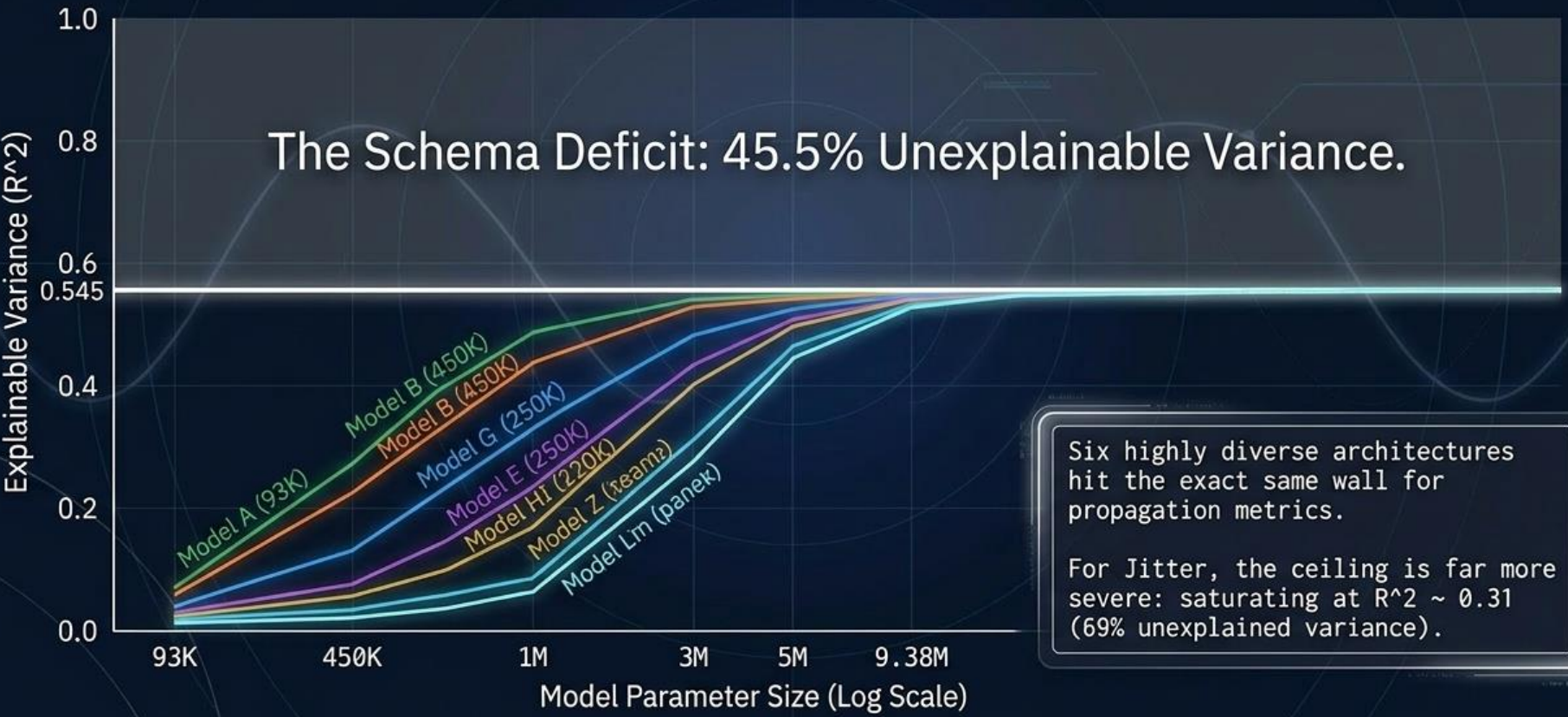
But general internet data does not behave like LEO orbital mechanics.

Pre-training without domain adaptation actively destroys predictability in satellite telemetry.



