NeurIPS 2024 Workshop on Tackling Climate Change with Machine Learning

Multi-scale Decomposition of Sea Surface Height Snapshots using Machine Learning

Jingwen Lyu[1*], Yue Wang[1*], Christian Pedersen[2], Spencer Jones[3], Dhruv Balwada[1]

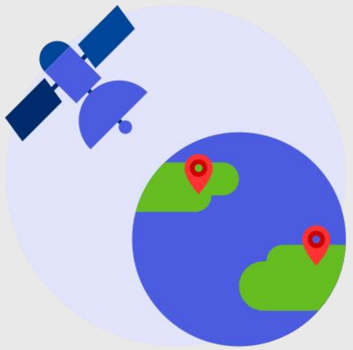
[1]Columbia University, [2]New York University, [3]Texas A&M University

*Equal contribution

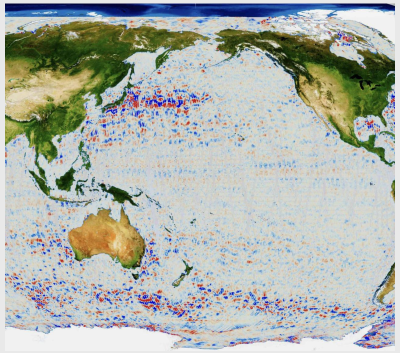
Ocean's Role in Climate Change

Climate Impact:

- 93% of Earth's excess heat storage
- 30% of CO_2 absorption
- Major climate regulator



Satellite Measurement



Sea Surface Height (SSH)

Ocean Flow Patterns

Our Task:

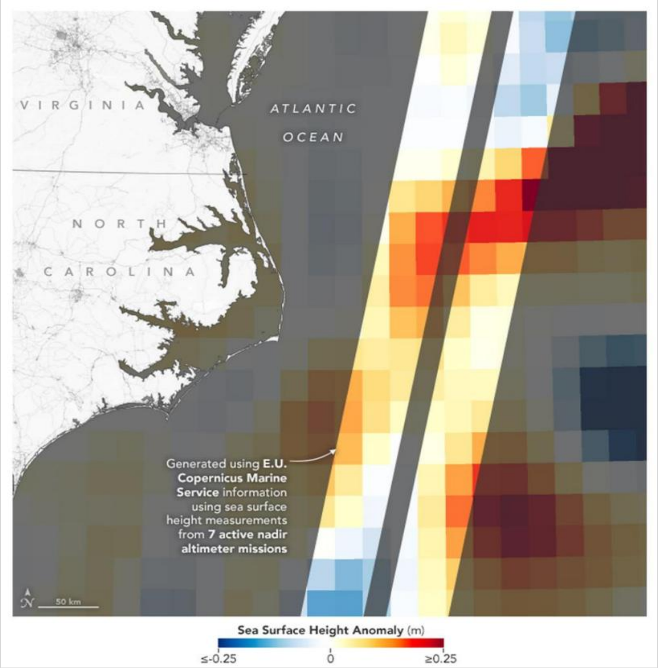
Accurately **decompose SSH snapshots** to understand ocean circulation patterns, enabling:

- **Better** climate predictions
- **Improved** carbon uptake estimates
- **Enhanced** climate model validation

