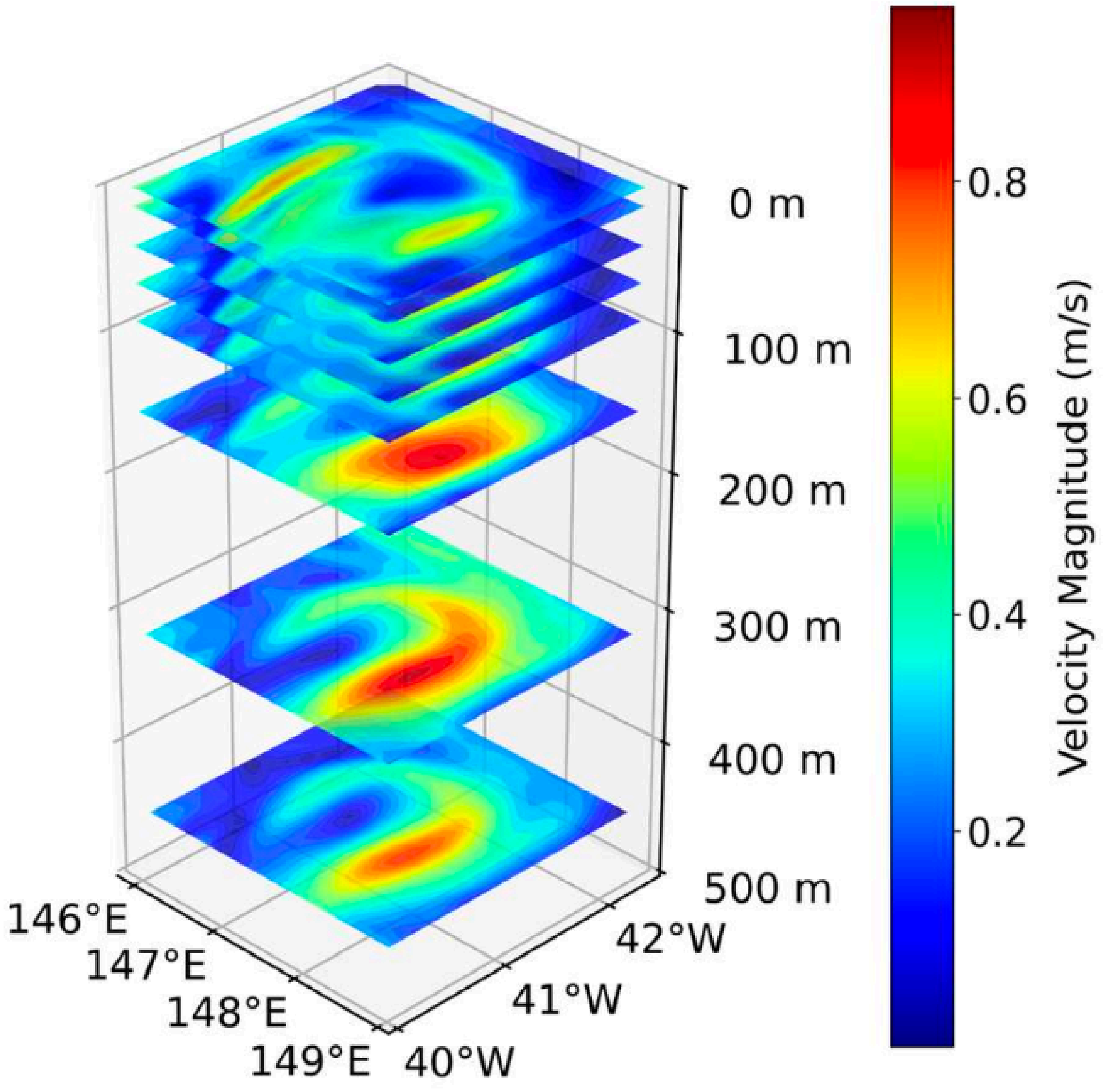
Reconstruction Of Three Dimensional Upper Ocean Fields

MLPR PROJECT 2026



Presented by :

Rohan Goyal | Vir Dang | Aditya Jaitly

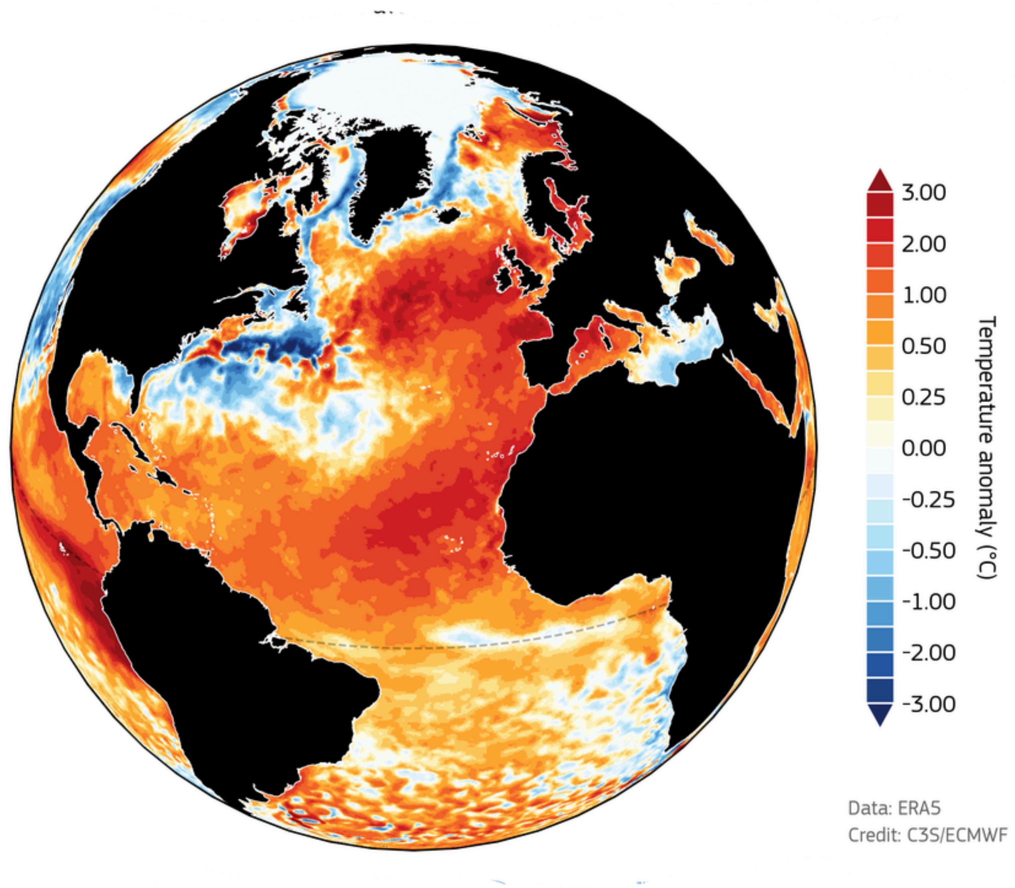
Problem Statement

Reconstruct subsurface ocean temperature *and* ***salinity fields*** *from* ***satellite-derived surface observations*** *using a* ***physics-informed 2DV-UNet framework****, while incorporating* ***Argo*** *float profile* ***data*** *for* ***validation*** *in dynamically complex ocean regions such as the* ***Bay of Bengal.***

Why Reconstruct Subsurface Ocean Fields?

Subsurface ocean data such as temperature, salinity, and currents are **critical** for **understanding** key **processes** such as:

- Ocean heat transport and climate variability
- Marine heatwaves and ENSO dynamics
- Ocean stratification and vertical mixing
- Mesoscale eddies and circulation patterns



Key Challenges for Reconstruction of Subsurface fields

Limited Subsurface Data

- Satellite missions provide dense surface measurements, but **subsurface** ocean data remain **sparse** and irregularly distributed.