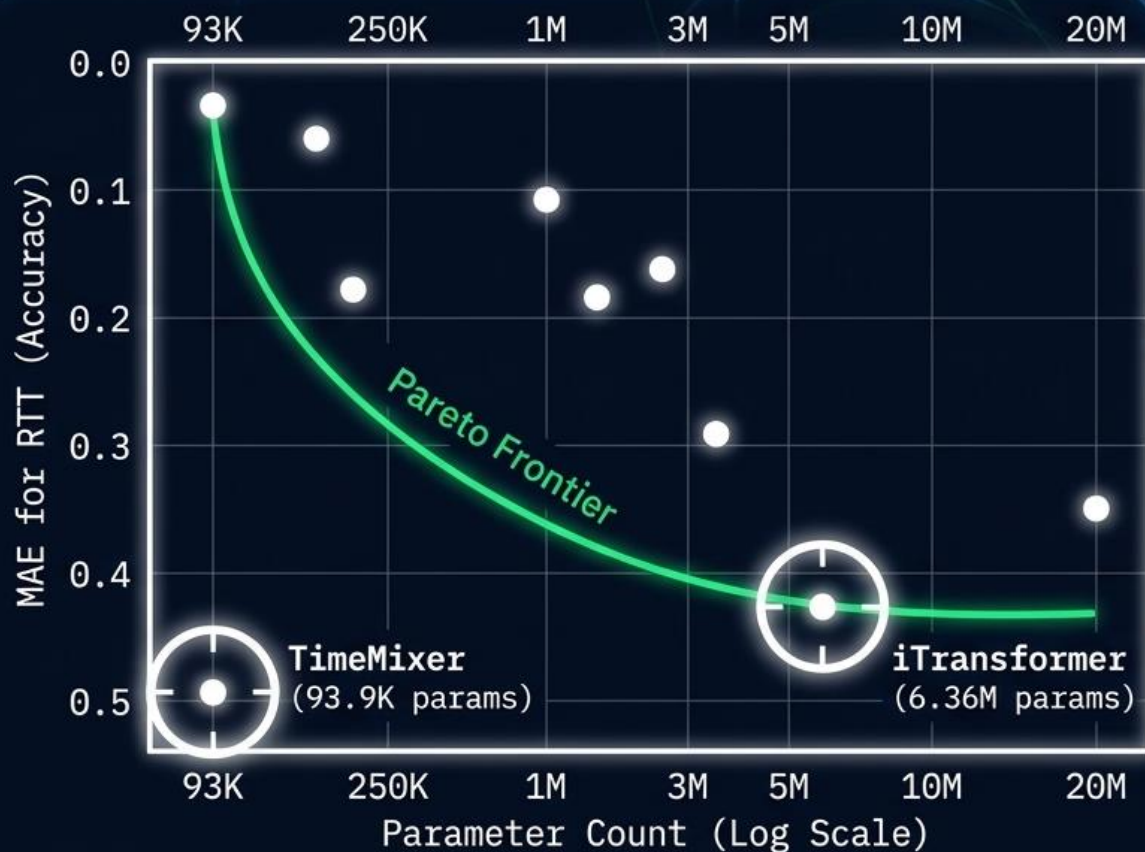
Results: Research Question 2

The Data-Schema Limit: An architecture-invariant predictability ceiling.



Results

The Pareto optimal deployment framework for LEO networks.