Satellite Revolution

From Traditional Satellite to SWOT

Traditional Altimeter



Difference

Spatial Resolution

~100km → **5-10km**

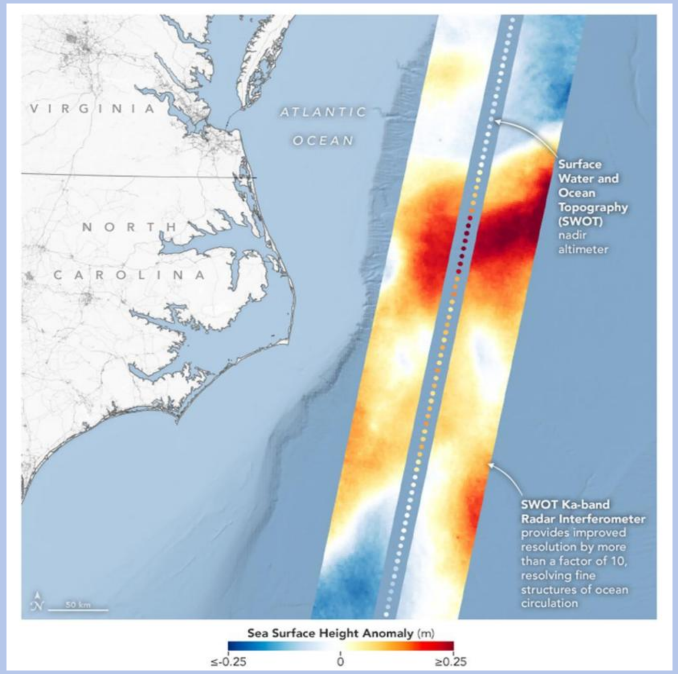
Coverage

1D Tracks → **2D Swath**

Temporal Resolution

Daily → **21 Days**

SWOT Satellite



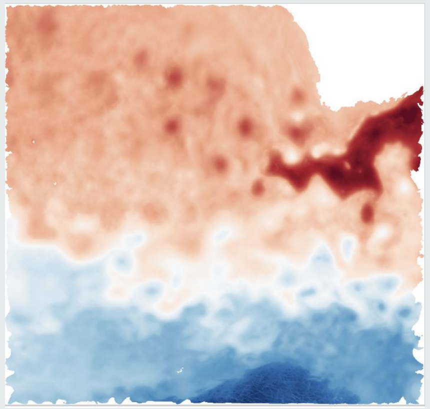
New Challenge:

Longer repeat time (21 days) requires new methods for analyzing single snapshots

Decomposing SSH

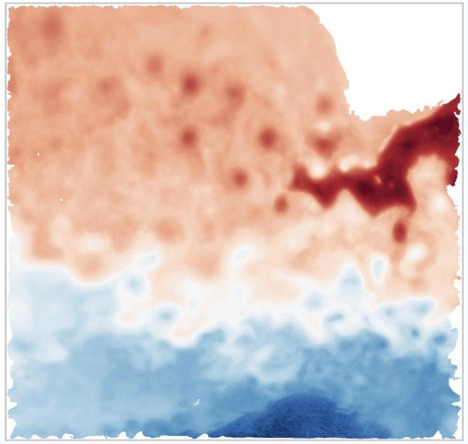
SSH = BM + UBM

Total SSH



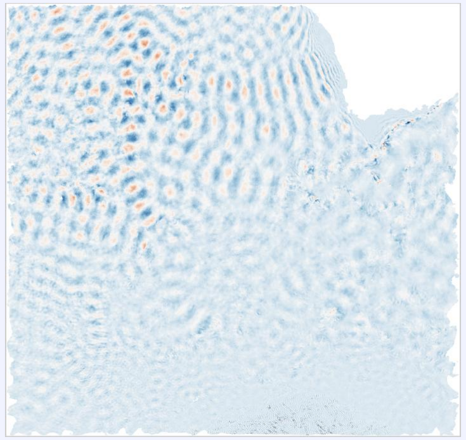
BM & UBM **coexist** at SWOT scales

Balanced Motion (BM)



- **Large-scale** circulation
- **Slow** evolution
- Critical for climate

Unbalanced Motion (UBM)