Complex Surface–Subsurface Relationships

- Ocean variables interact through highly **nonlinear** processes, making it difficult to infer subsurface structure directly from surface conditions.

Physical Consistency

- **Purely data-driven** reconstruction **models** can generate **physically unrealistic fields** without additional physical constraints.

Computational Constraints

- High-resolution deep learning reconstruction models require significant memory, storage, and computational resources, making large-scale deployment challenging in resource-constrained operational environments.

LITERATURE REVIEW

ARMOR3D Reconstruction Model (Baseline Model)

A **widely used** ocean map that tries to **guess** what the **deep ocean** looks like **using satellite imagery** of the surface of the ocean.

How It Works:

- **Looks at the Surface:** It **collects satellite data** like surface temperature (SST) , saltiness (SSS), and sea height (SSH).
- **Makes a Statistical Guess:** Uses **multivariate linear regression** to project satellite surface data (SST, SSS, SSH) into estimate subsurface values.

Major Shortcoming:

- **Ocean** processes involve **nonlinear interactions** between temperature, salinity, andurrents, which **statistical models cannot** capture effectively.

RMSE for Temperature (0-200m depth) :
≈ 0.30∘C to 1.10∘C

RMSE for Salinity (0-200m depth) :
≈ 0.15psu to 0.35psu

S. Guinehut, A. Larnicol, and P.-Y. Le Traon, “The global ocean reanalysis and simulation system ARMOR3D,” Journal of Atmospheric and Oceanic Technology, vol. 29, no. 10, pp. 1439–1450, 2012.

3DV-UNet Reconstruction (Flagship Model)

Uses a **hybrid CNN–Transformer** architecture to reconstruct three-dimensional ocean variables from satellite observations.

How it Works:

- Uses a **U-Net convolutional architecture** to learn spatial features
- Incorporates a **3D Vision Transformer bottleneck** to model **cross-depth & cross-variable dependencies**
- Reconstructs multiple variables including: temperature, salinity, zonal & meridional velocity

.

Major Shortcoming:

- **Lack** of explicit **physical constraints,** making the model primarily data-driven rather than physically guided.

RMSE for Temperature (0-200m depth) :
≈ 0.30∘C

RMSE for Salinity (0-200m depth) :
≈ 0.11psu

X. Zhu, Y. Li, H. Wang, and J. Liu,
“3DV-UNet: Eddy-resolving reconstruction of three-dimensional ocean variables from satellite observations,”
Remote Sensing, vol. 17, no. 19, 2025.

Research Gap

While both statistical baselines and global deep learning models offer valuable insights, they leave critical gaps when applied to highly volatile, stratified regions like the Bay of Bengal specially in the upper water column.

Where ARMOR3D Fails:

- Assumption of Linearity: Uses **linear statistical equations** that **fail to map the complex, non-linear** mixing of heat and salt.
- The Profile Bottleneck: **Processes data using columns** or localized smoothing, completely **ignoring the horizontal fluid dynamics** and eddy boundary paths.

Where 3DV-UNet Fails:

- **Physically unrealistic predictions: Lacking explicit physical constraints**, the model operates on a **purely data-driven** basis rather than a physically guided one. This can yield predictions that are mathematically optimized but physically impossible.
- **Global Generalisation:** Designed as a **global model**, it **averages out the hyper-local,** extreme coastal **anomalies** unique to localised basins (Struggles with **massive river runoff,** Ganges-Brahmaputra, in the **Bay of Bengal**, where intense surface salinity capping **breaks standard data patterns**) .
- **Data & Compute Heavy:** The model is **highly computationally expensive** to **train, deploy** and makes **real-time inference challenging** compared to compact models in **resource-constrained environments**.

Addressing The Research Gap

1. **Physics-informed loss functions:**

- Integrates dynamic thermodynamic parameters to enforce physical consistency and eliminate unphysical profiles (e.g., density inversions).

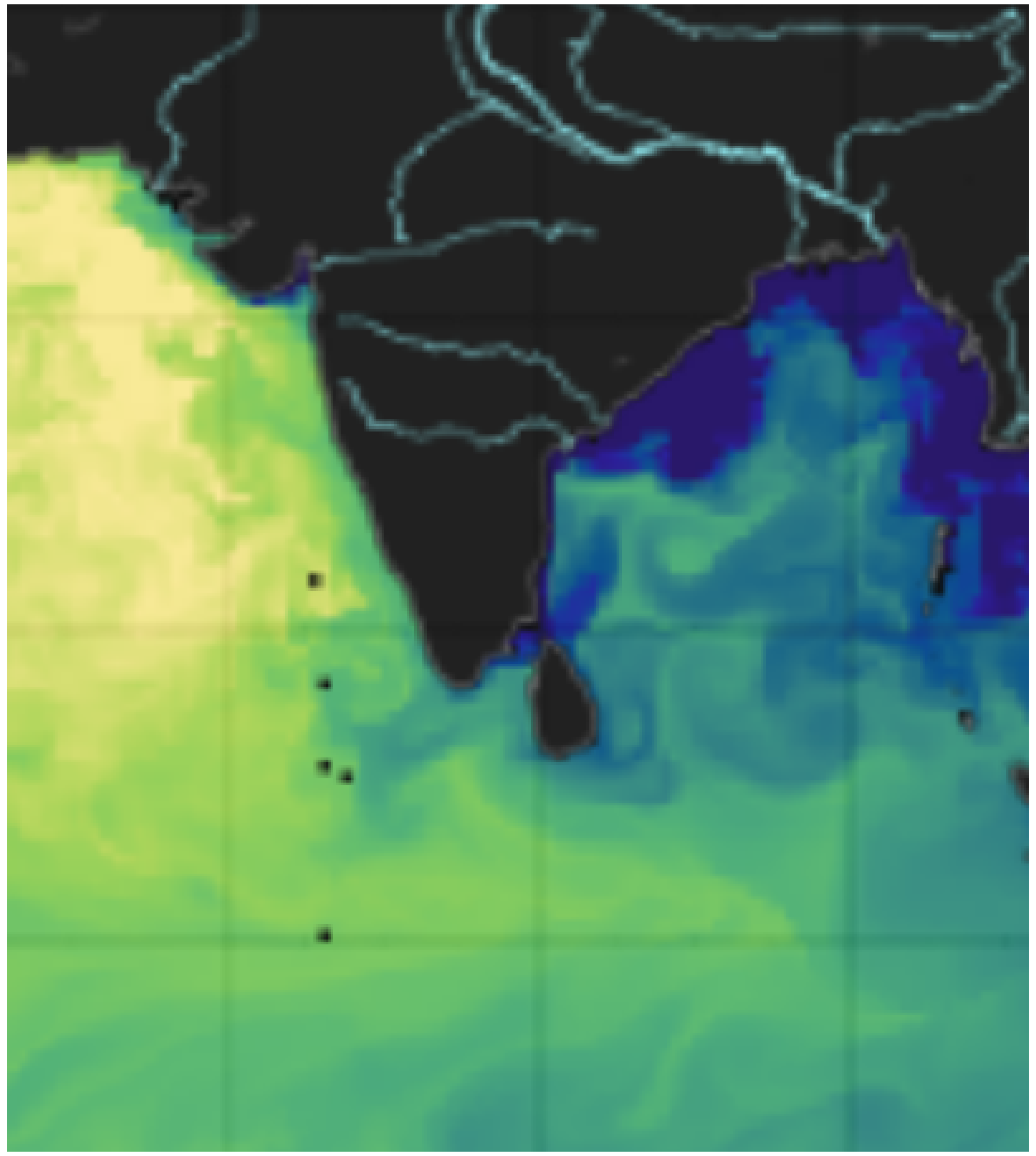
2. **Multi-Predictor Matrix:**

- Leverages an **8-channel input array**—including absolute **geostrophic vectors** and **wind forces**—to strengthen surface-to-subsurface mapping and **bypass traditional satellite observation limits.**

3. **Region-Specific Modeling:**

- Tailored for the **Bay of Bengal** (0–200m), optimizing the network to learn complex, hyper-local coastal ocean dynamics.