Edge-Tier

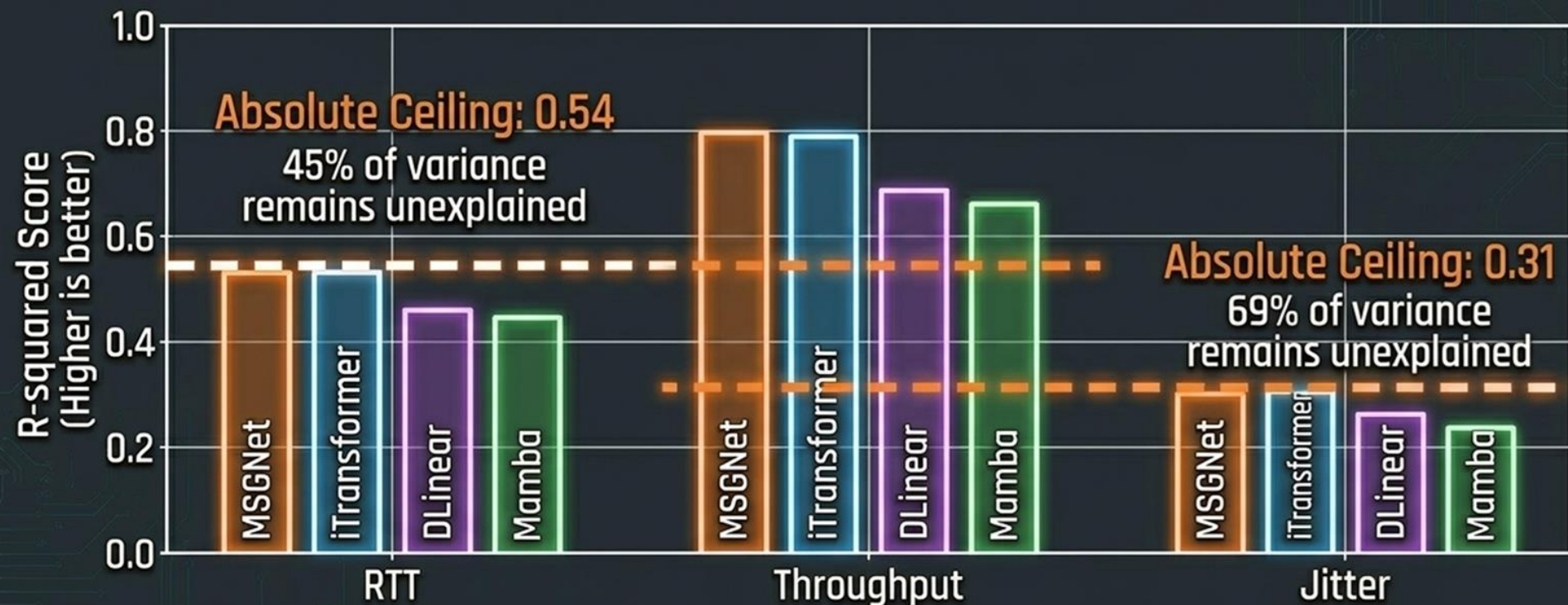
TimeMixer achieves globally optimal MAE at 67x smaller footprint than the best transformer.

Cloud-Tier

iTransformer defines the retraining-latency frontier for ground-station adaptation within a single 5-8 minute orbital pass (13.8min retraining).

Results

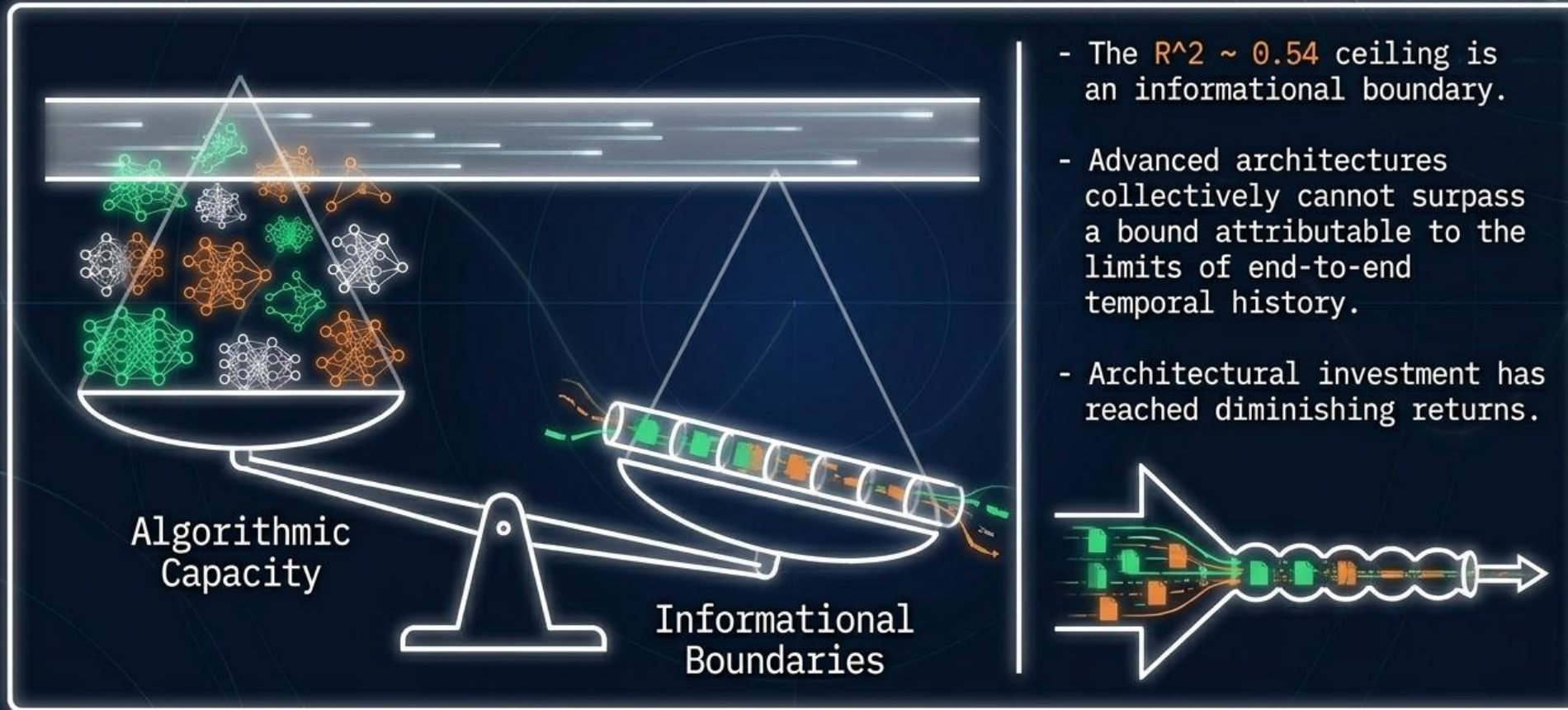
Hitting the Wall: The Absolute Predictability Limits



