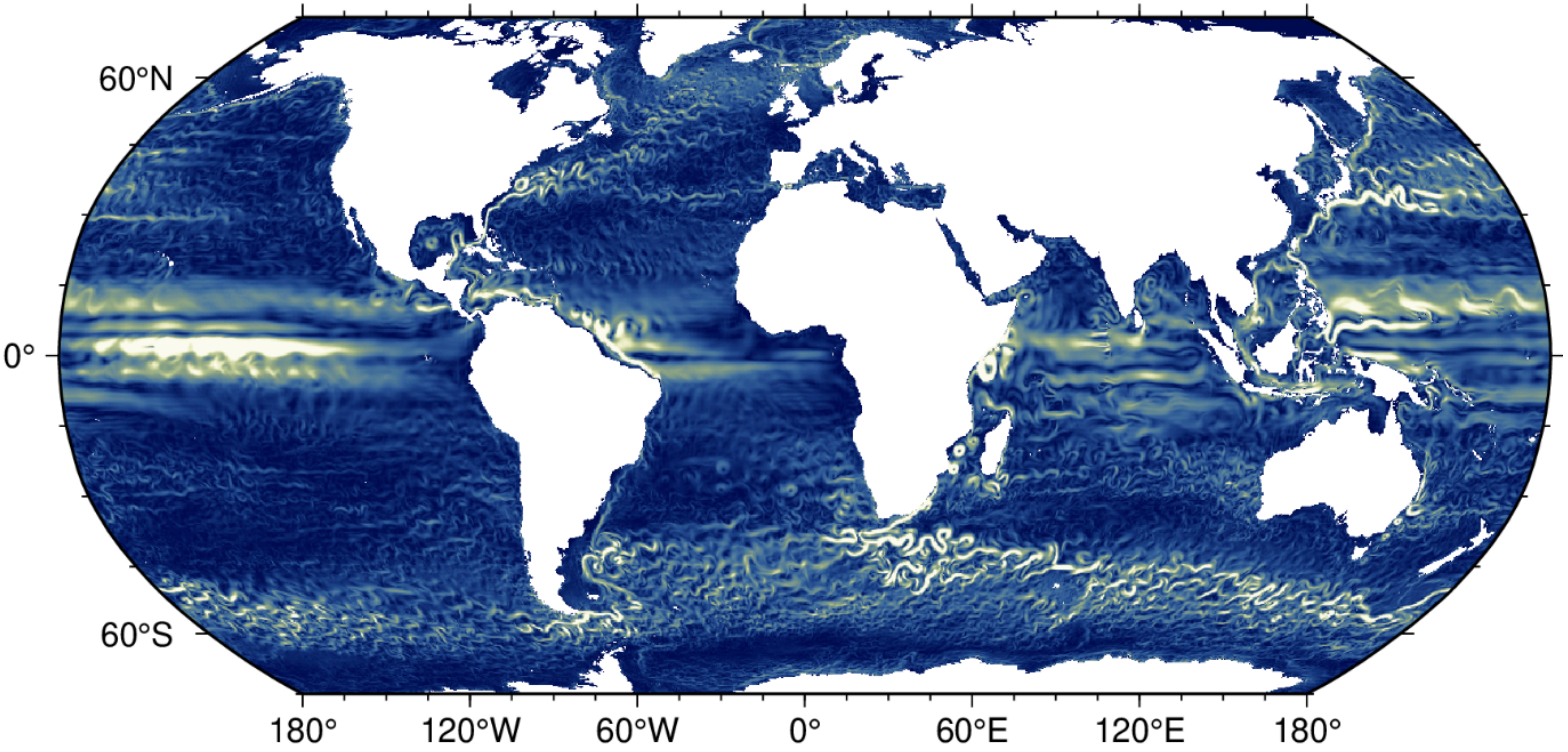
Inferring the ocean subsurface using generative AI

Andre Souza

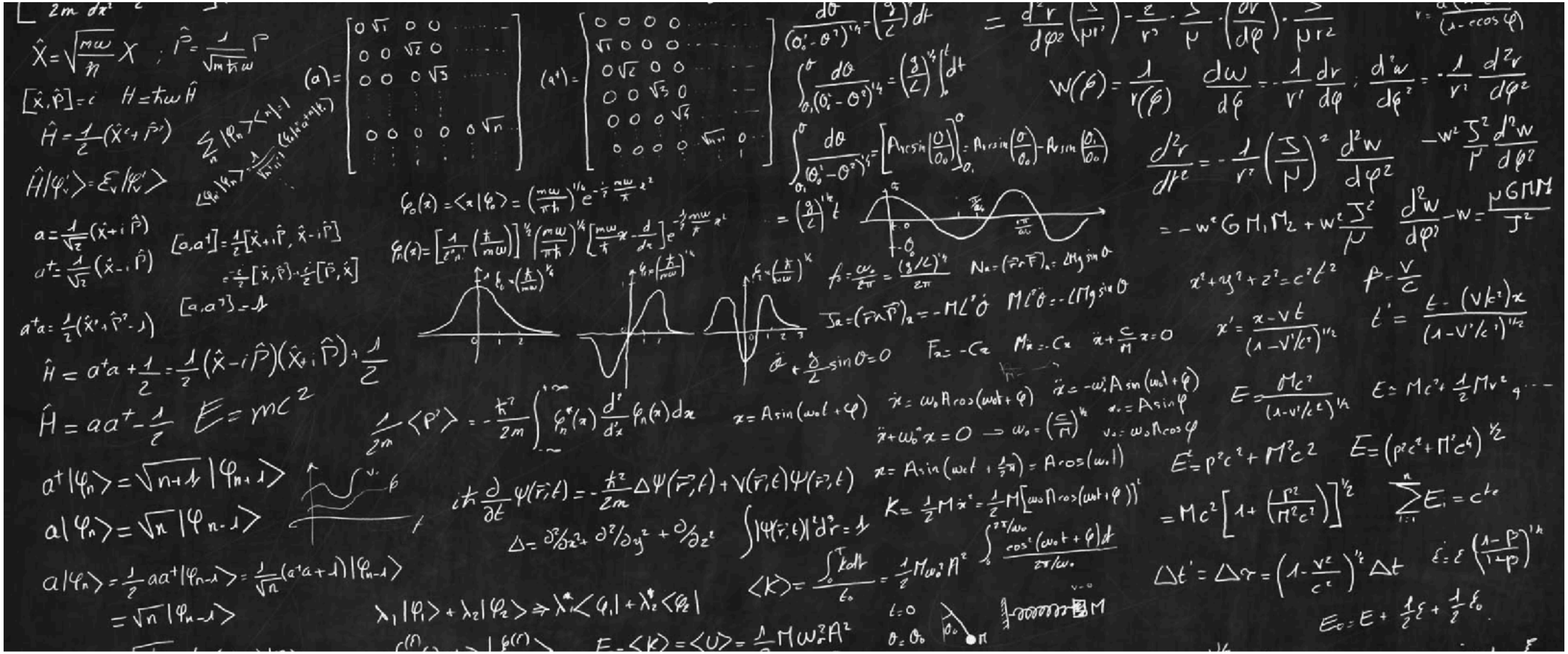
Recap

We’ve seen Satellites

We’ve seen Ocean



We’ve seen Models



We’ve seen various ways of combining models with data

We’ve seen various ways of combining models with data

This talk will discuss (yet) another way

Combining Models with Data

- Physical theories let us make inferences using less information (Maxwell's equations)
- For complex systems (atmosphere, ocean, chaotic systems) the best we can do are numerical models (computational irreducibility)
- Using computational models with data-driven approaches allow us to extend the theory + data inferences

Combining Models with Data

- Physical theories let us make inferences using less information (Maxwell’s equations)
- For complex systems (atmosphere, ocean, chaotic systems) the best we can do are numerical models (computational irreducibility)
- Using computational models with data-driven approaches allow us to extend the theory + data inferences

What kind of questions do we ask?

Information we want

$$y = f(x)$$

A mapping

Information we have

What kind of questions do we ask?

Information we want

$$\rho(y|x)$$

A Statistical Inference

Information we have

What kind of questions do we ask?

$$y = f(x)$$

$$\rho(y|x)$$

$$\rho(y|x) = \delta(y - f(x))$$

What kind of questions do we ask?

$$y = x^2$$

$$\rho(y|x)$$

$$\rho(y|x) = \delta(y - x^2)$$

What kind of questions do we ask?

$$y = x^2$$

$$\rho(x|y)$$

$$\rho(x|y) = \frac{\delta(x - \sqrt{y}\,) + \delta(x + \sqrt{y}\,)}{2}$$

What kind of questions do we ask?

Useful Information

$$\rho(y|x) = \delta(y - f(x))$$

"Useless" Information

$$\rho(y|x) = \rho(y)$$

What kind of questions do we ask?

ill-posed

$$y = f(x)$$

well-posed

$$\boxed{\rho(y|x)}$$

What kind of questions do we ask?

ill-posed

$$y = f(x)$$

well-posed

$$\boxed{\rho(y|x)}$$

connection

$$f(x) = \mathbb{E}_y \rho(y|x)$$

What can we do now that we couldn't do before?

$$\rho(y|x)$$

high-dimensional

What can we do now that we couldn't do before?