Dataset and Features Preprocessing

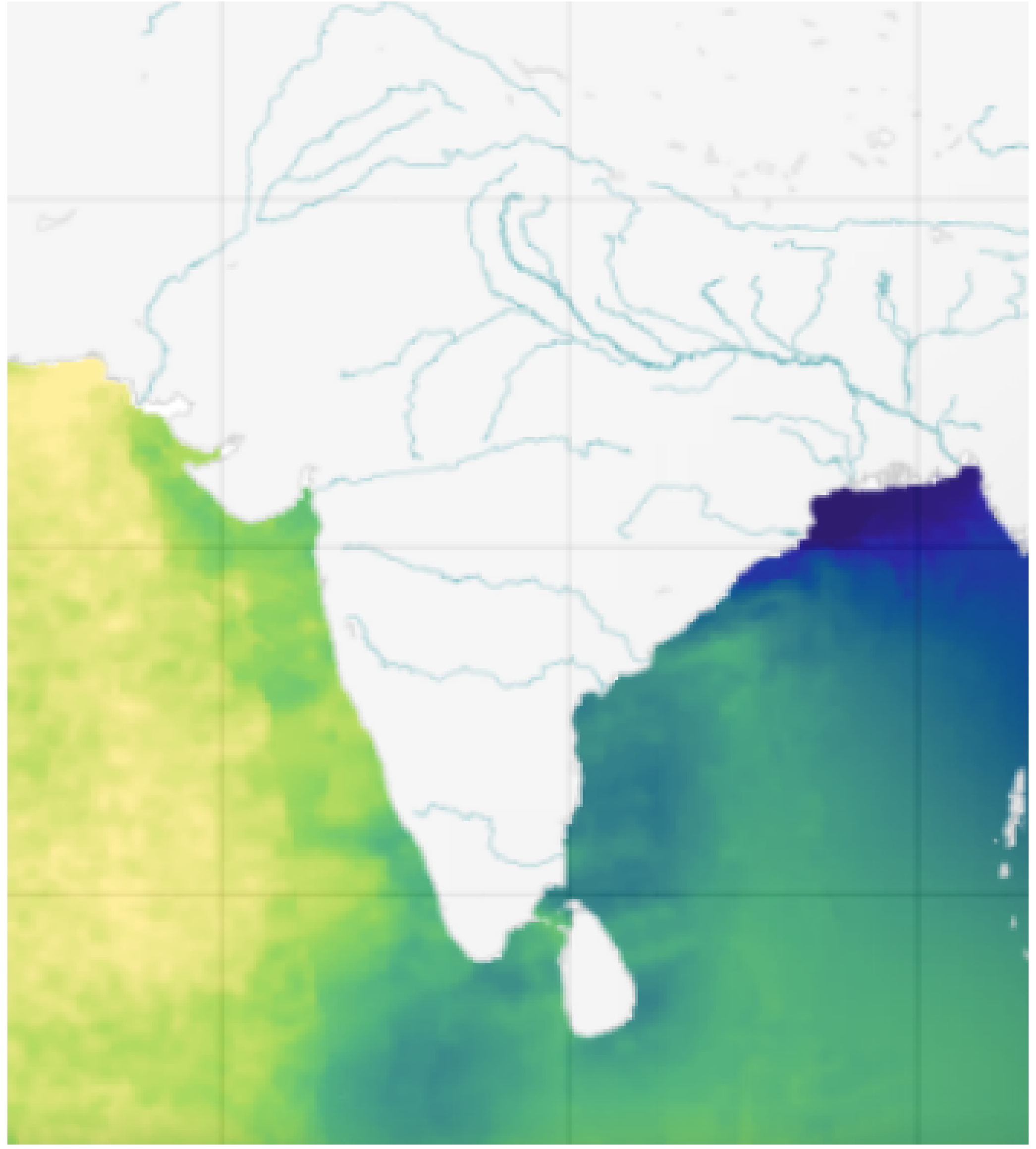


GLORYS12V1 (Global Ocean Reanalysis)

GLORYS12V1 is a high-resolution global ocean reanalysis produced by Mercator Ocean. Its nature is a numerical model integrated with real-world observations via data assimilation. The primary advantage is its 1/12° horizontal resolution and 50 vertical levels, providing a seamless 3D physical state of the ocean where observations are sparse.

- **Collection:** Data is "collected" by assimilating satellite altimetry, SST, and in-situ T/S profiles into the NEMO platform using a reduced-order Kalman filter.
- **Ethics:** No ethical concerns; it uses public-domain environmental data and follows open-science protocols.

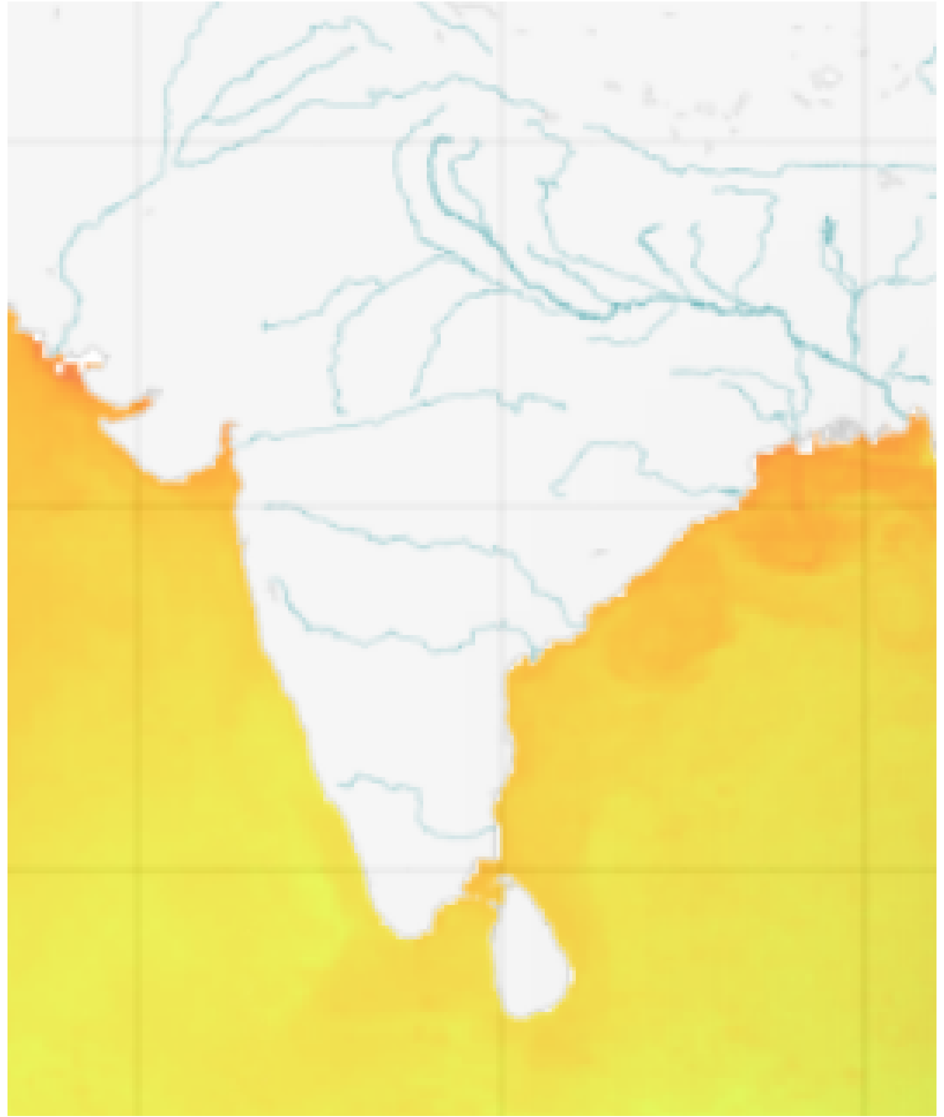
CMEMS MultiObservation Product



The CMEMS MultiObservation SSS/SSD is a globally gridded dataset providing sea surface salinity and density fields. Its nature is a statistically merged observational product. The primary advantage is its ability to provide a complete, high-quality map by bridging the gaps between different satellite missions.