No algorithm can breach this barrier.
This is not an algorithmic failure; it is an informational limit.

Results: The Necessity of Schema Augmentation

The bottleneck is the 9-column schema, not algorithmic expressivity.

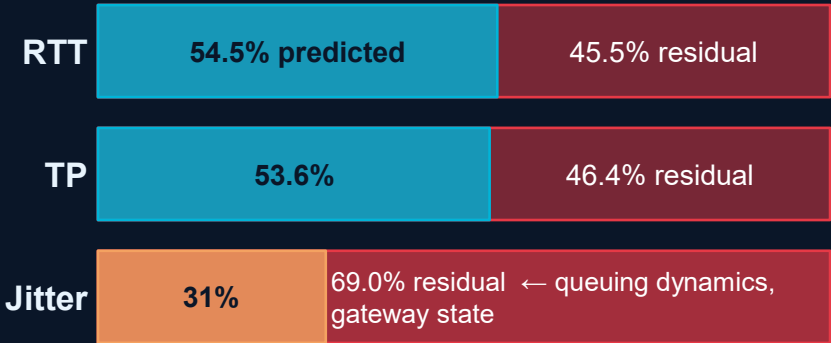


The Data-Schema Bottleneck: A Fundamental Information Limit

What the 9-column schema cannot capture:

- ⚠ Physical-layer SINR / beam ID
- ⚠ Gateway congestion state
- ⚠ Satellite ephemeris
- ⚠ Cross-user traffic interference
- ⚠ Atmospheric / weather context

Explained vs. Unexplained Variance



Architecture-Invariant Ceiling
Top-6 cluster: $\Delta R^2 = 0.011$ (4 freq. + 2 Transformer models)
Convergence evidences information limit, not architectural failure

System design implication:

No further architectural refinement operating on this 9-column schema will breach:

- RTT ceiling at $R^2 \approx 0.545$
- Jitter ceiling at $R^2 \approx 0.310$

The decisive next intervention is schema augmentation, not a better model.

"This plateau represents a fundamental informational limit of the data schema itself, not an algorithmic deficiency."

Experimental Limitations

Experimental constraints and conservative biases.

⚠ Scope Limit

Single residential terminal in Vancouver over a 30-day window. Excludes seasonal and global constellation-growth non-stationarity.

VANCOUVER TERMINAL:
SINGLE UNIT

WINDOW:
30 DAYS

CONSTELLATION GROWTH:
EXCLUDED (NON-STATIONARY)



⚠ Harmonization Bias

Applying harmonized hyperparameters actively underestimated the ceiling potential of large-capacity Transformers.