$$\rho(y|x)$$

Method: Score-Based Diffusion Models

Architecture: Modified U-Net

Applications

$$\rho(y|x)$$

Information we want

Information we have

Application: Ocean Inference

$$\rho(y|x)$$

Ocean Interior

Ocean Surface

What we have

What we want

Application: Ocean Inference

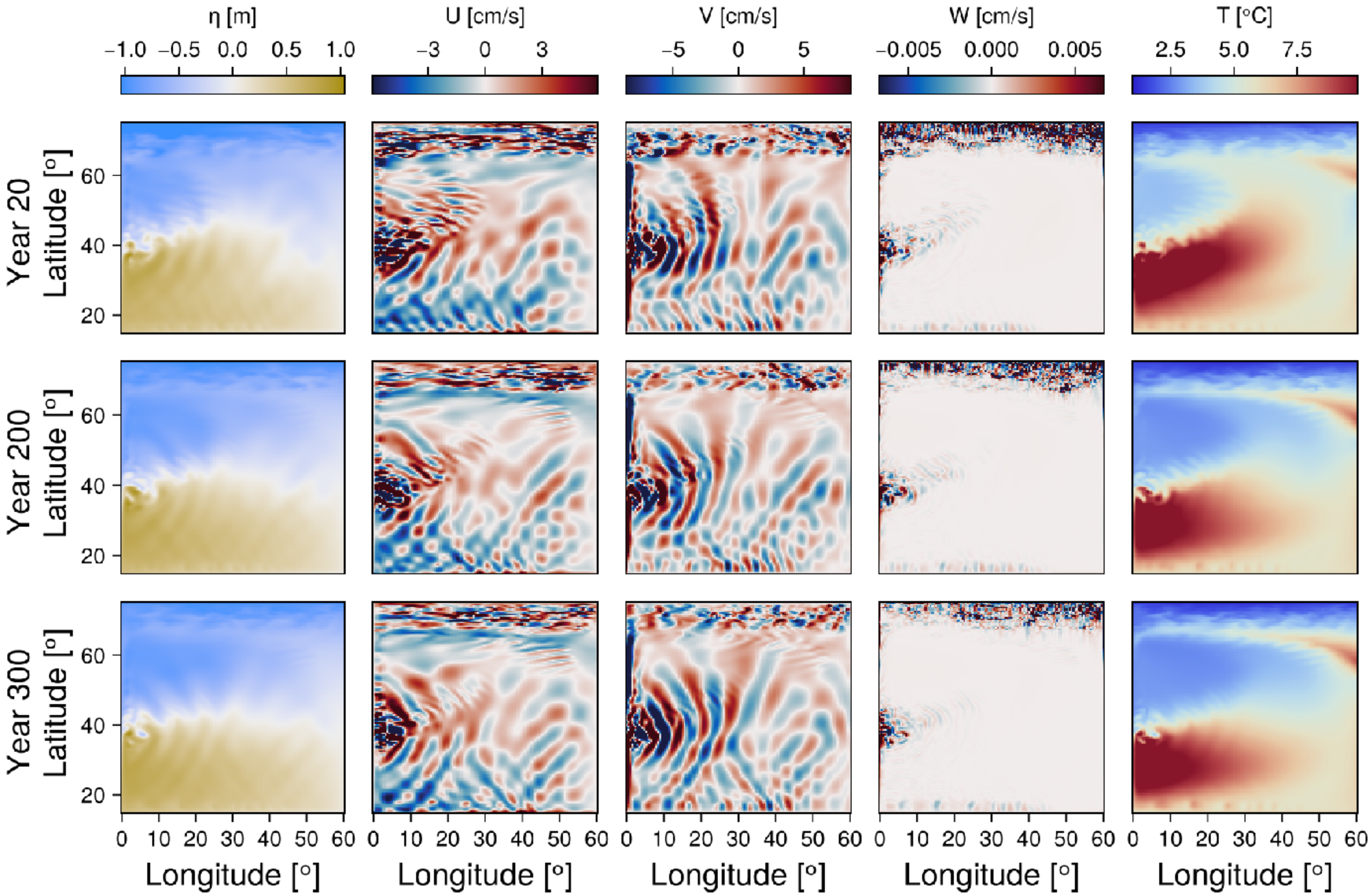
$$\rho(y|x)$$

Model: 3D Navier-Stokes with 256x256x15x3 degrees of freedom (non-stationary)

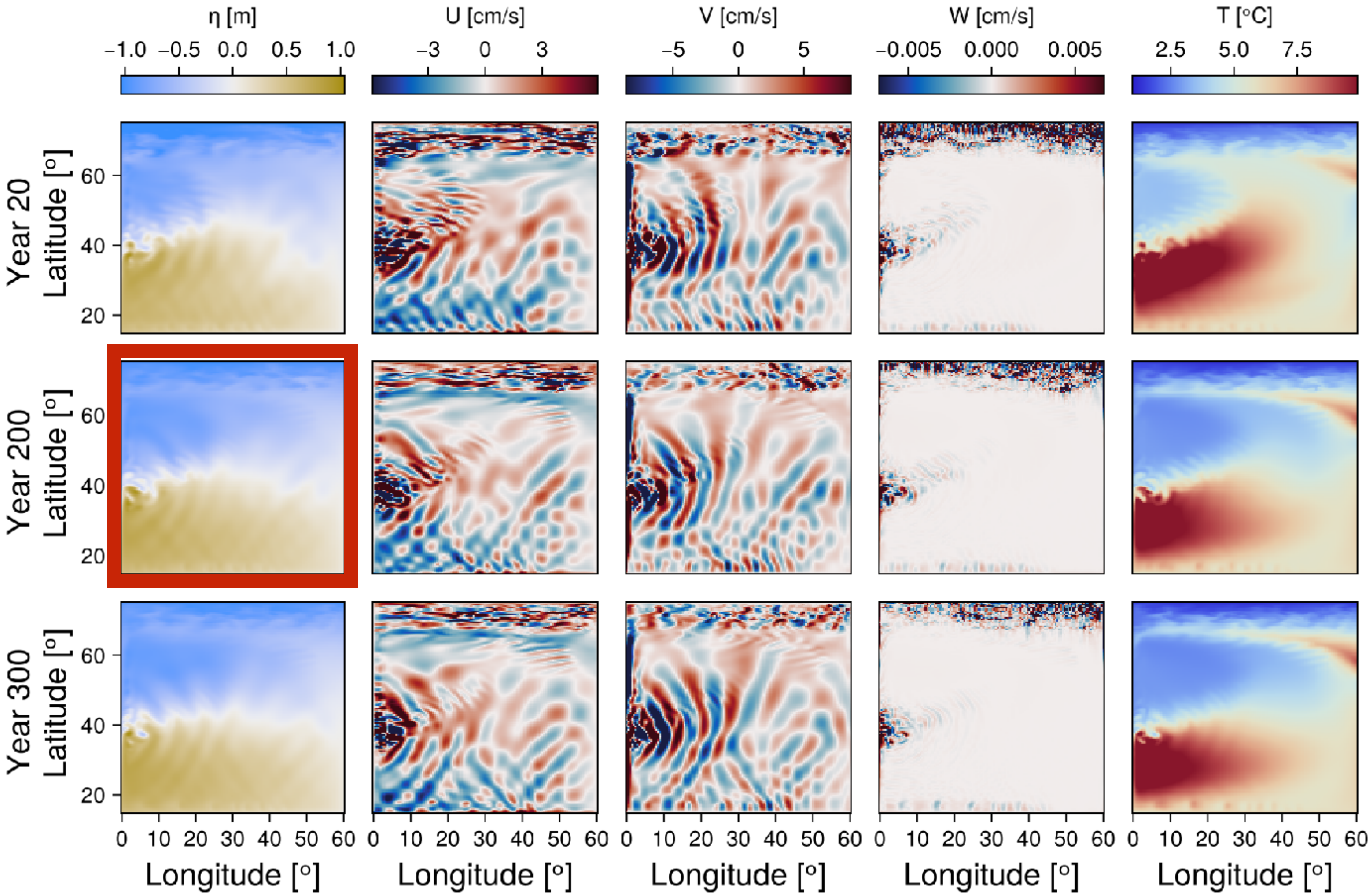
Given information "x": Free Surface Height (different levels of coarse-graining)

Predicted Information "y": The 3D Ocean Interior State

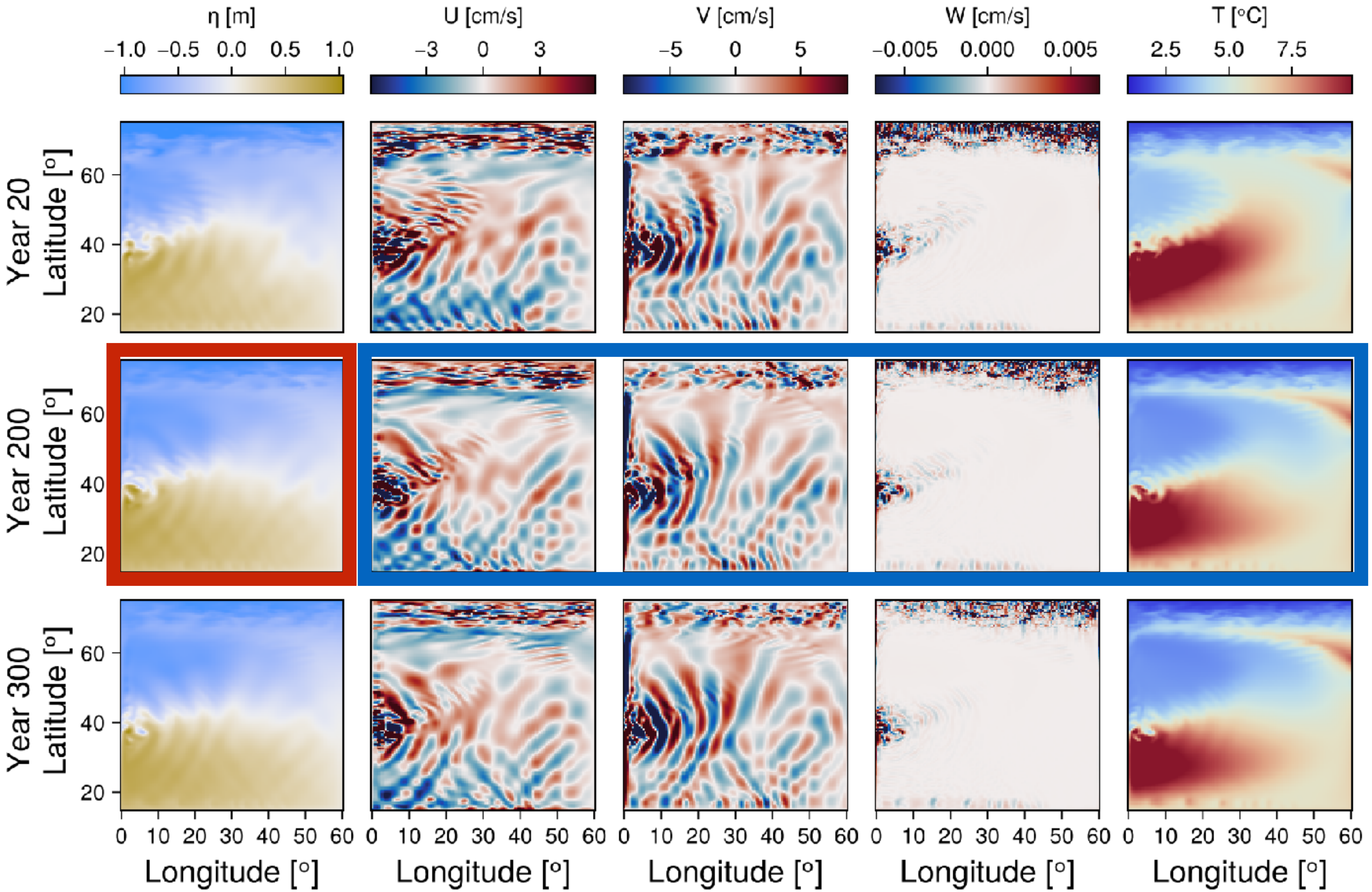
Application: Ocean Inference Task (400m depth)



Application: Ocean Inference Task (400m depth)