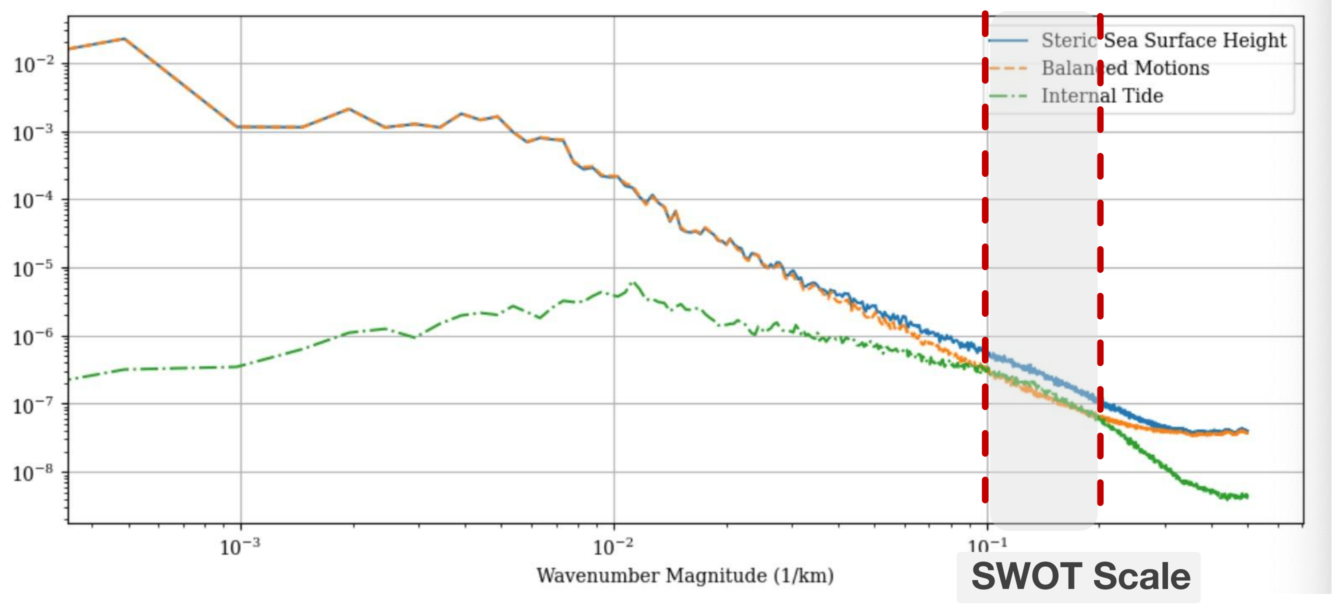
- **Small-scale** waves
- **Rapid** evolution
- Internal waves & tides

Require Separation for Analysis

Challenge: How to seperate BM & UBM from a **single** SSH snapshot?

Challange: A Multi-Scale Problem

Power Spectral Density Analysis



Key Challenges

Signal overlap

BM/UBM coexist at SWOT scales

Power law decay

PSD decays exponentially with wavenumber

Current Solution: Gradient-Enhanced U-Net

A Step Forward, But Not Enough

State-of-the-Art Approach

U-Net + Gradient Loss

$$\mathcal{L}(\eta_{UB}, \tilde{\eta}_{UB}) = \|\eta_{UB} - \tilde{\eta}_{UB}\|^2 + \alpha\|\nabla\eta_{UB} - \nabla\tilde{\eta}_{UB}\|^2$$

α: gradient loss weight (requires tuning)

Key Limitations

- Heavy computational overhead
- Sensitive to α hyperparameter tuning
- High-frequency noise in prediction

Required: A More Fundamental Solution

1. Self-adaptive framework that eliminates gradient regularization tuning
2. Enhanced spectral fidelity at high wavenumbers without noise artifacts

ZCA Whitening: A Fundamental Solution

ZCA Core Process

1. **Feature Decorrelation**
   Separate features across different spatial scales
2. **Power Normalization**
   Normalize power across all frequencies
3. **Spatial Recovery**
   Return to original coordinate space

Key Advantages

- **Enhanced Feature Detection**
  Reveals previously ignored small-scale features
- **Improved Separation**
  Clear BM/UBM distinction at all scales
- **Simplified Training**
  Standard MSE loss becomes effective

Innovation: ZCA transforms multi-scale learning problem into standard learning task

ZCA Whitening: Mathematical Formula

1. Data Preprocessing

1 Input Matrix

$$X_{\eta_{\mathrm{UB}}} \in \mathbb{R}^{n \times d}$$

n: number of samples
d: 108 × 108 (spatial dimensions)