HYPERPARAMETERS:
HARMONIZED

UNDERESTIMATION:
CEILING POTENTIAL

MODEL TYPE:
LARGE-CAPACITY TRANSFORMERS



⚠ Deployment Reality

Evaluation in full 32-bit precision lacks the inference latency profiling and quantization noise inherent to true edge deployment.

PRECISION:
FULL 32-BIT

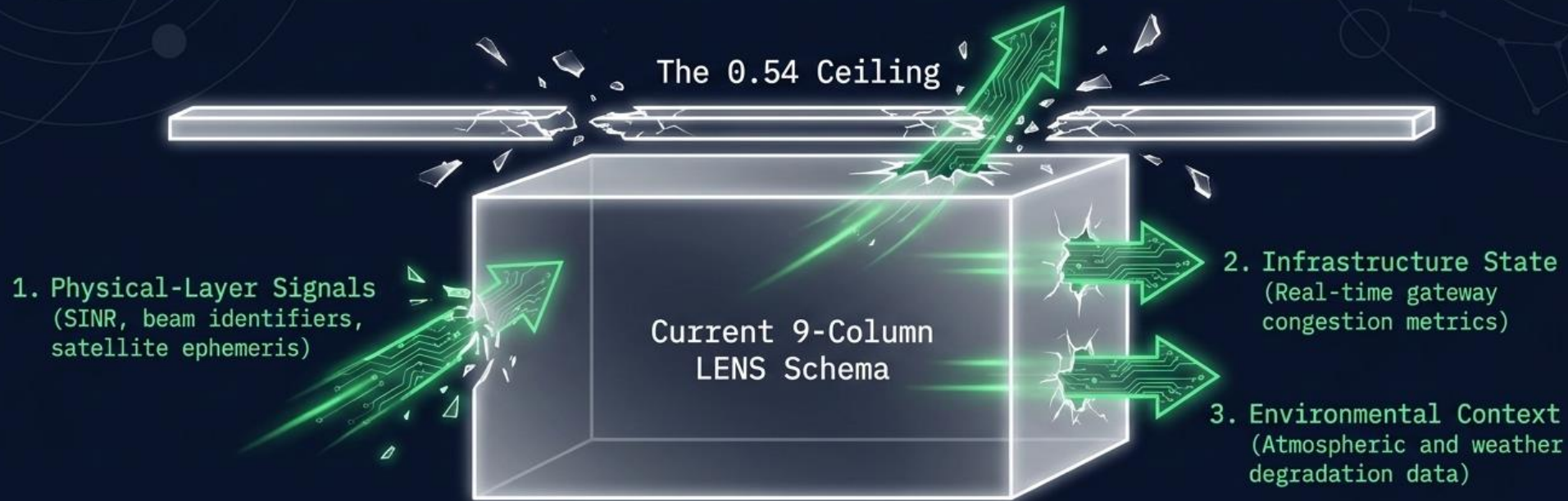
LATENCY PROFILING:
LACKING

QUANTIZATION NOISE:
INHERENT (EDGE DEPLOYMENT)



Future Work — Multi-Modal Feature Fusion

Breaking the ceiling requires cross-layer schema augmentation.



To reclaim the 45.5% unexplained RTT variance and 69.0% unexplained jitter, next-generation AI pipelines must fuse temporal history with cross-layer data.

Future Work: Continuous Learning

Adapting to minute-scale variability across the global constellation.



Geographic Generalization

Testing the data-schema limit across diverse **orbital shells** and regional **demand profiles**.