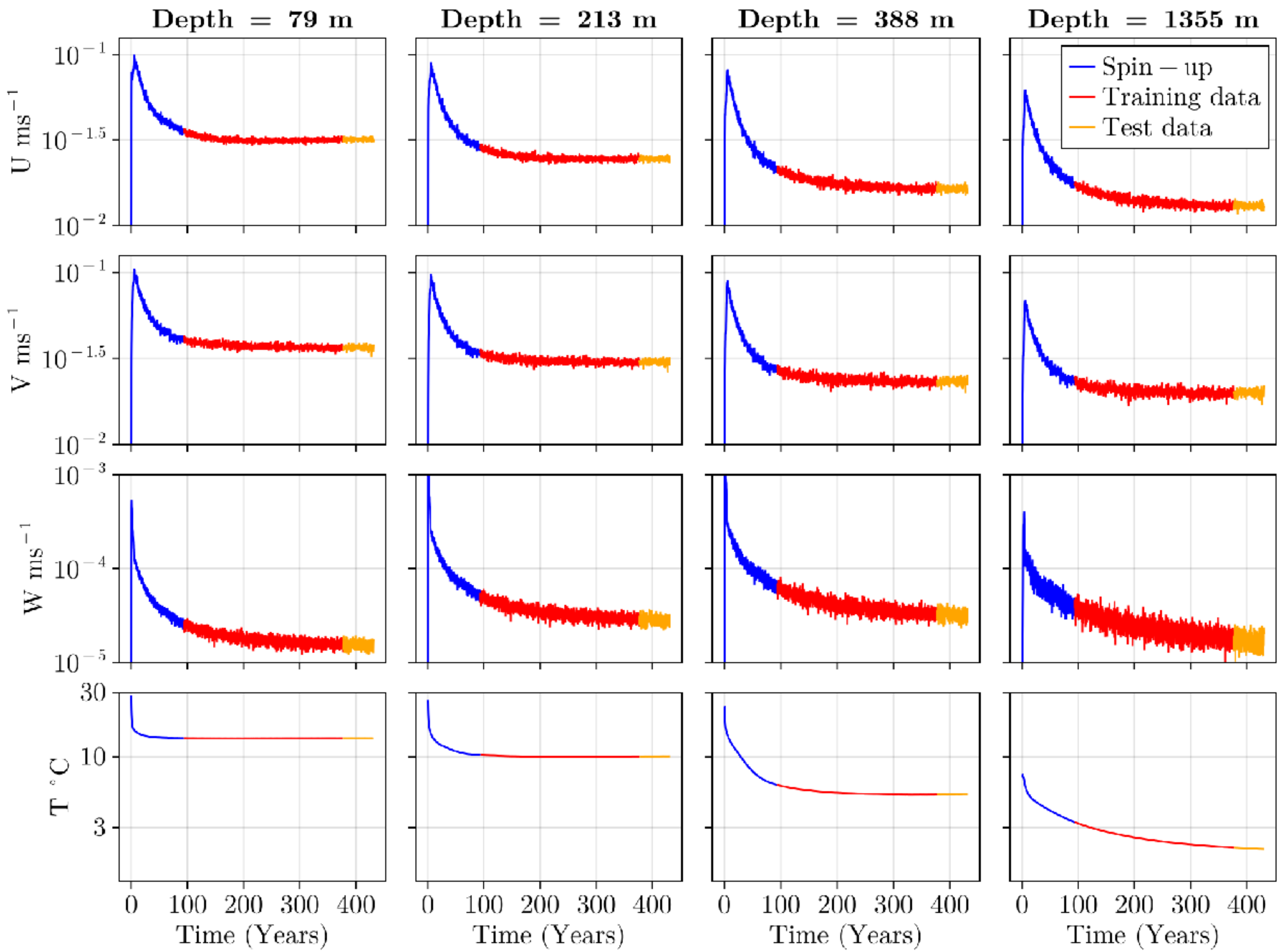


Application: Ocean Inference Task (400m depth)



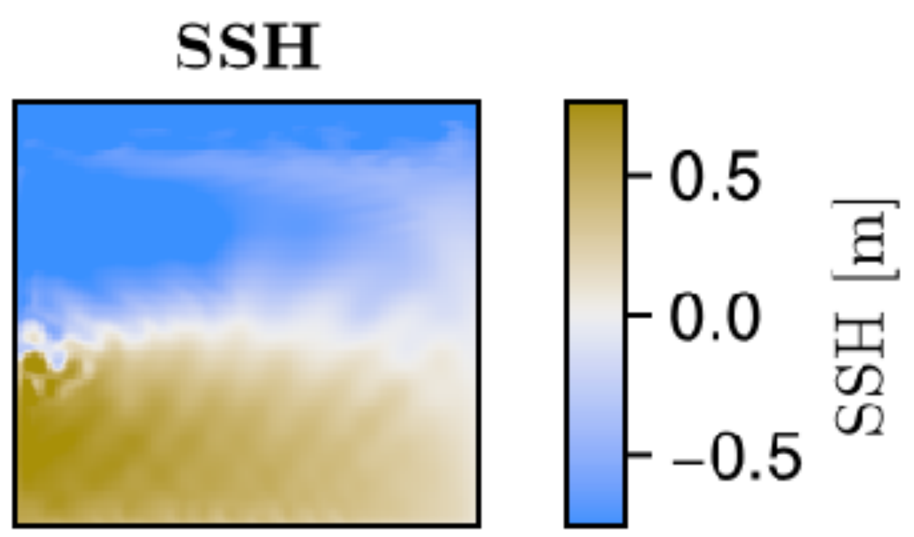
Training and Test Set: Non-stationarity

Application: Ocean Inference

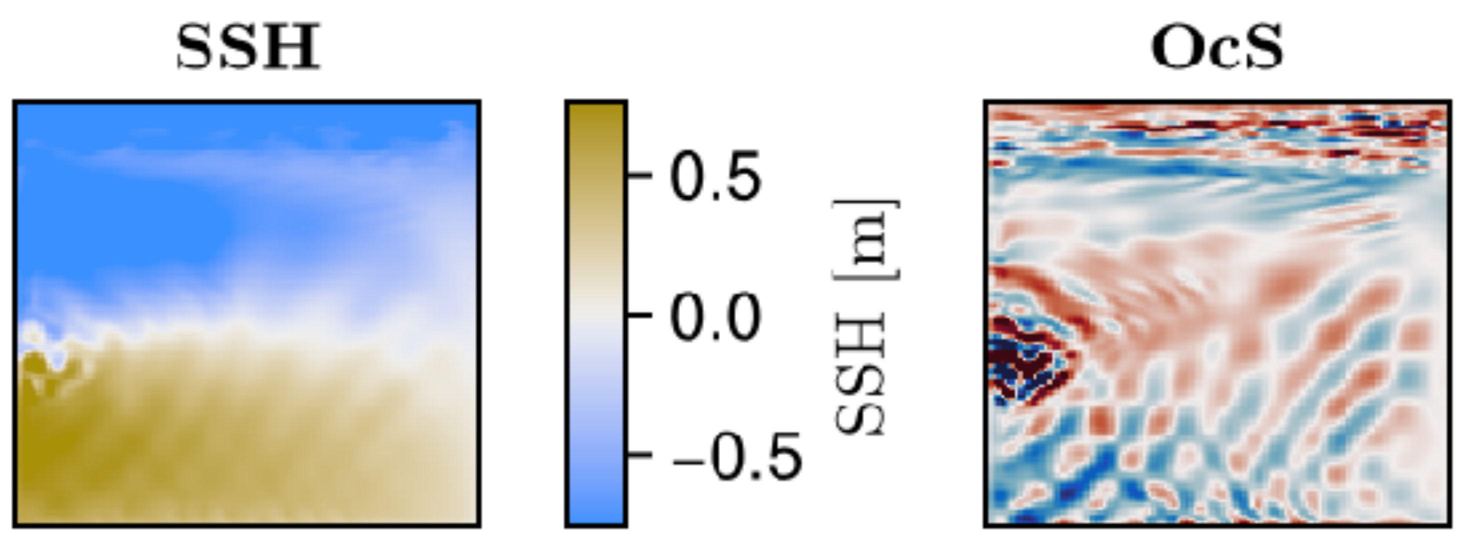


Results

Application: Ocean Inference Result (400m depth)



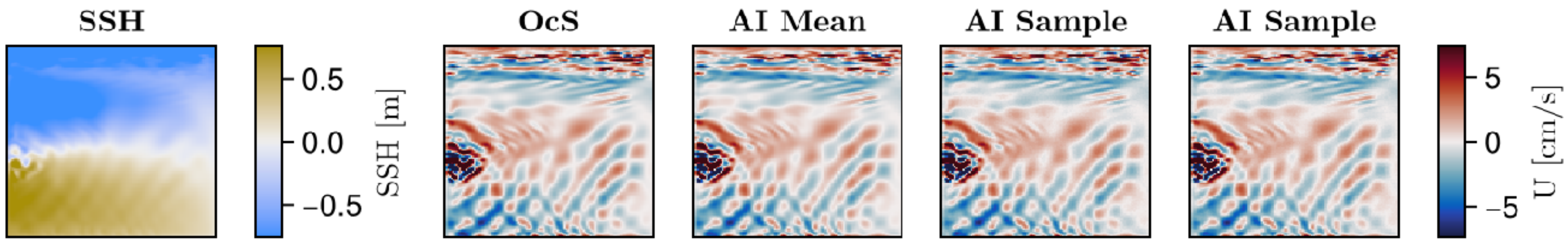
Application: Ocean Inference Result (400m depth)



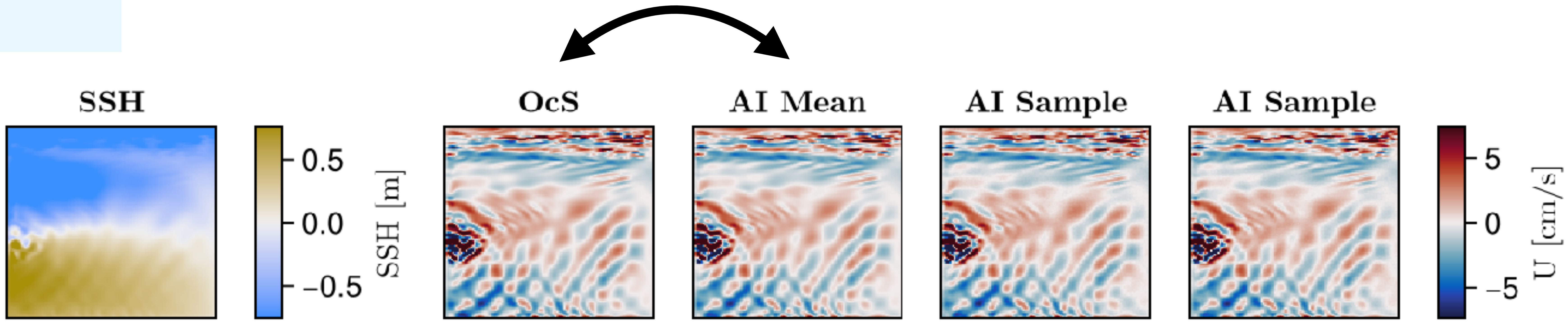
Application: Ocean Inference Result (400m depth)

Application: Ocean Inference Result (400m depth)

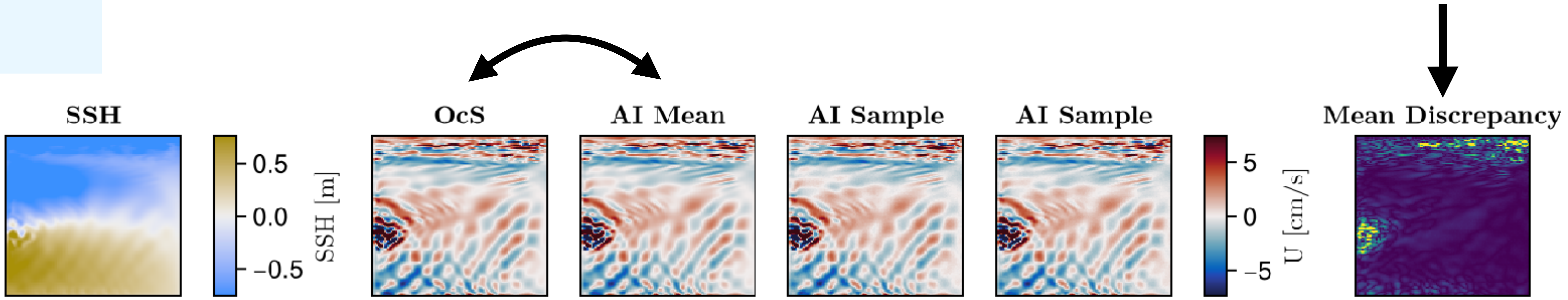
Application: Ocean Inference Result (400m depth)