2 Center Data

$$X^c_{\eta_{\mathrm{UB}}} = X_{\eta_{\mathrm{UB}}} - \mu_{\eta_{\mathrm{UB}}}$$

μ: mean across UBM samples

3 Compute Covariance

$$\Sigma = \frac{1}{n-1}(X^c_{\eta_{\mathrm{UB}}})^T X^c_{\eta_{\mathrm{UB}}}$$

Captures feature relationships

2. ZCA Transform

1 Eigendecomposition

$$\Sigma = U \Lambda U^T$$

U: eigenvectors
Λ: diagonal eigenvalue matrix

2 Whitening Matrix

$$W_{\mathrm{ZCA}} = U(\Lambda + \epsilon I)^{-1/2} U^T$$

ε: small constant for numerical stability

3 Apply Transform

$$X_{\mathrm{ZCA}} = X^c_{\eta_{\mathrm{UB}}} W_{\mathrm{ZCA}}$$

Final whitened representation

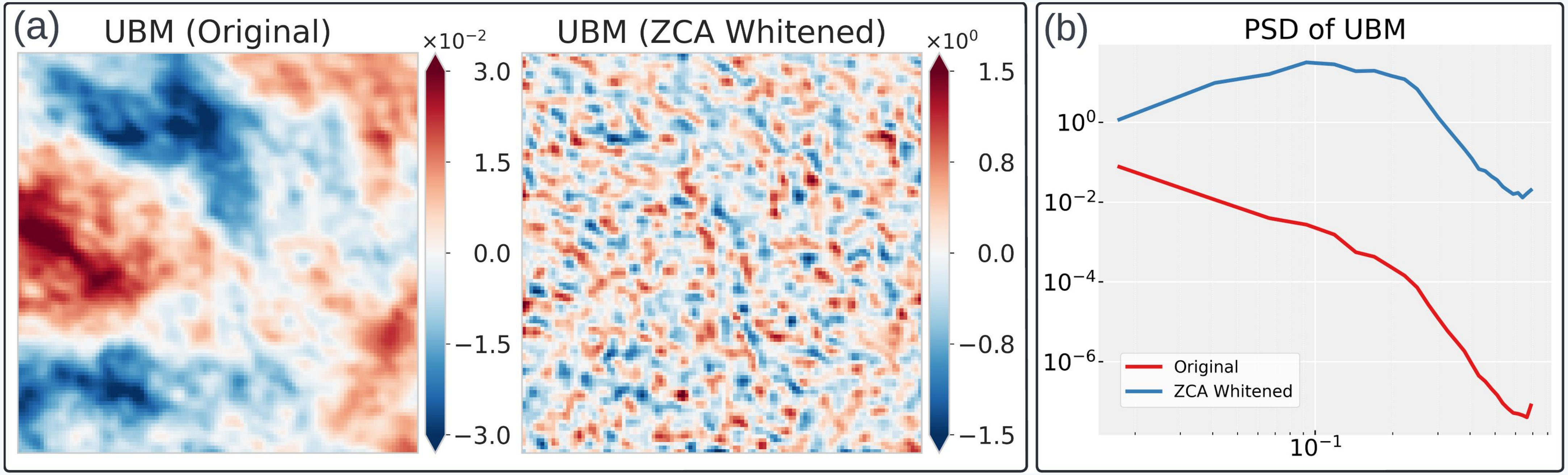
Key Properties

Spectral Equalization
Flat power distribution across all wavenumbers

Decorrelation
Cov(X_ZCA) ≈ I
Features become statistically independent

Invertibility
Exists analytical inverse transformation

ZCA Whitening: Effect Visualization



Original UBM

- **Large-scale** features dominate
- **PSD**: Steep power decay
- **Amplitude scale**: $O(10^{-2})$

After ZCA Whitening

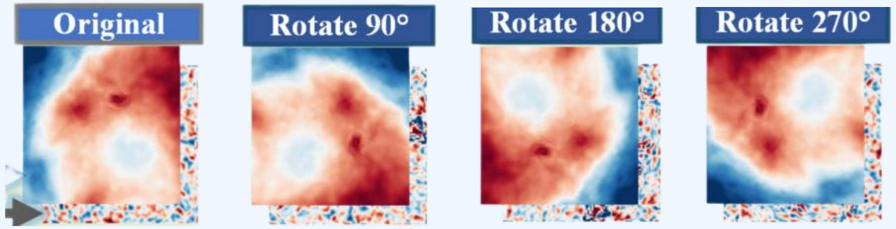
UBM (ZCA Whitened)

- **Small-scale** features visible
- **PSD**: Flatter power distribution
- **Amplitude scale**: $O(10^0)$