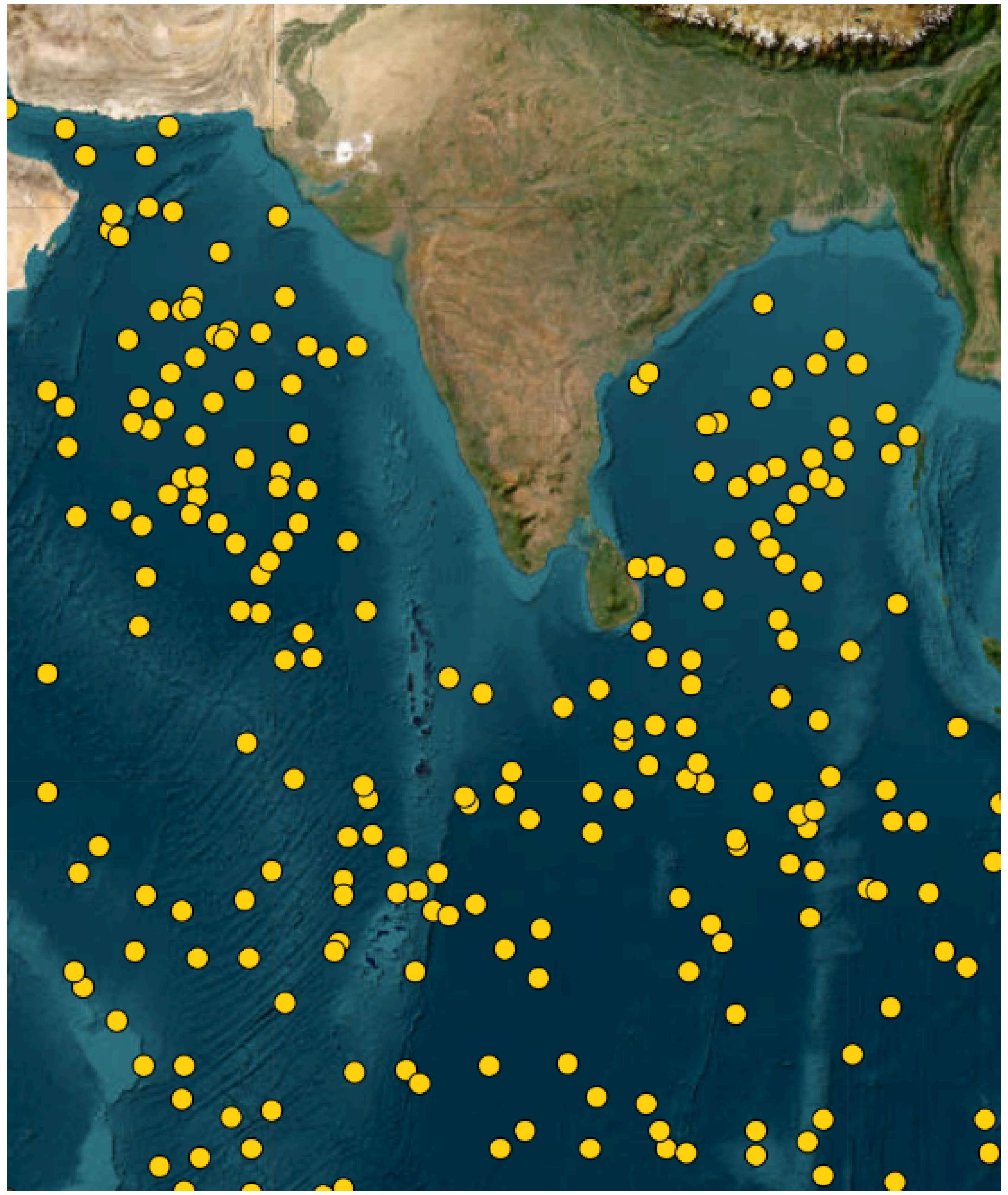
- Collection: Data is collected by blending satellite-derived salinity measurements (from missions like SMOS, SMAP, and Aquarius).
- Ethics: No ethical concerns; it relies on non-commercial, environmental Earth observation data distributed under Copernicus open-data policies.

GHRSST L4 REMSS



GHRSST L4 REMSS is a globally complete, high-resolution daily sea surface temperature dataset produced by Remote Sensing Systems. The primary advantage is its synthesis of different sensor types to provide a continuous, reliable SST map regardless of weather or cloud cover.

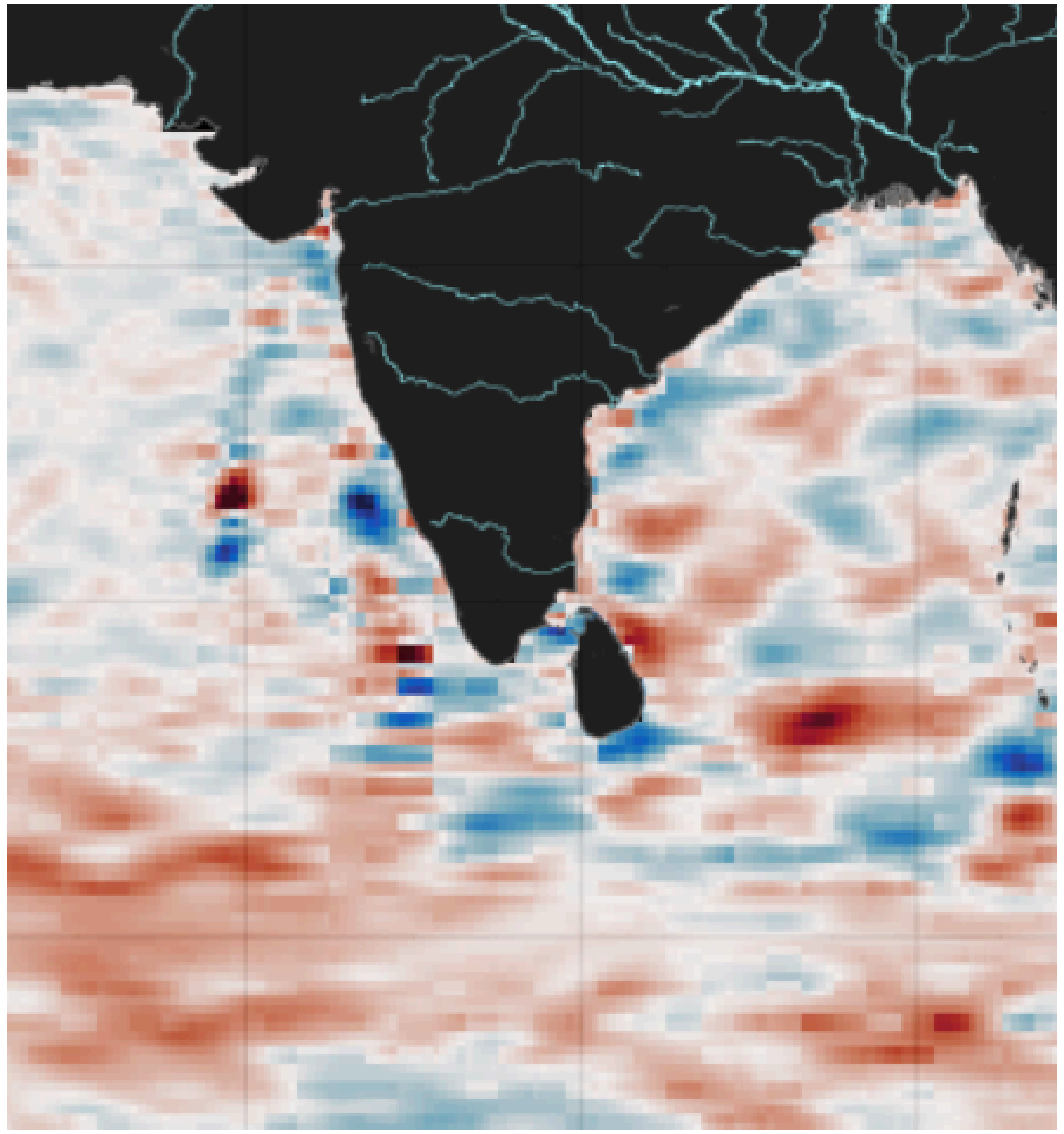
- Collection: Data is "collected" by mathematically merging infrared satellite readings with microwave satellite readings (which penetrate clouds) .
- Ethics: No ethical concerns; it utilizes publicly funded, public-domain Earth observation satellite data .