Application: Ocean Inference Result (400m depth)



Application: Ocean Inference Result (400m depth)



Application: Ocean Inference Result (400m depth)

Application: Ocean Inference Result (400m depth)

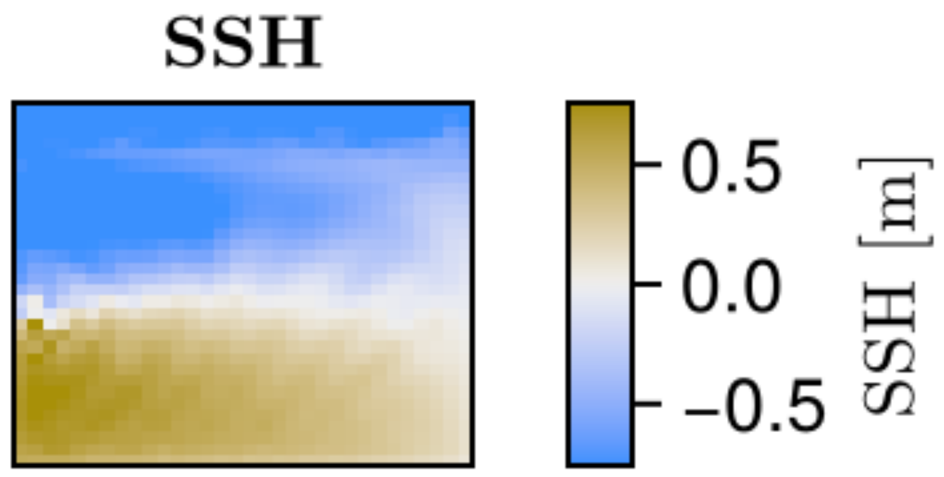
Application: Ocean Inference Result (400m depth)

Application: Ocean Inference Result (400m depth)

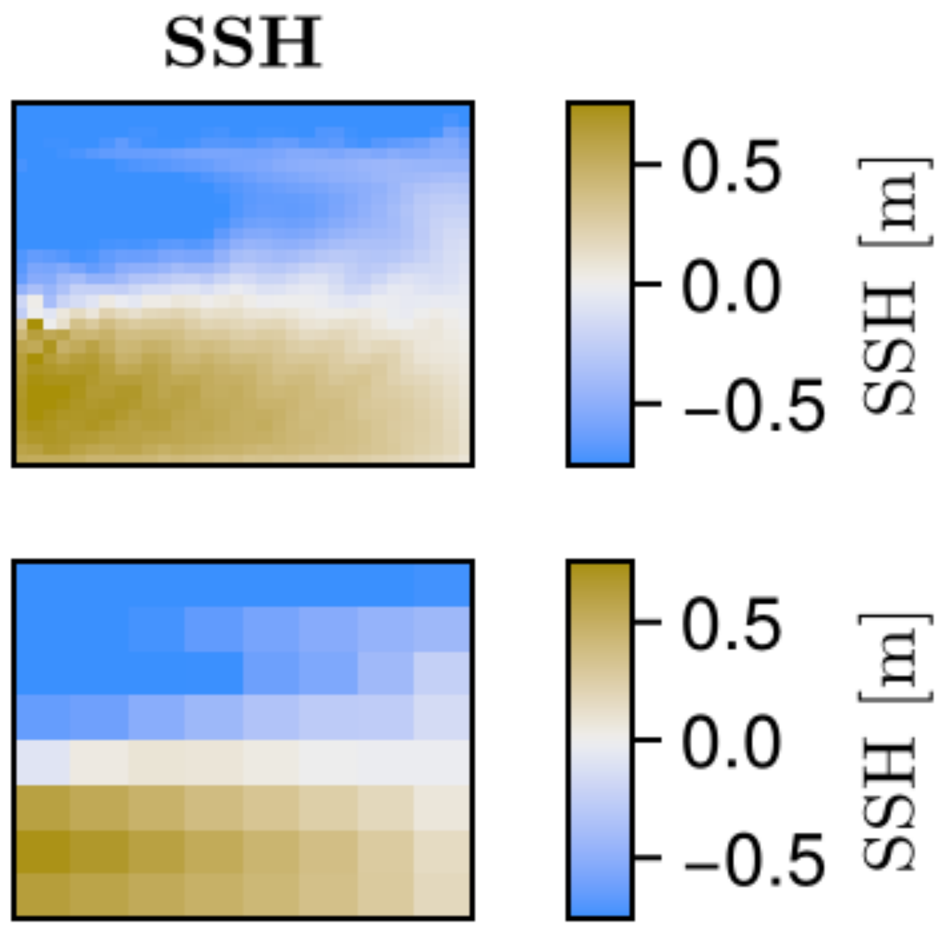
Application: Ocean Inference Result (400m depth)

(with Coarse-Graining)

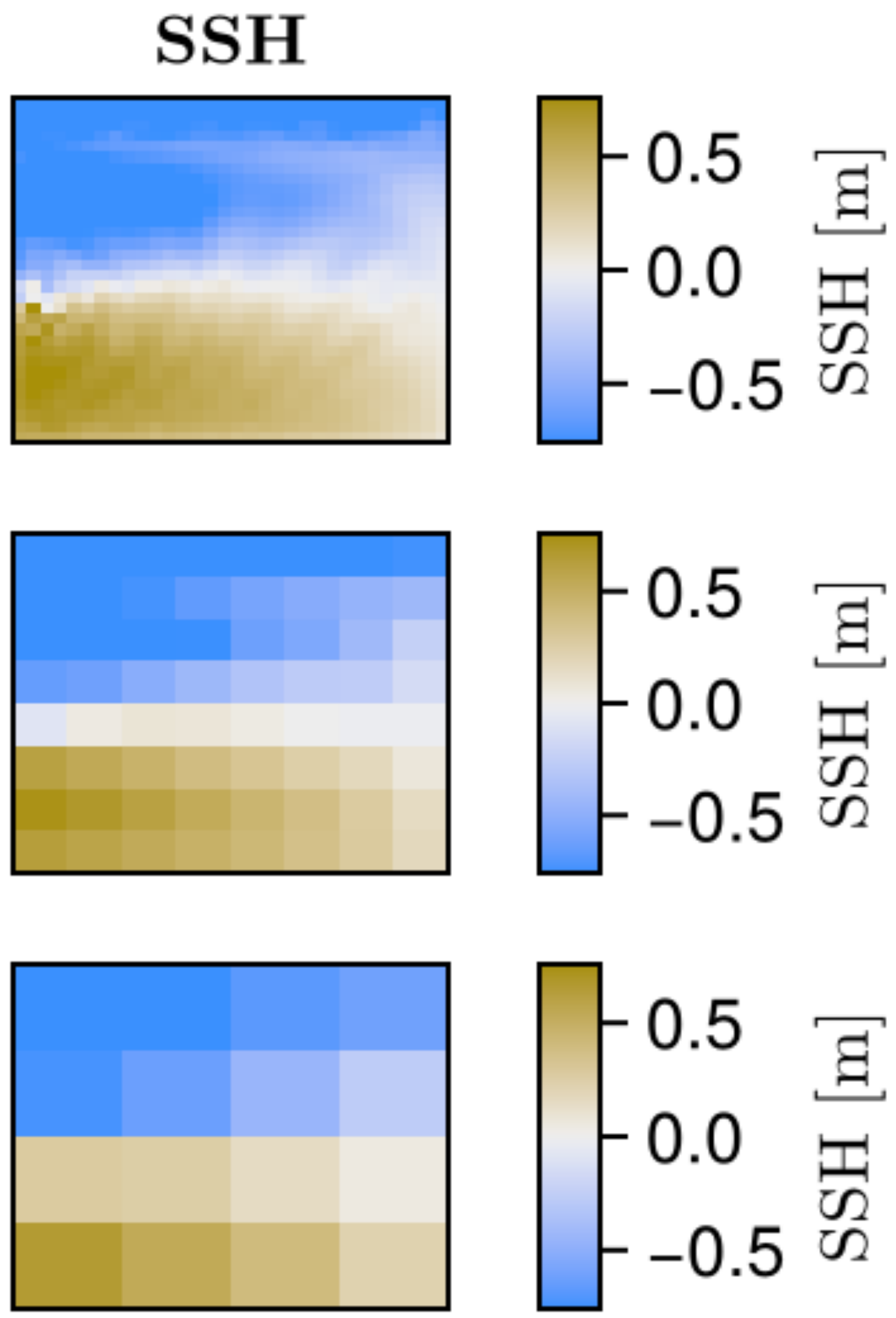
Application: Ocean Inference Result (400m depth)



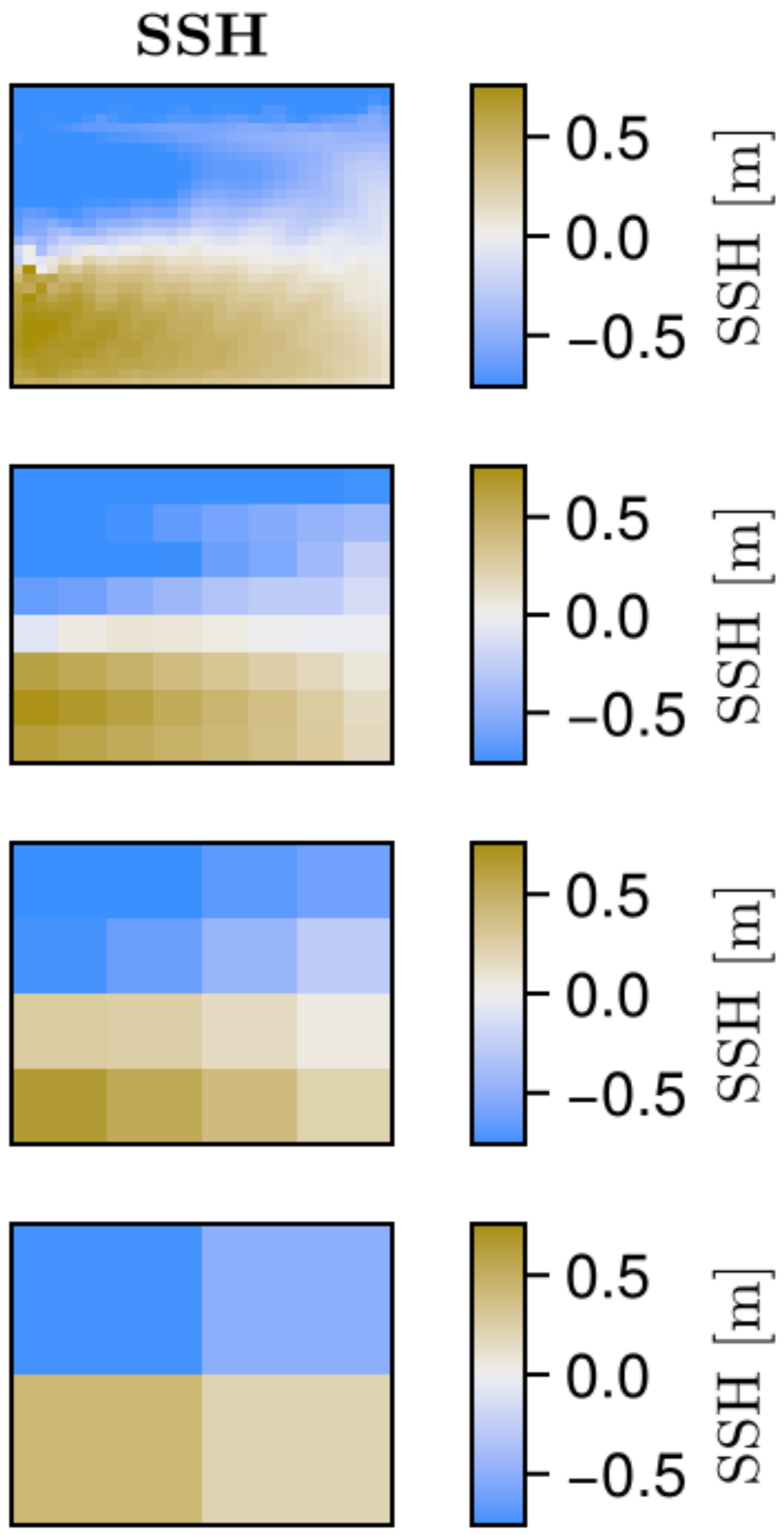
Application: Ocean Inference Result (400m depth)