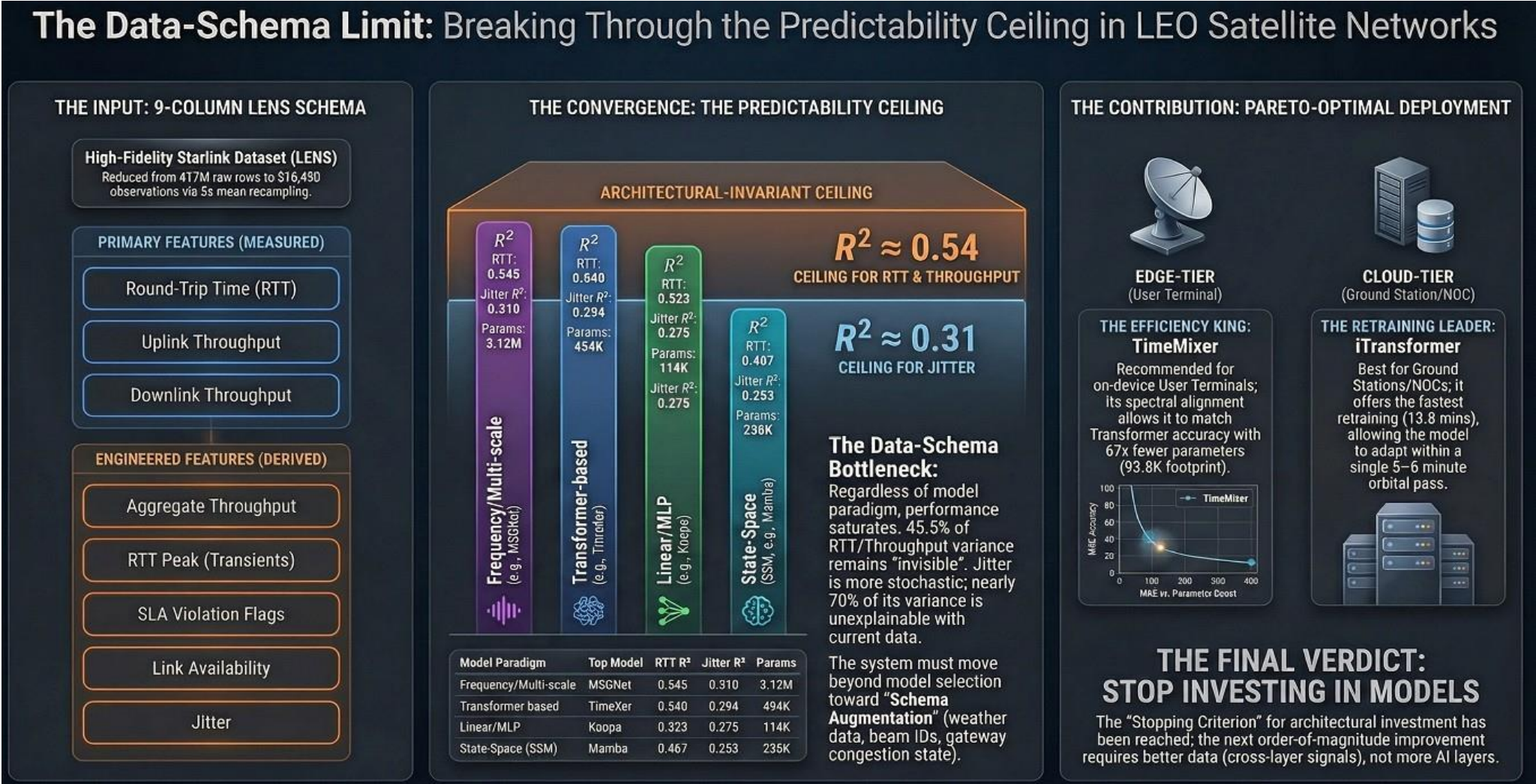
Continual Online Learning

Static trained models will **rapidly degrade** as satellites are replenished and **routing policies** shift.

Operational Constraints

The future requires incremental fine-tuning that adapts within strict **minute-scale** operational **latency constraints**.

Conclusion & Core Takeaways



Conclusion & Core Takeaways

Three deployment principles for next-generation LEO networks.

1

Edge-tier requires compact spectral models

Deploy sub-100K parameter frequency-aware models (e.g., TimeMixer) for on-device QoS prediction without cloud dependency.

2

Cloud-tier requires latency-constrained retraining

Utilize inverted-variate transformers (e.g., iTransformer) for rapid, within-pass adaptation at the ground station.

3

Stop chasing architectures; augment the schema

The absolute limit of temporal sequence forecasting has been reached ($R^2 \sim 0.54$). The next order-of-magnitude leap requires physical-layer feature fusion.



Thank you for your time and attention.

Any questions/comments?