Argo Program (In-Situ Profiling)

The Argo dataset is a global array of autonomous robotic floats. Its nature is direct in-situ measurement, making it the "ground truth" for oceanography. The main advantage is that it captures the subsurface ocean (down to 2,000m) in real-time, regardless of weather conditions that might block satellites.

- **Collection:** Floats descend to a "parking depth," drift for 10 days, and then record T/S profiles as they rise to the surface to beam data to satellites.
- **Ethics:** Managed by the International Argo Steering Team; they address "maritime territory" concerns by notifying coastal states when floats enter their Exclusive Economic Zones (EEZ).

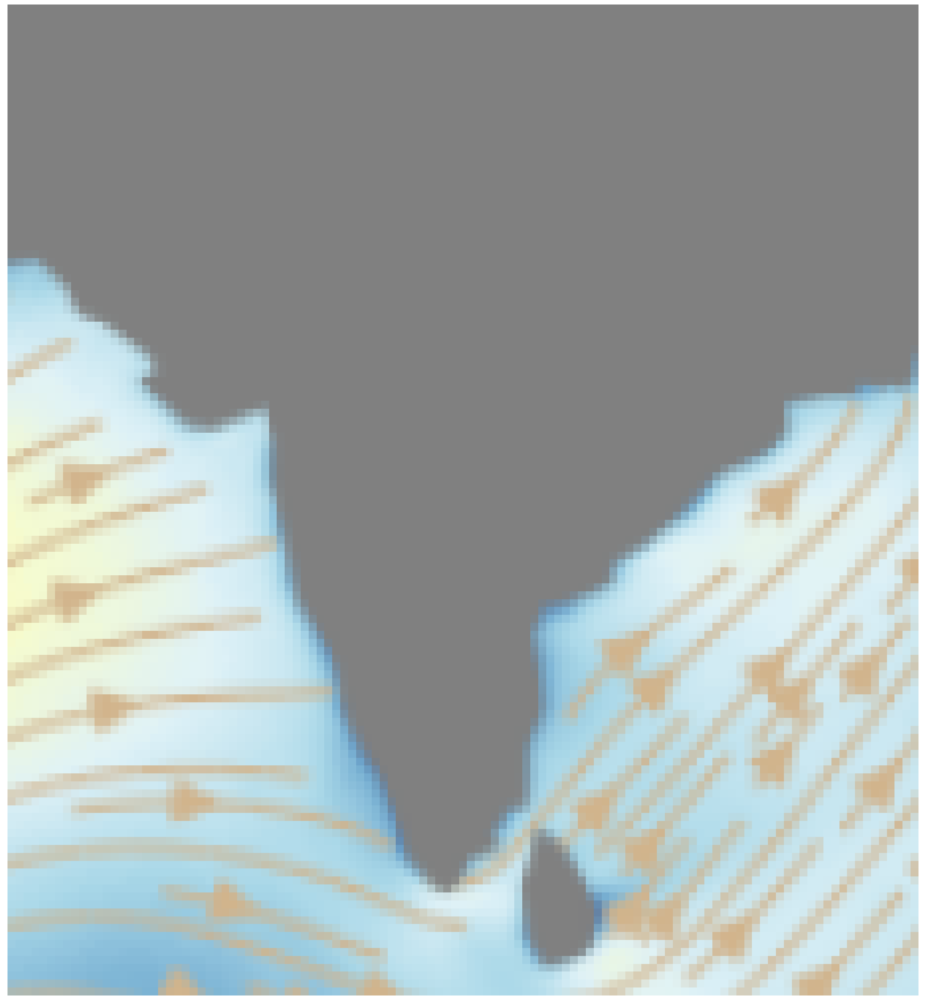


DUACS (SEALEVEL_GLO_PHY_L4_MY)

It is a statistical reconstruction (Level 4) that provides a continuous 3D-like view of the ocean surface . The primary advantage is its stability and homogeneity, making it the "climate-quality" benchmark for tracking sea-level rise and mesoscale circulation (eddies) over decades.

- Collection: Data is synthesized by merging "along-track" measurements from all available altimeter missions using Optimal Interpolation. This process fills the gaps between satellite orbits.
- Stats: Features UGOS/VGOS/SLA.

CCMP v3.1 (UWND, VWND)



CCMP v3.1 (Cross-Calibrated Multi-Platform) is a globally gridded Level 4 dataset providing a consistent, gap-free time-series of vector winds over the world's oceans. The primary advantage is its ability to provide spatially complete, 6-hourly wind maps at a 0.25° resolution by seamlessly integrating data from uniquely different satellite sensors.

- Collection: Data is "collected" by mathematically merging 10-meter surface wind retrievals from multiple satellite microwave sensors
- Ethics: No ethical concerns; it utilizes publicly funded, public-domain Earth observation satellite data.
- Stats: Curl calculated using UWND/VWND where missing

Table

GLORYS	CMEMS	REMSS	ARGO	DUACS	CCMP
Ground Truth: T,S	SSS	SST	In situ: T,S	UGOS VGOS	UWND VWND

Data Pre-processing

- All datasets were interpolated to a daily temporal resolution and a spatial resolution of 1/4 ◦ using bilinear interpolation, with spatial bounds in the Bay of Bengal

- All input and output data are processed using sea–land masking to ensure that model errors are calculated only on valid ocean data, thereby ensuring fair and consistent evaluation.

- All input and label data were cleaned through the following steps
    a. Outliers detected and removed.
    b. Units were standardised (e.g., K→°C)

- All Data was Z scored normalised to remove feature dimensionality bias

- Strict temporal split: Train = Jan 2023 – Dec 2024 (731 days), Test = Jan 2025 – Dec 2025 (365 days). No random shuffling — purely chronological to simulate real-world forecasting conditions and prevent temporal leakage.