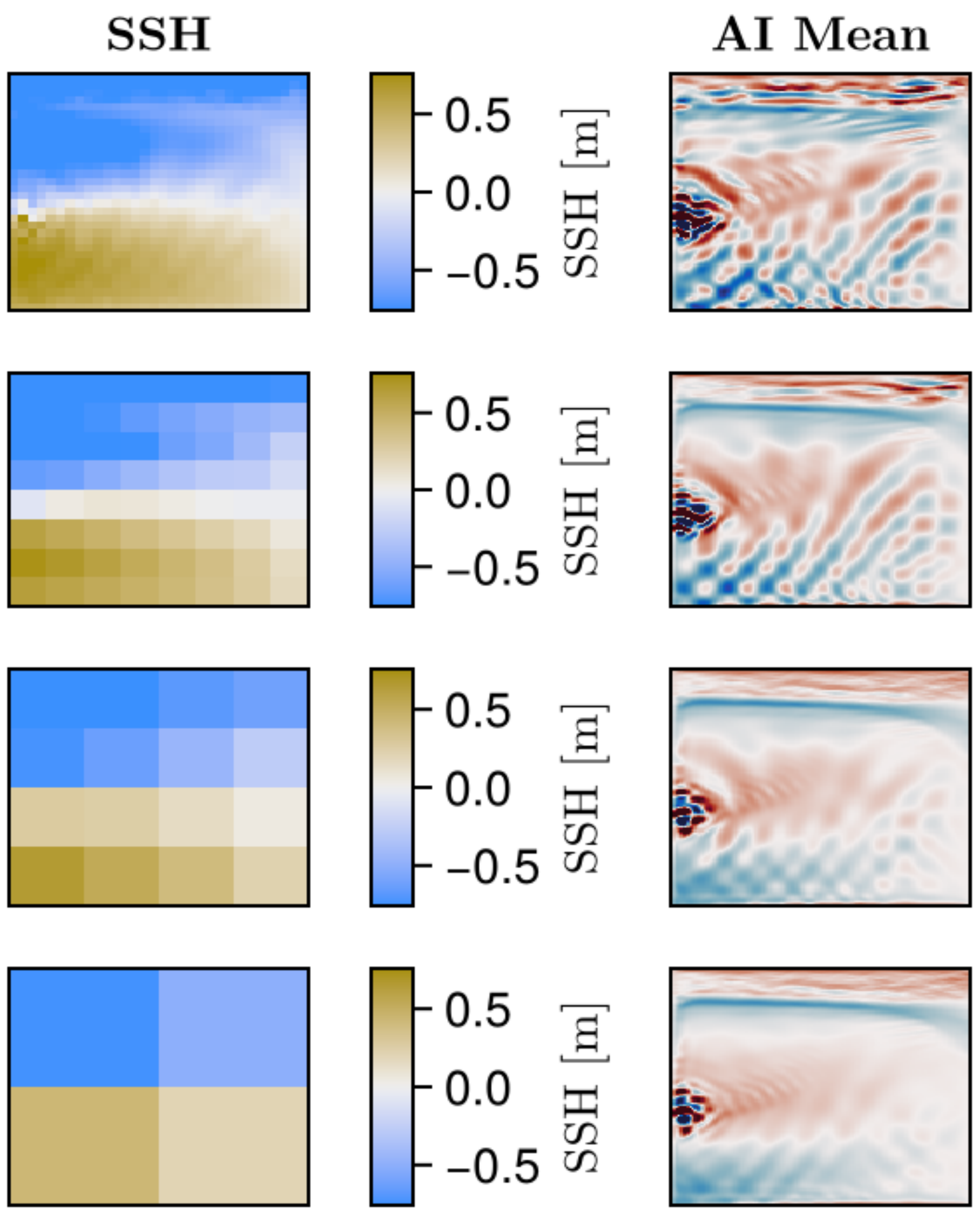
Application: Ocean Inference Result (400m depth)



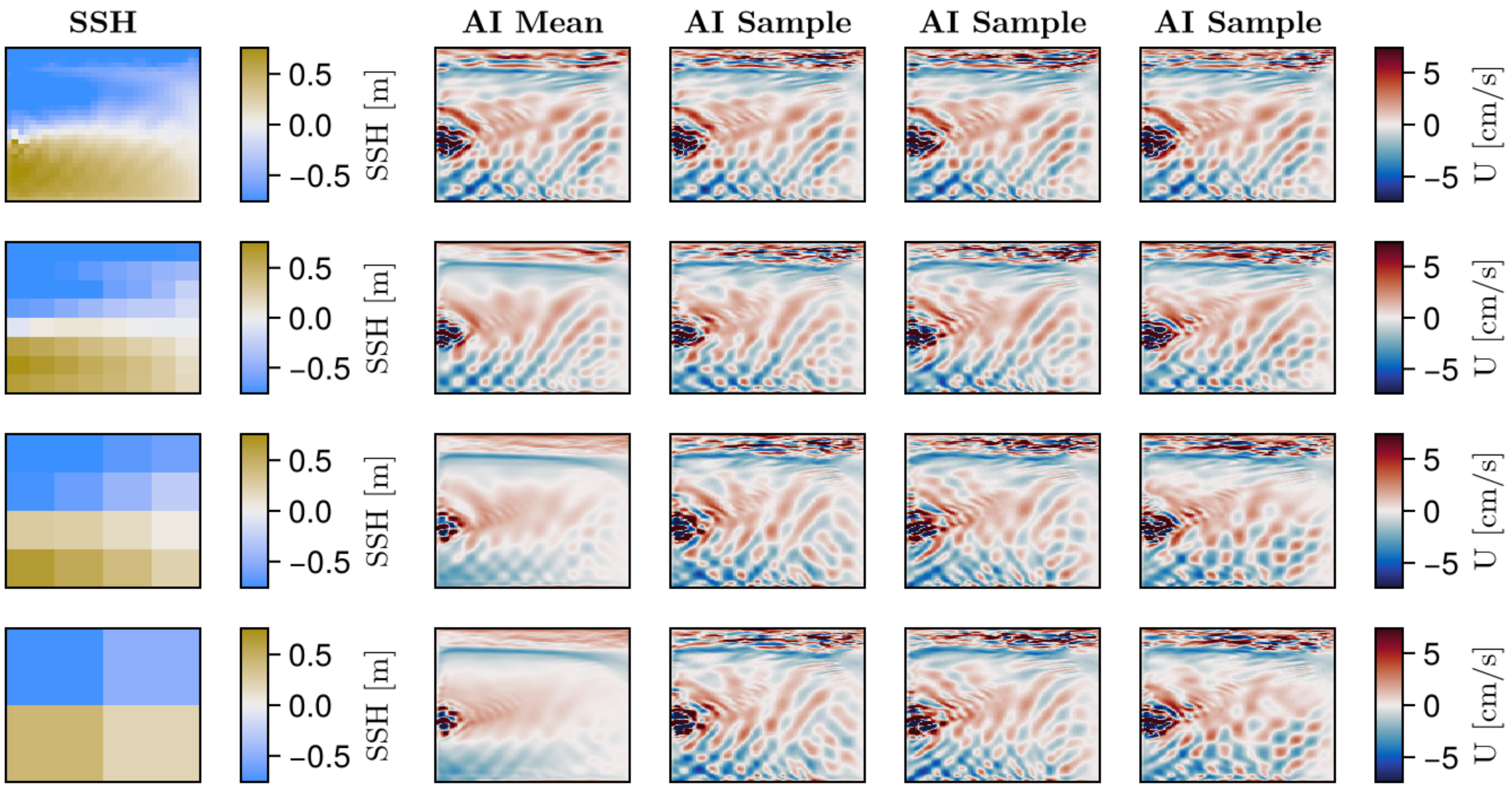
Application: Ocean Inference Result (400m depth)



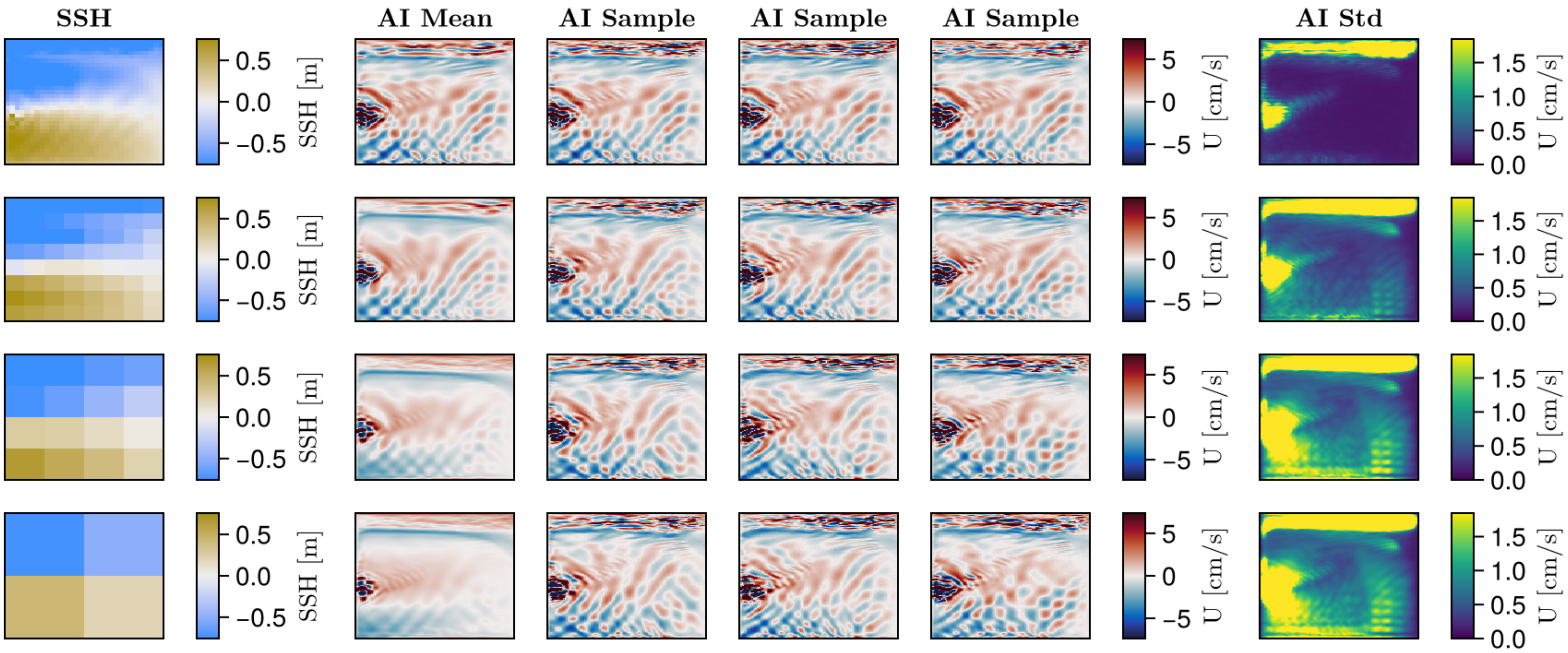
Application: Ocean Inference Result (400m depth)