Data Augmentation: Solution for Data-Efficiency

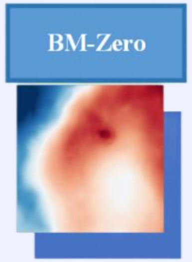
Maximizing Learning from Limited Ocean Data

Rotational Augmentation



- **4 orientations** per sample
- Preserves physical relationships
- Rotation-invariant ocean dynamics

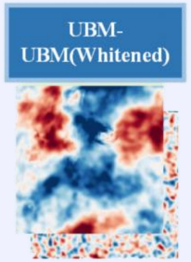
Synthetic Sample Generation



Pure BM Samples

$(\eta_{BM}, 0)$

Large-scale circulation patterns



Pure UBM Samples

$(\eta_{UBM}, \eta_{UBM\_whitened})$

Small-scale wave patterns

Training Enhancement

Original Dataset	After Augmentation	Improvement
24,000	**144,000**	**6×**
training samples	training samples	

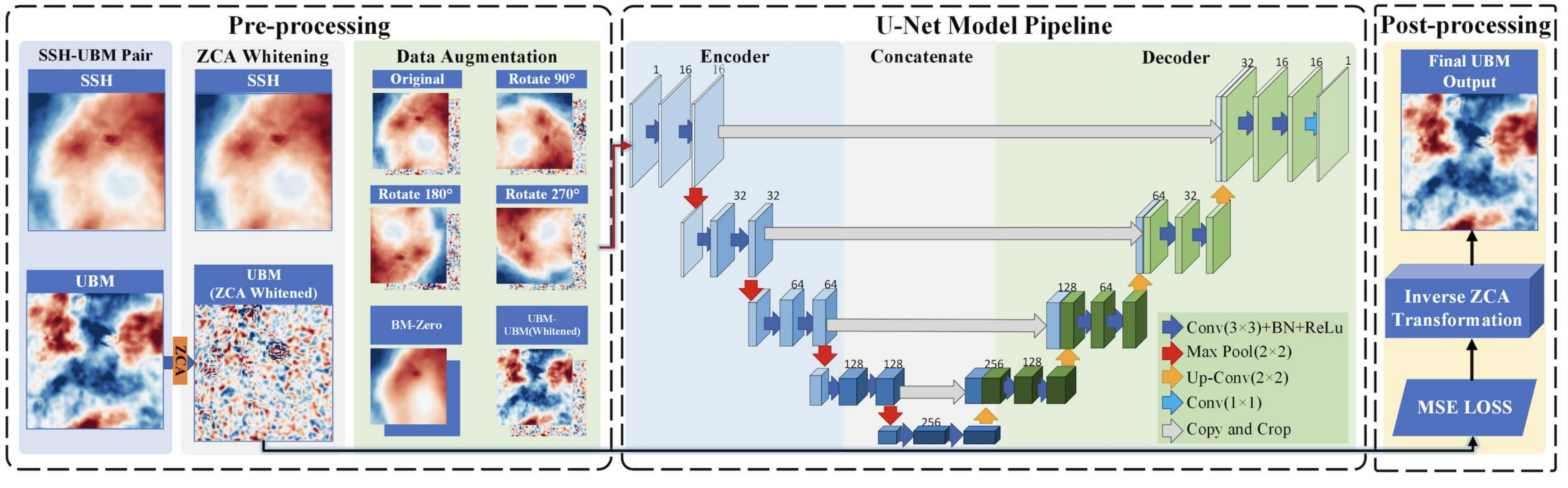
Key Benefits:

↑ Model robustness

↑ Feature recognition

↑ Generalization

Our Approach Overview



Key Innovations

1 ZCA Whitening

Enhances high frequency information capture

2 Data Augmentation

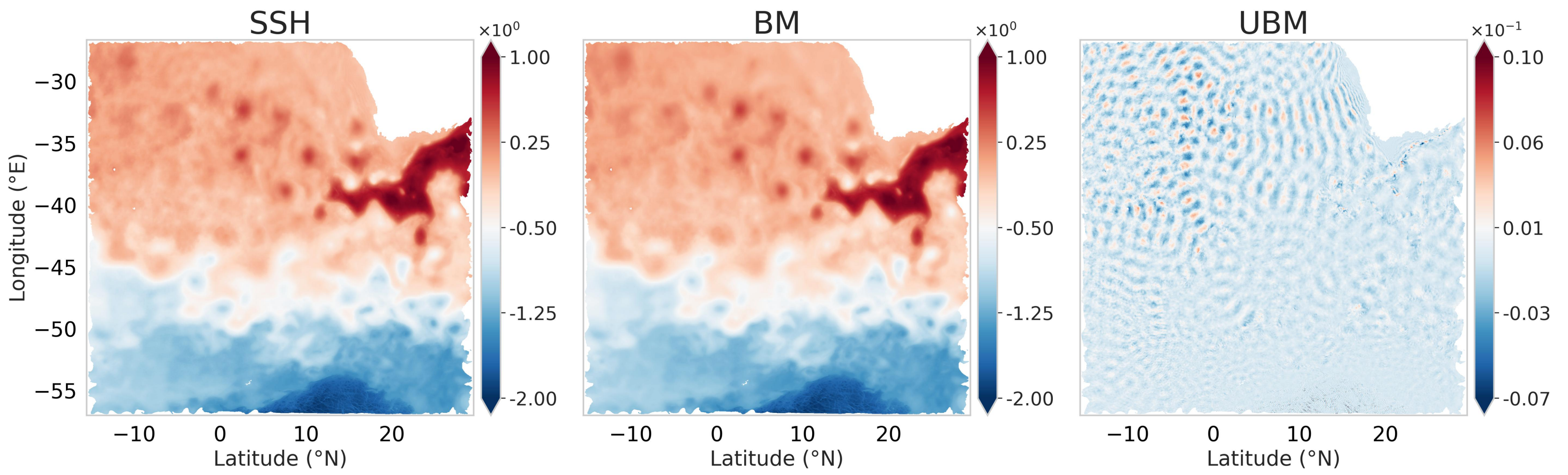
Mitigates limited training samples

3 U-Net

Multi-scale feature processing

Dataset: LLC4320 Ocean Simulation

High-Resolution Global Ocean Model for SSH Decomposition



Dataset Overview

Temporal Coverage

70 daily snapshots
(Sept-Nov 2011)

Spatial Coverage

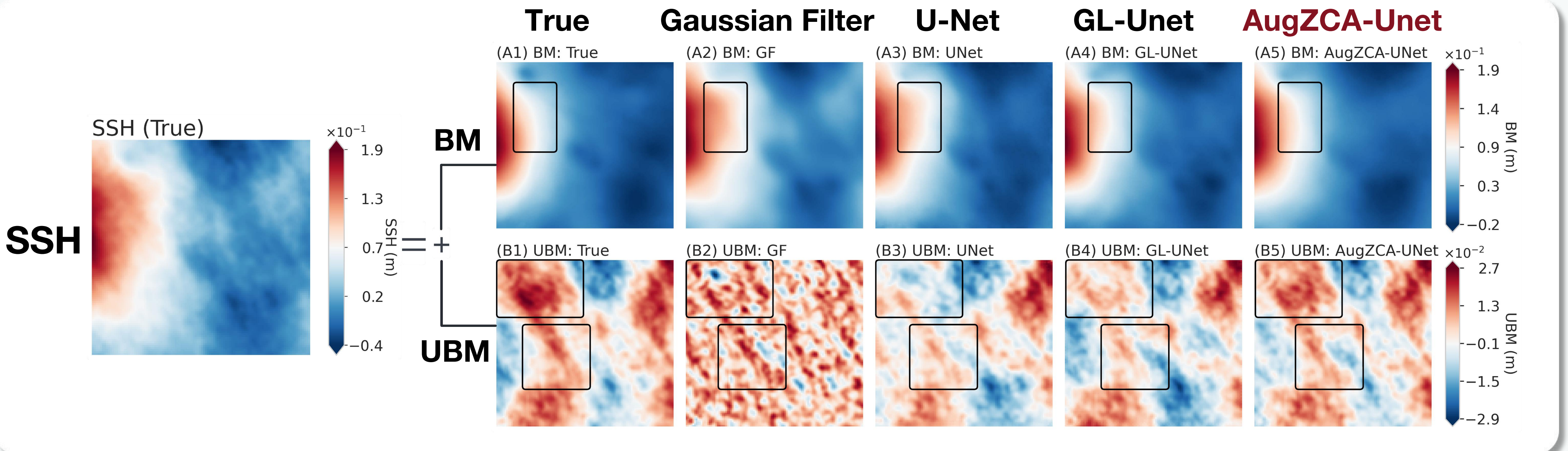
Agulhas retroflection region
(15° W-29° E, 27° S-57° S)