Methodology

Baseline Models

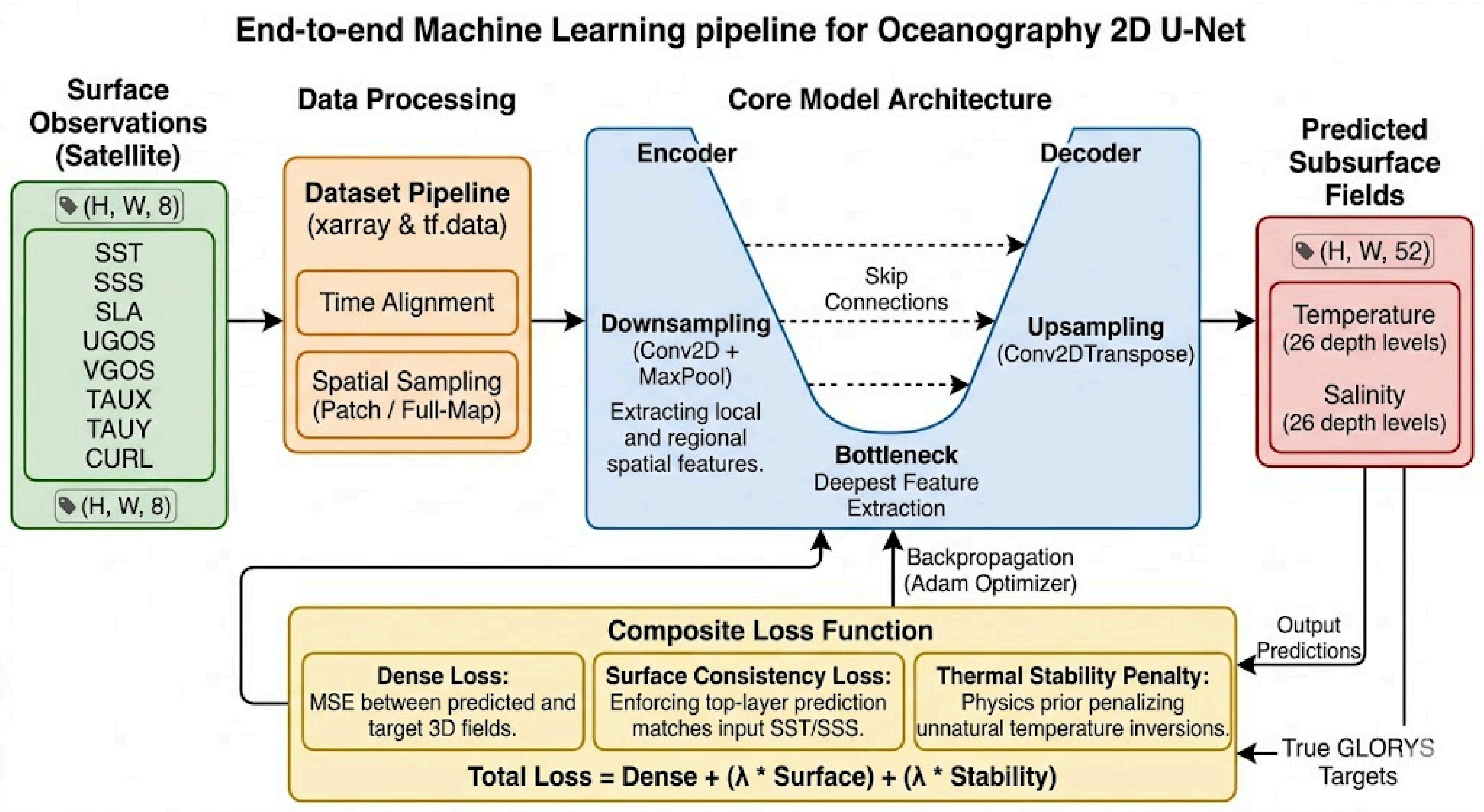
Why Our Model?

The following Models were trained on a subset of our data to perform as a baseline against our model in our specific context:

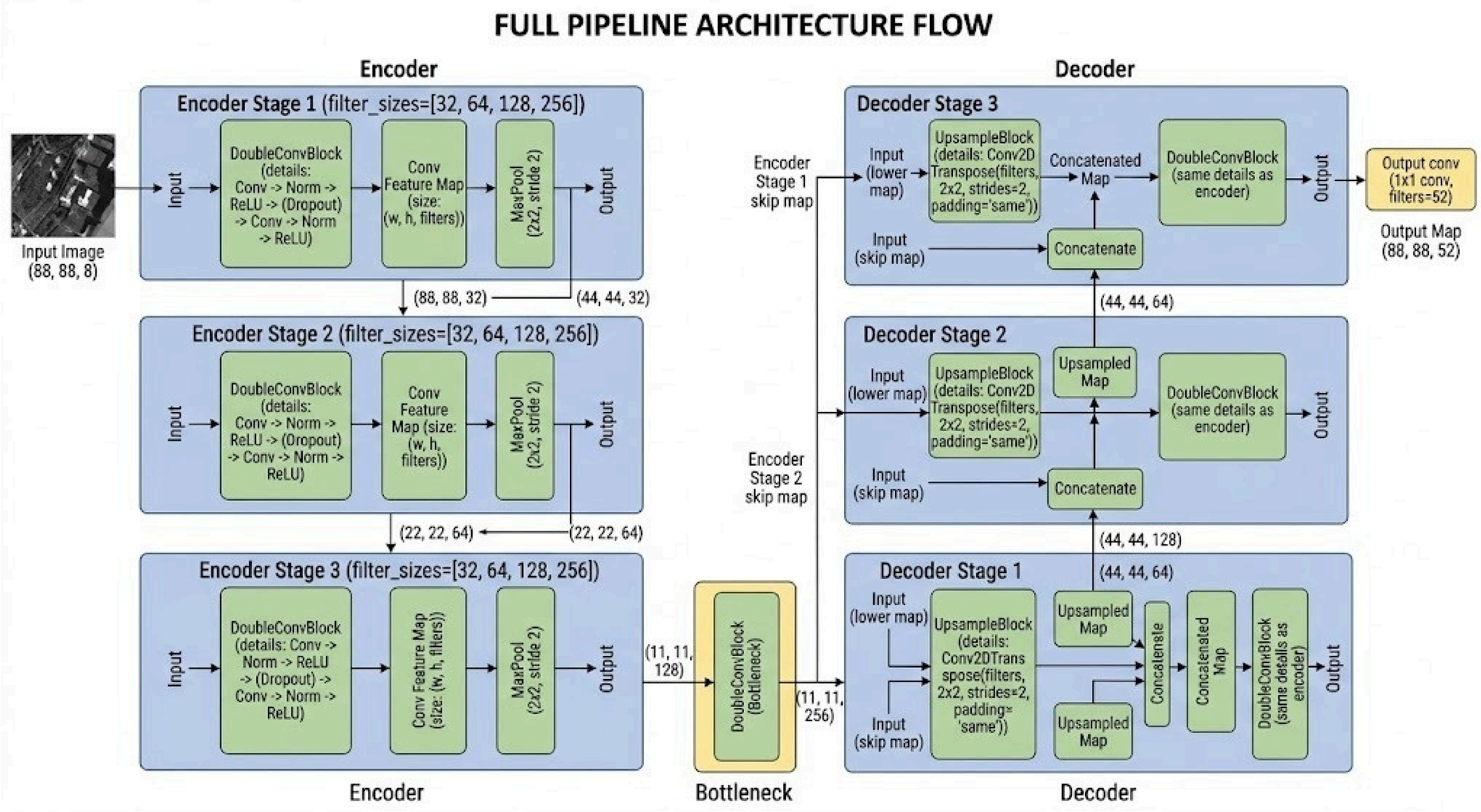
- Ridge: linear regression with L2 regularization reduce overfitting.
- Random Forest: ensemble of decision trees built on bootstrap samples
- SVR : kernel-based method that fits a function within an ε-insensitive margin, using support vectors to control complexity and handle nonlinear relationships.

This provides a standard benchmark for how much U-net improves over simpler regressors.

Our Model


End-to-end Machine Learning pipeline for Oceanography 2D U-Net
Surface Observations (Satellite)
(H, W, 8)
SST
SSS
SLA
UGOS
VGOS
TAUX
TAUY
CURL
(H, W, 8)
Data Processing
Dataset Pipeline (xarray & tf.data)
Time Alignment
Spatial Sampling (Patch / Full-Map)
Core Model Architecture
Encoder
Decoder
Skip Connections
Downsampling (Conv2D + MaxPool)
Extracting local and regional spatial features.
Upsampling (Conv2DTranspose)
Bottleneck Deepest Feature Extraction
Predicted Subsurface Fields
(H, W, 52)
Temperature (26 depth levels)
Salinity (26 depth levels)
Backpropagation (Adam Optimizer)
Output Predictions
True GLORYS Targets
Composite Loss Function
Dense Loss: MSE between predicted and target 3D fields.
Surface Consistency Loss: Enforcing top-layer prediction matches input SST/SSS.
Thermal Stability Penalty: Physics prior penalizing unnatural temperature inversions.
Total Loss = Dense + (λ * Surface) + (λ * Stability)

UNet Structure



Challenges Faced

Challenge	How Addressed	Details
Overfitting	L2 Weight Decay, Dropout, Physics Constraints, Train/Test monitoring	Multiple regularization strategies combined; per-epoch train-vs-test gap monitoring detects divergence early.
Poor Quality Data / Missing Values	NaN masking, valid-pixel filtering	Land pixels and sensor gaps produce NaN values. These are systematically handled via nan_to_num imputation and valid_mask filtering in classical models.