Application: Ocean Inference Result (400m depth)



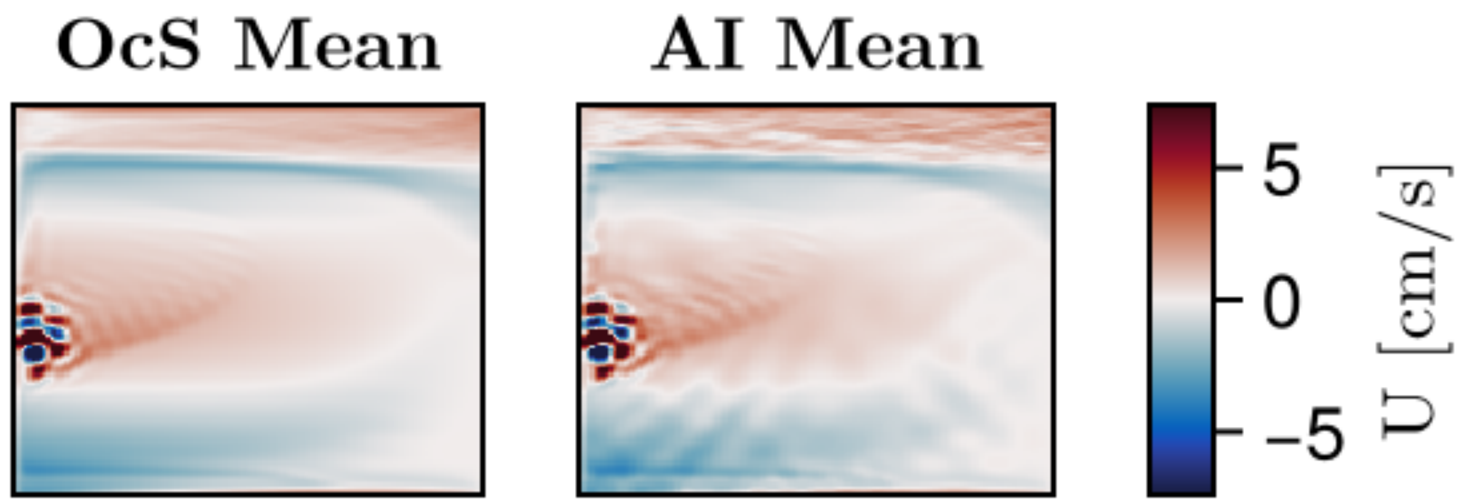
Application: Ocean Inference Result (400m depth)



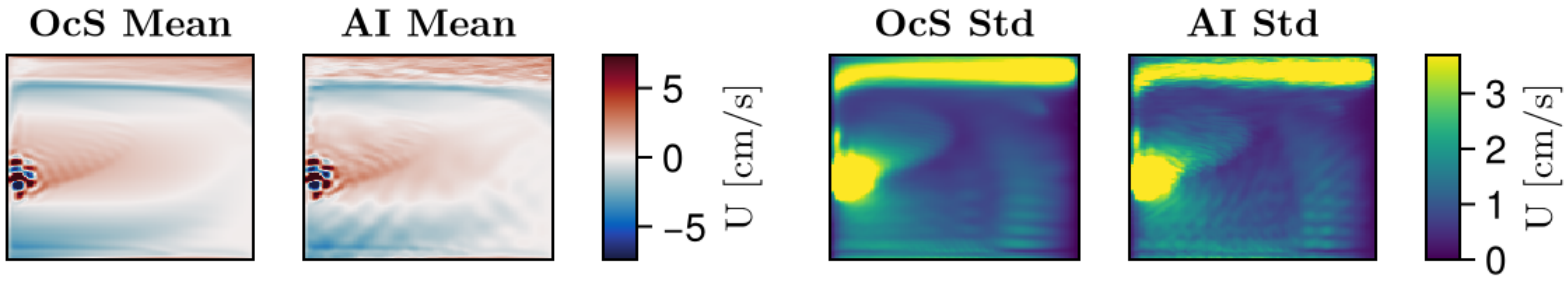
Application: Ocean Inference Result (400m depth)

(with Extreme Coarse-Graining)

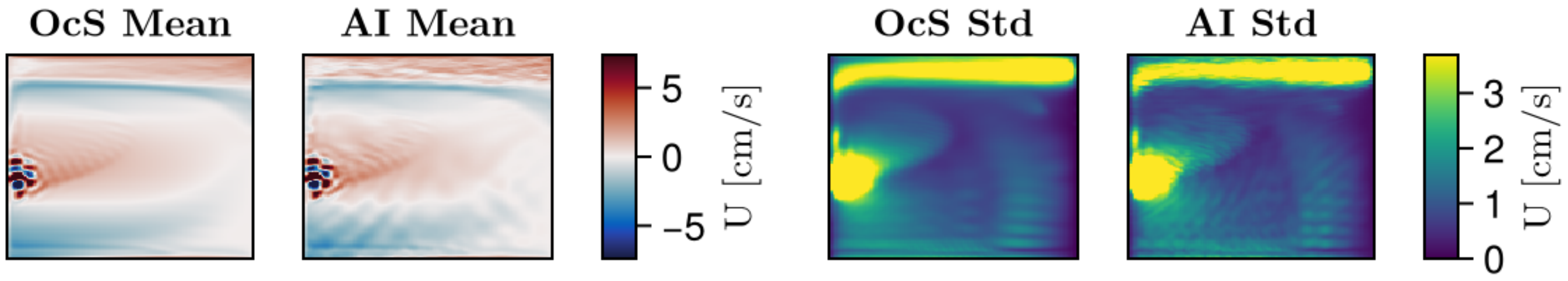
Application: Regression to Ensemble Mean