Resolution

Temporal: 1 Day
Spatial: ~2km

Results: Visual Comparison

Model Performance Across Different Methods



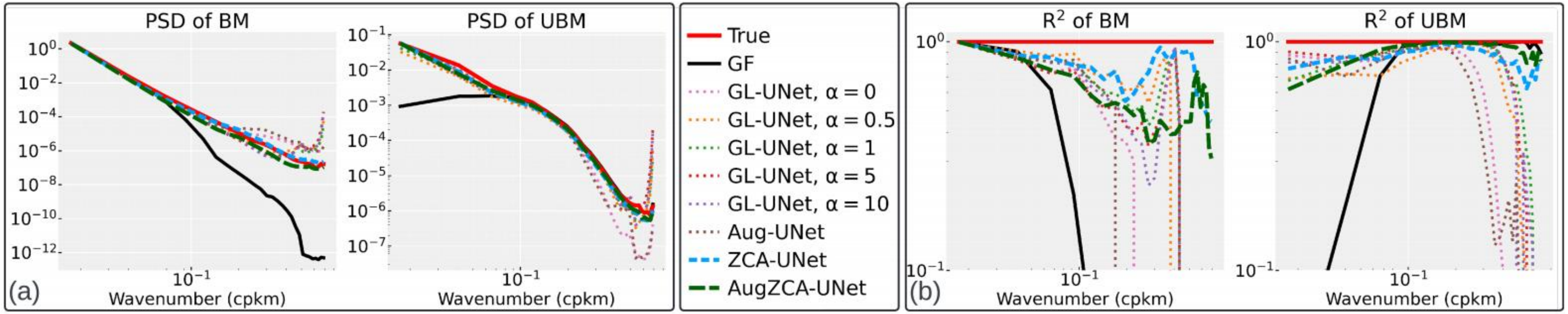
Balanced Motion (BM)

- Preserves **large-scale** features without over-smoothing

Unbalanced Motion (UBM)

- Better recovery of **fine-scale** structures

Results: Spectral Analysis



Low Wavenumbers

- All methods perform well for large-scale features

High Wavenumbers

- AugZCA-UNet maintains accuracy at small scales

Results: Performance Metrics

Quantitative Analysis of Model Performance

Table 1: Pixel-wise Absolute Error Distribution Measures for $\hat{\eta}_{\mathrm{B}}$ ($\times 10^{-2}$) and $|\nabla \hat{\eta}_{\mathrm{B}}|$ ($\times 10^{-3}$)

	Measures	GF	GL-UNet					Aug-UNet	ZCA-UNet	AugZCA-UNet
			$\alpha = 0$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 5$	$\alpha = 10$			
$\hat{\eta}_{\mathrm{B}}$	Median	0.406	0.404	0.420	0.397	0.398	0.406	0.384	0.393	**0.372**
	P95 (95%)	1.68	1.52	1.49	1.48	1.51	1.53	1.50	1.52	**1.40**
$\lvert\nabla \hat{\eta}_{\mathrm{B}}\rvert$	Median	0.505	0.464	0.447	0.450	0.450	0.453	0.460	0.450	**0.427**
	P95 (95%)	2.20	1.80	1.70	1.72	1.74	1.75	1.78	1.71	**1.64**

Error Reduction

↓ **8.4%** BM Error Reduction
vs GL-UNet

↓ **15.4%** Gradient Error Reduction
vs GF

Key Advantages

- **No** gradient loss tuning needed
- **Consistent performance** across scales
- **Lower error** in both BM and gradients

Conclusion

Key Contributions

- AugZCA-UNet: Superior multi-scale decomposition without gradient tuning
- ZCA whitening enables consistent cross-scale performance
- Effective processing of SWOT satellite data

Future Work

- ❑ Address ZCA memory scaling for larger images
- ❑ Investigate transfer learning for generalization
- ❑ Incorporate physical constraints in ML frameworks