Hardware Hell

What was the Largest Challenge in Implementation?

We were heavily limited by our hardware and as such had to reduce our scope in the following aspects:

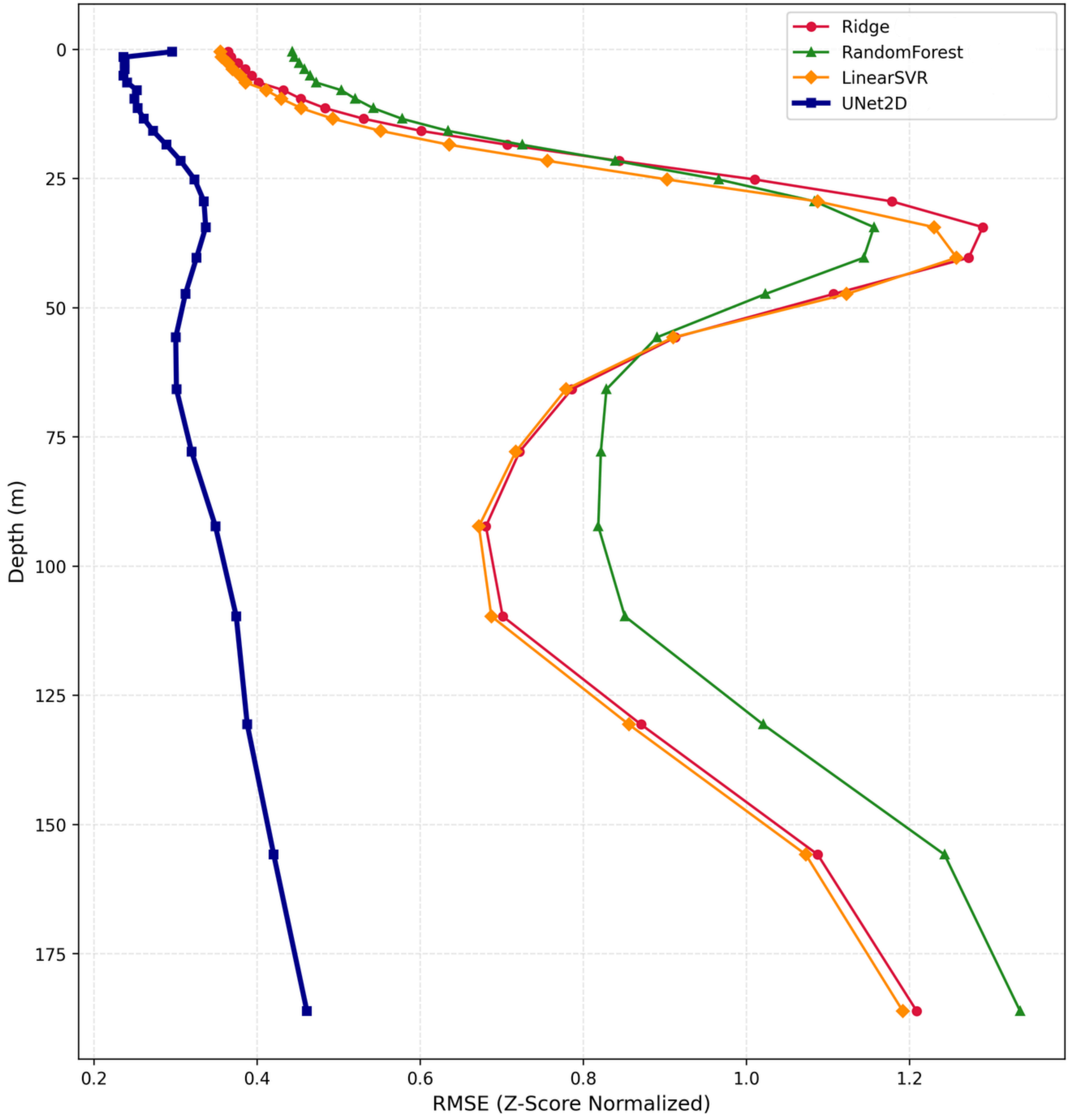
- Temporal : Data reduced from 10 years to 3 years
- Spatial : Constricted to Bay of Bengal for our training
- Complexity : Had to integrate lazy loading for pre processing

This bottleneck significantly altered our implementation plan and reduced our efficiency and output.