Application: Regression to Ensemble Mean

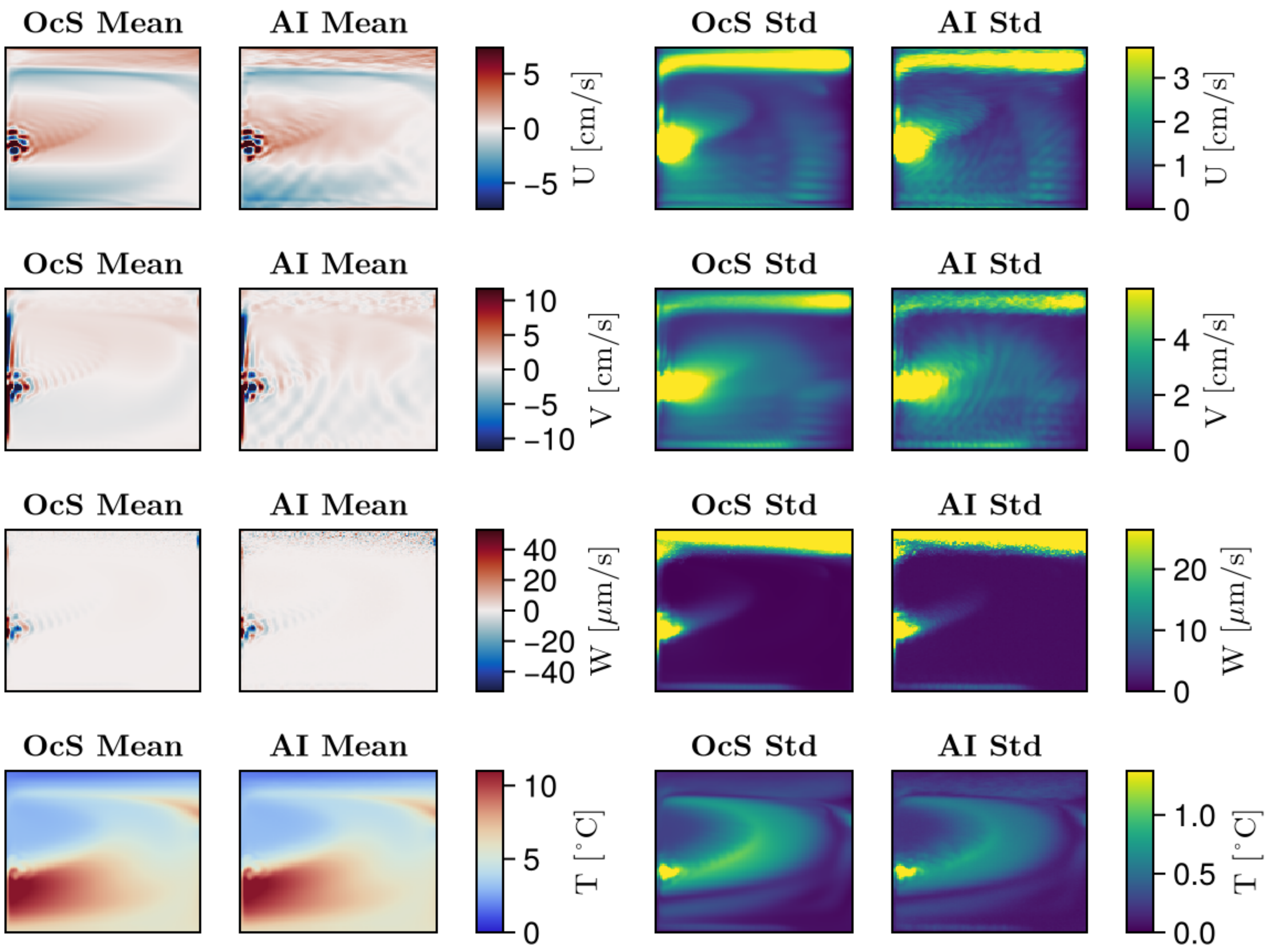


Application: Regression to Ensemble Mean



$$\rho(y|x) = \rho(y)$$

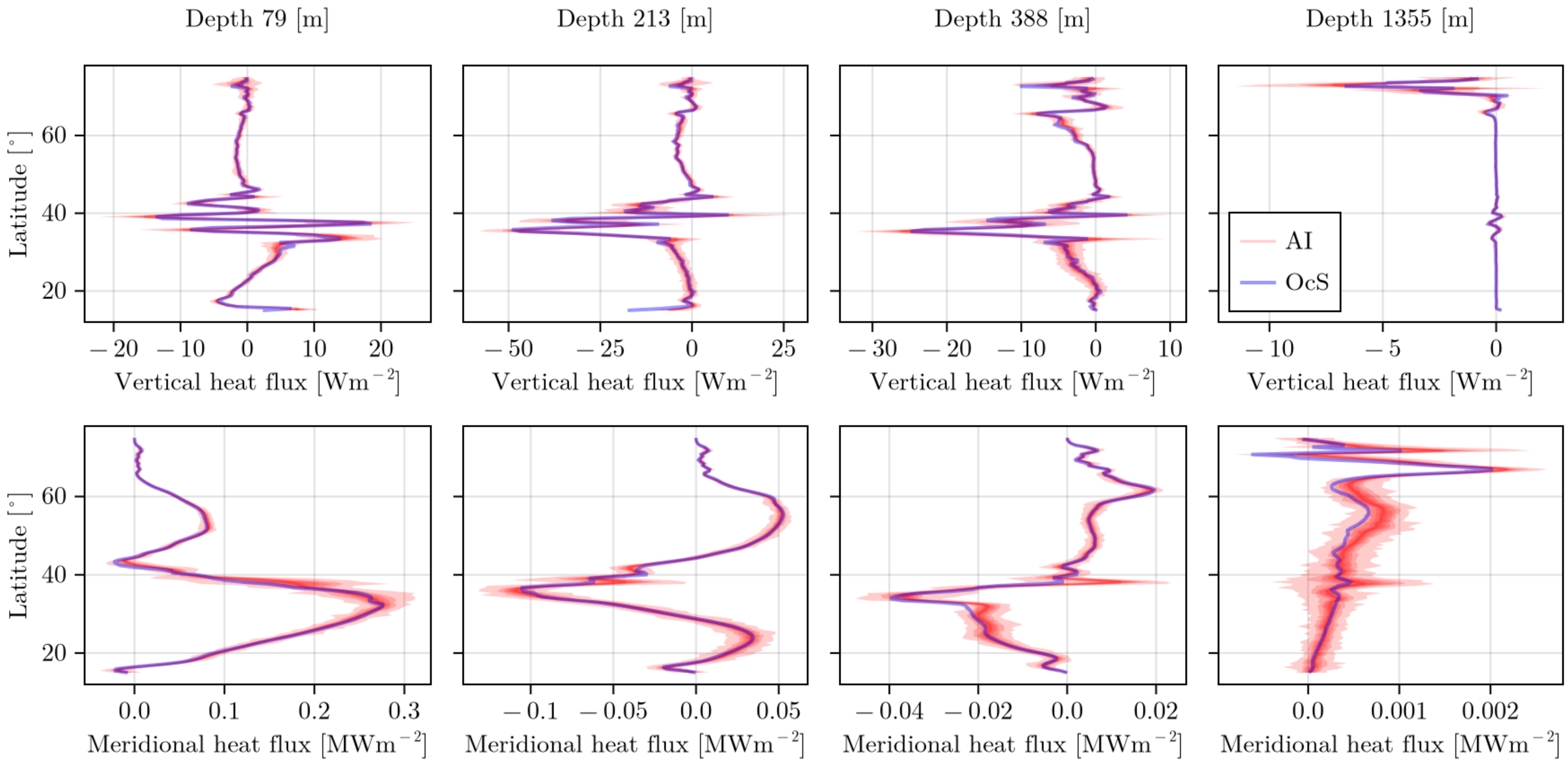
Application: Regression to Ensemble Mean



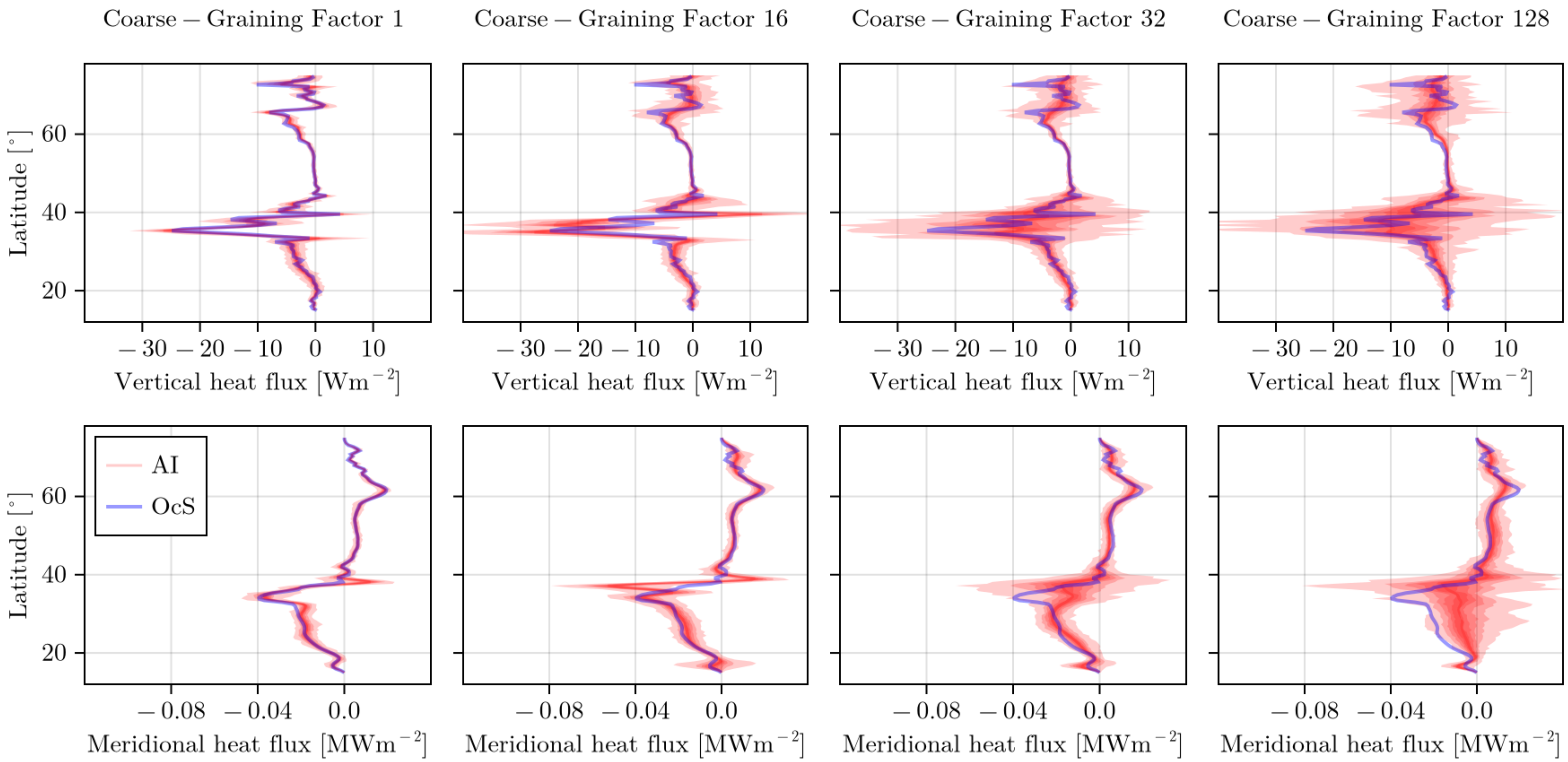
$$\rho(y|x) = \rho(y)$$

Results: Covariances

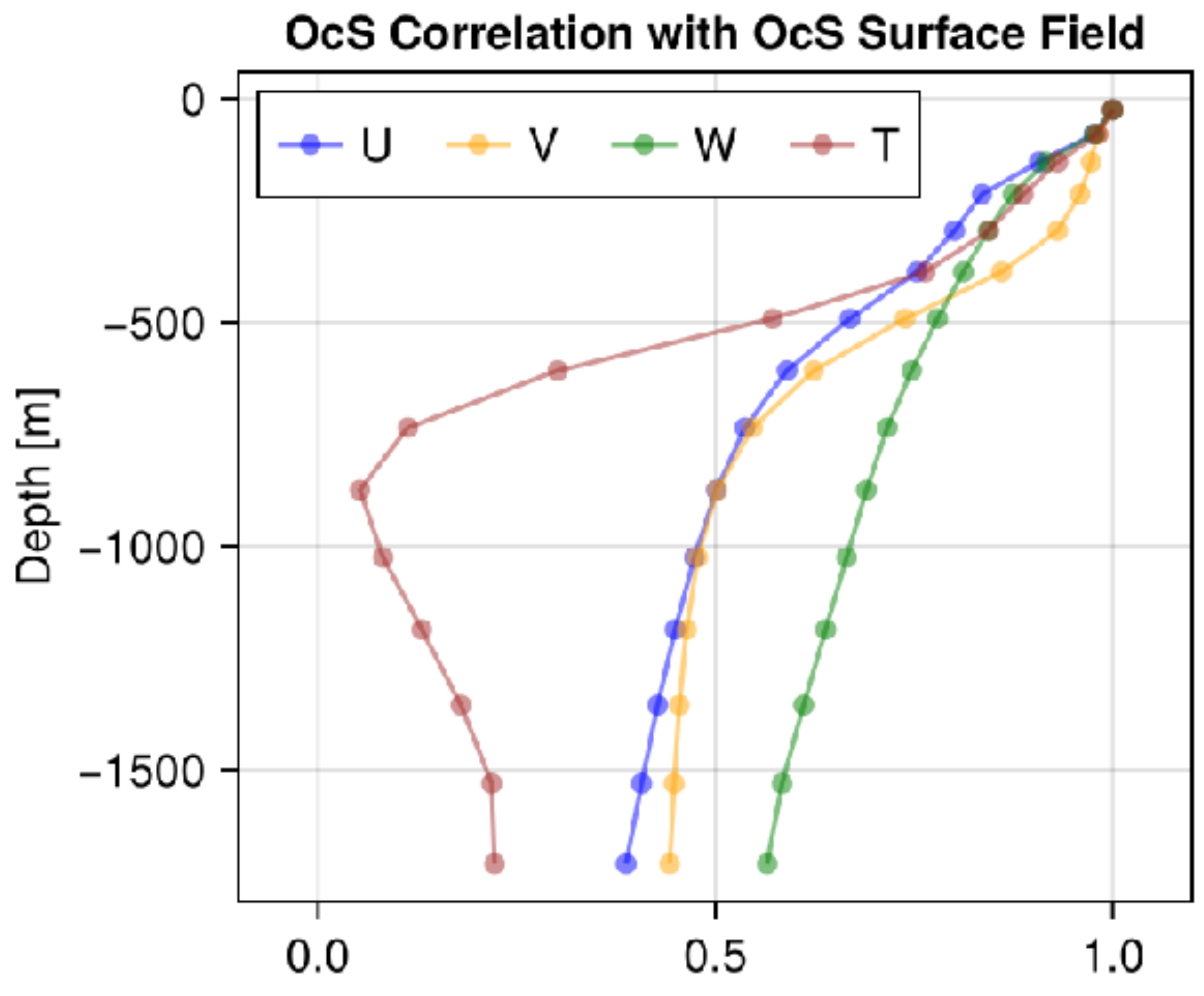
Application: Ocean Inference Covariances