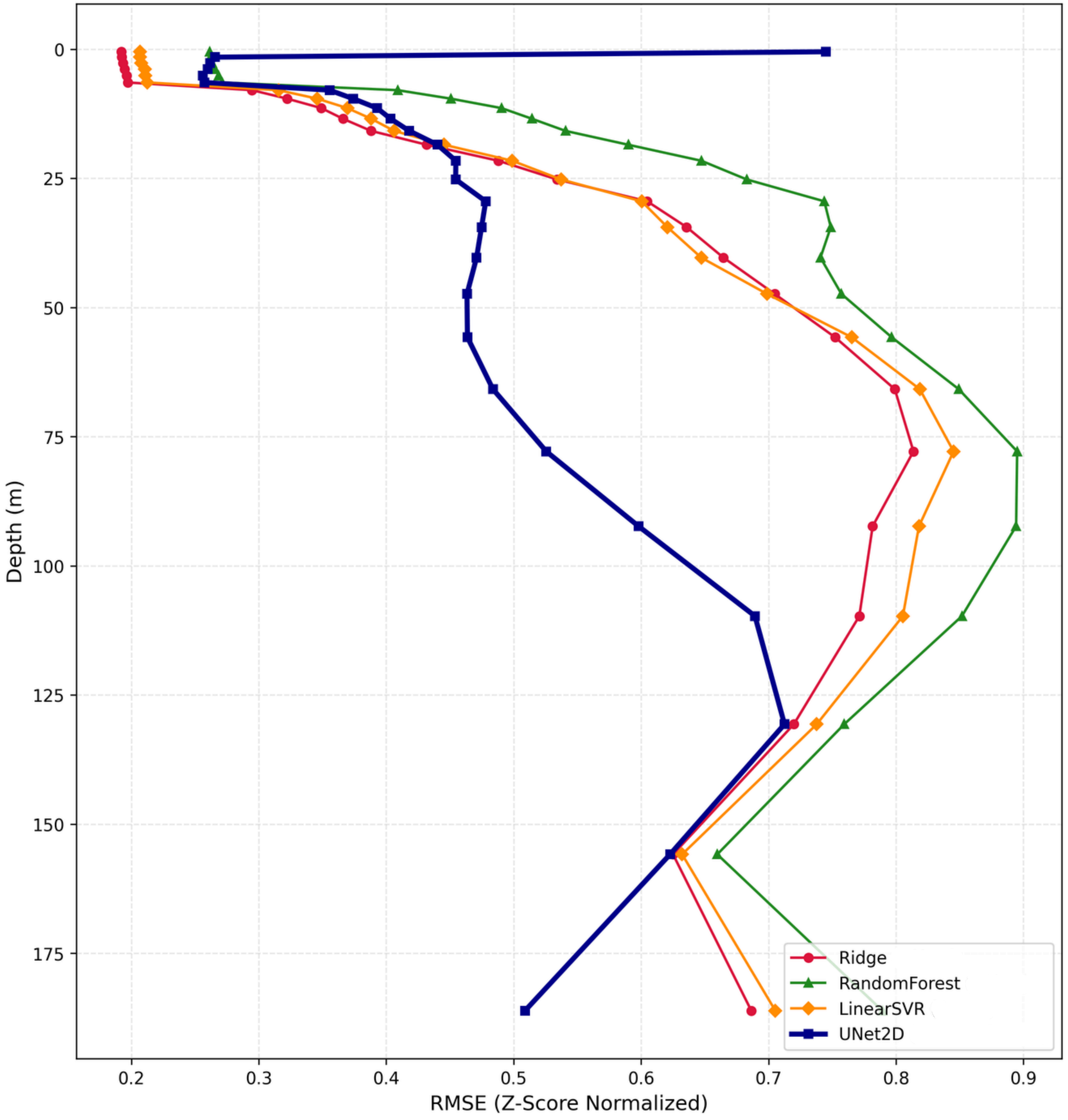
Performance Metrics

Baseline Models

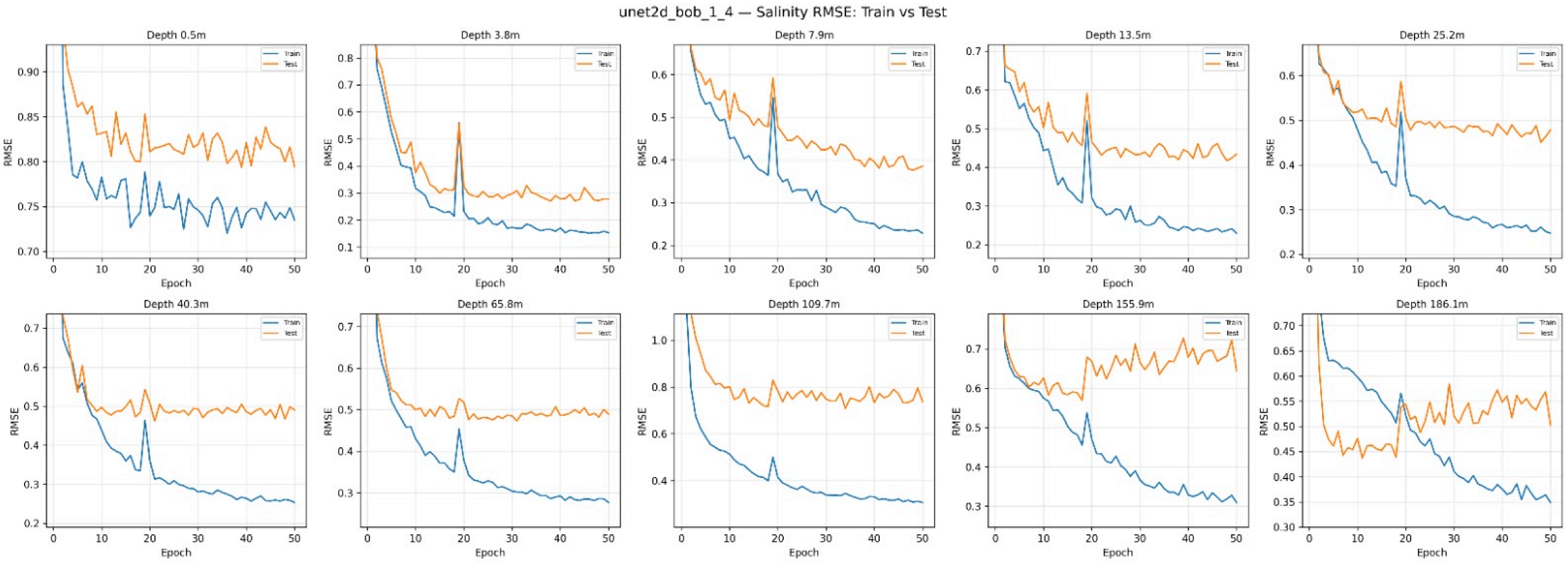
Depth vs. RMSE: Physics UNet2D vs. Classical Baselines (Temperature)



Depth vs. RMSE: Physics UNet2D vs. Classical Baselines (Salinity)

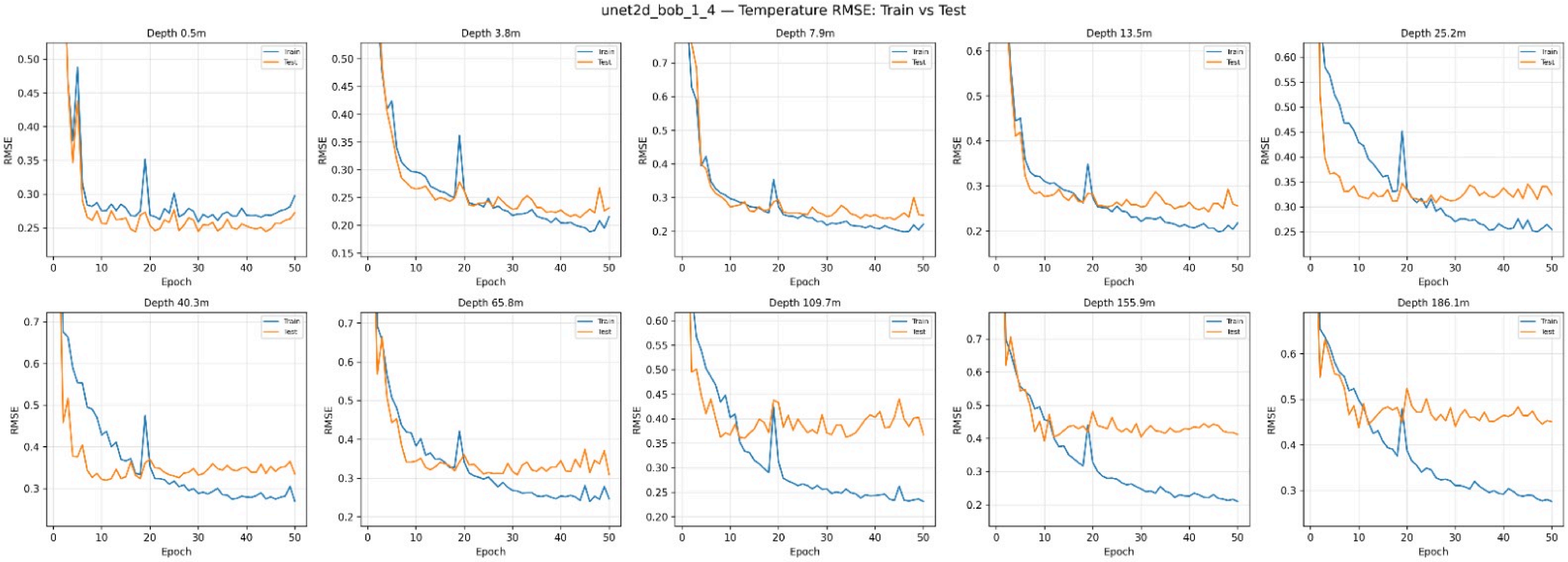


Salinity



unet2d_bob_1_4 — Salinity RMSE: Train vs Test
Depth 0.5m
Depth 3.8m
Depth 7.9m
Depth 13.5m
Depth 25.2m
Depth 40.3m
Depth 65.8m
Depth 109.7m
Depth 155.9m
Depth 186.1m
Train
Test
RMSE
Epoch

Temperature



Deployment

Advantages going forward!

Our model was built with edge computation in mind! By leveraging a 2D U-Net rather than a heavy 3D architecture, the model is highly efficient.

Allowing it to run locally on low-power maritime devices—like shipboard computers, marine buoys, or autonomous underwater vehicles (AUVs)

Challenges to Face!

This mobility comes has results in the following tradeoffs:

- Highly localized - Trained for the specific geo-profile of the Bay of Bengal
- Highly specialized - Focuses on depths where other models struggle however predicts to a lesser depth

Thank You