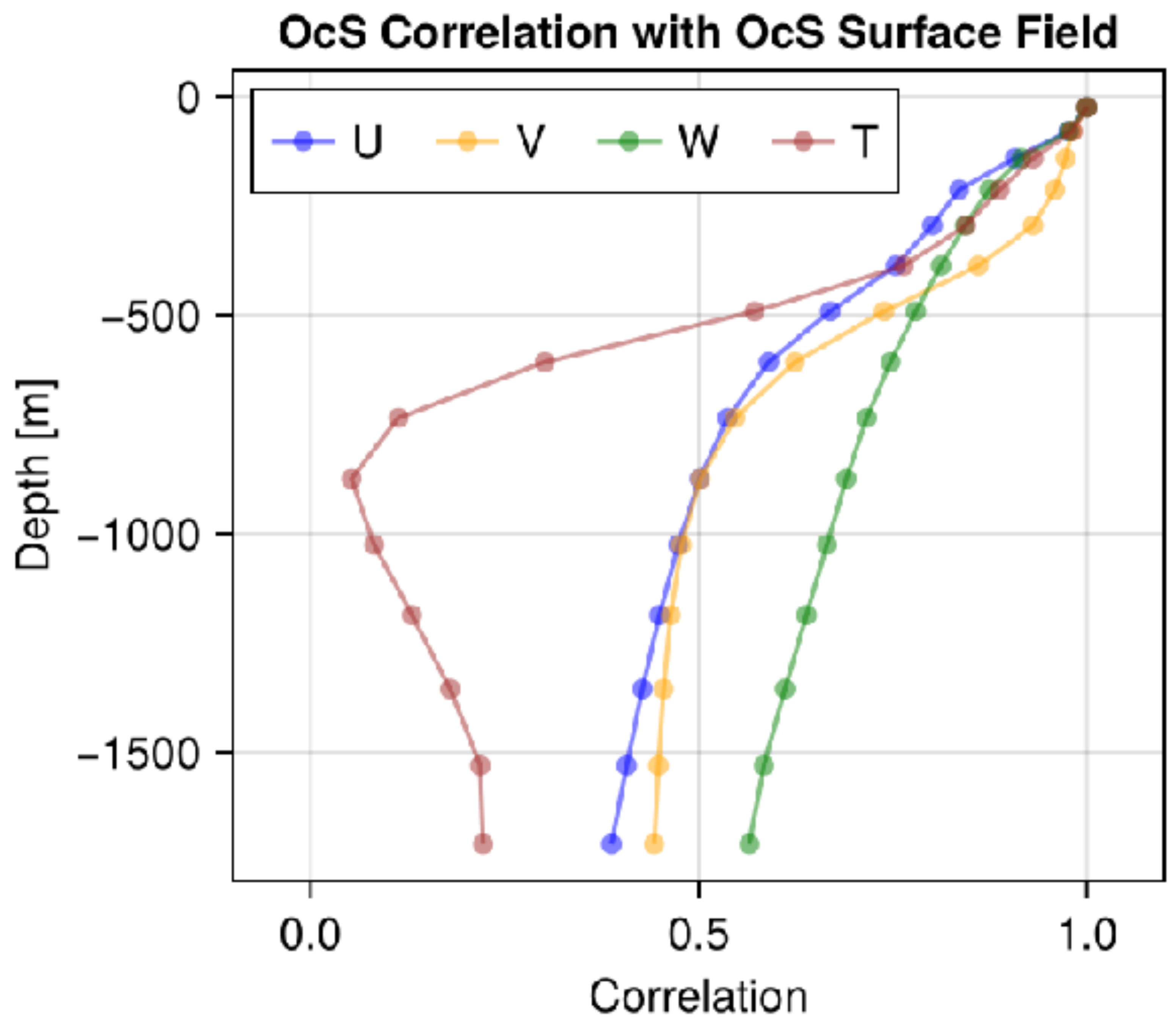


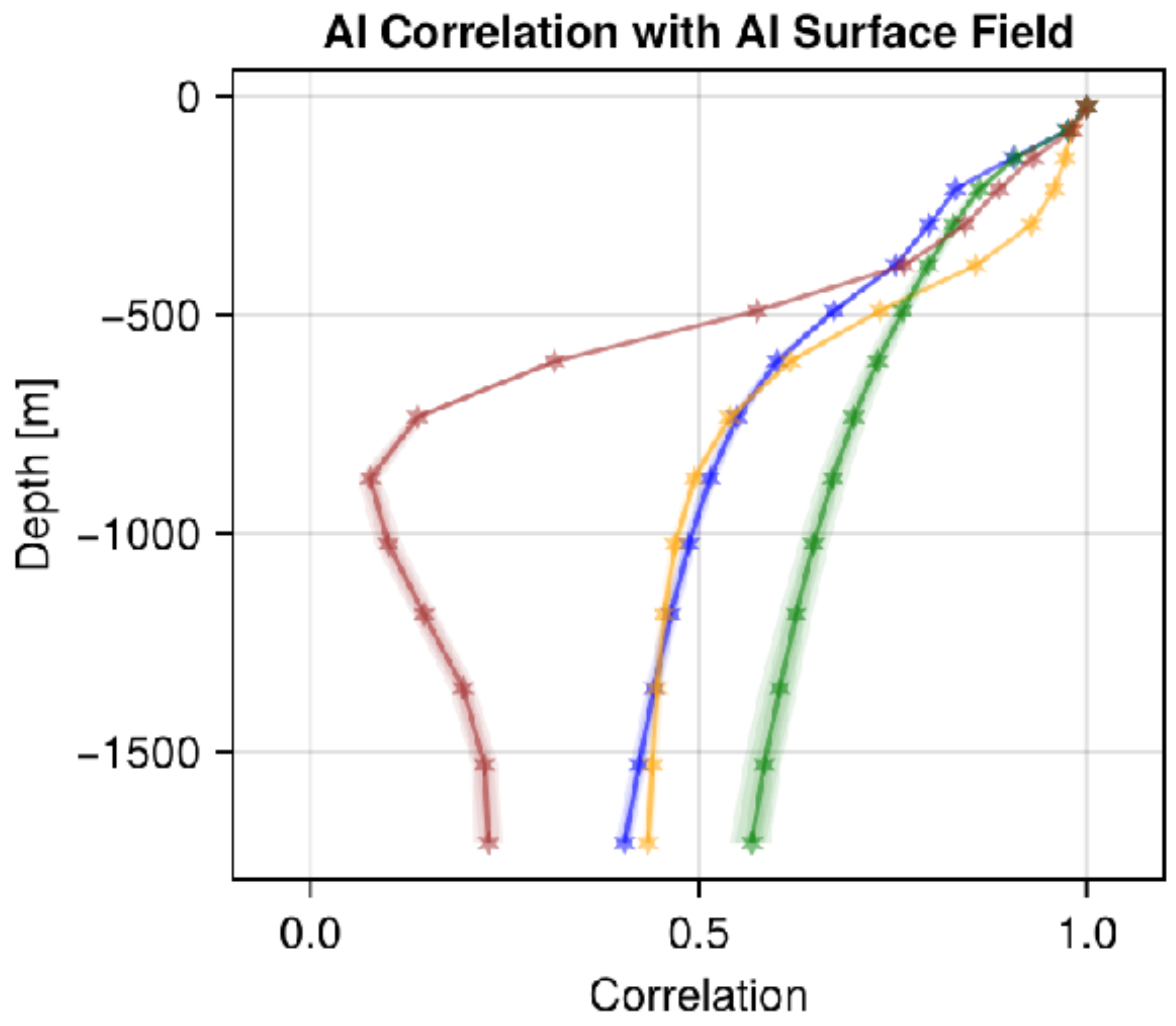
Application: Ocean Inference Covariances

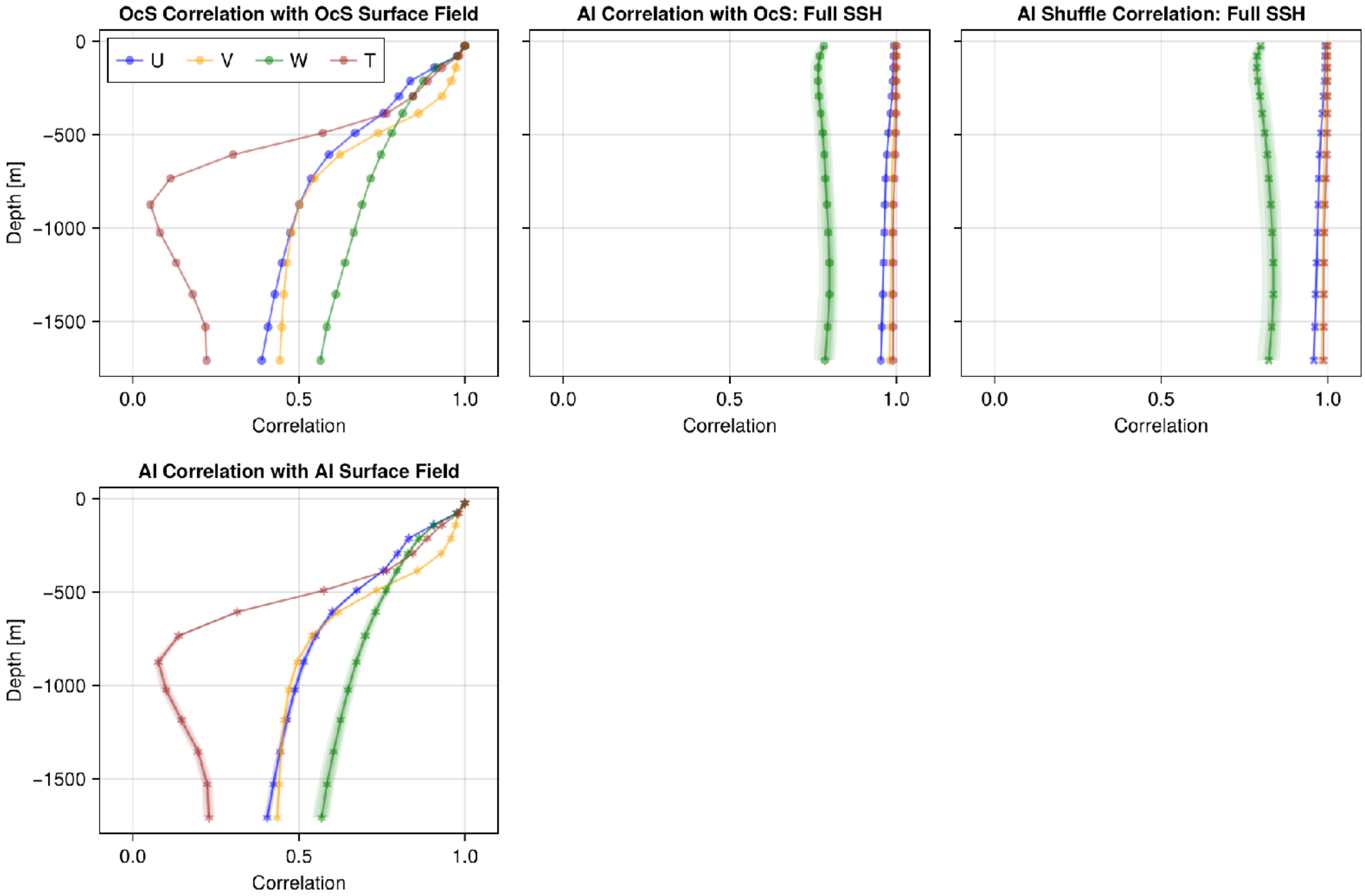


Results: Depth Correlations

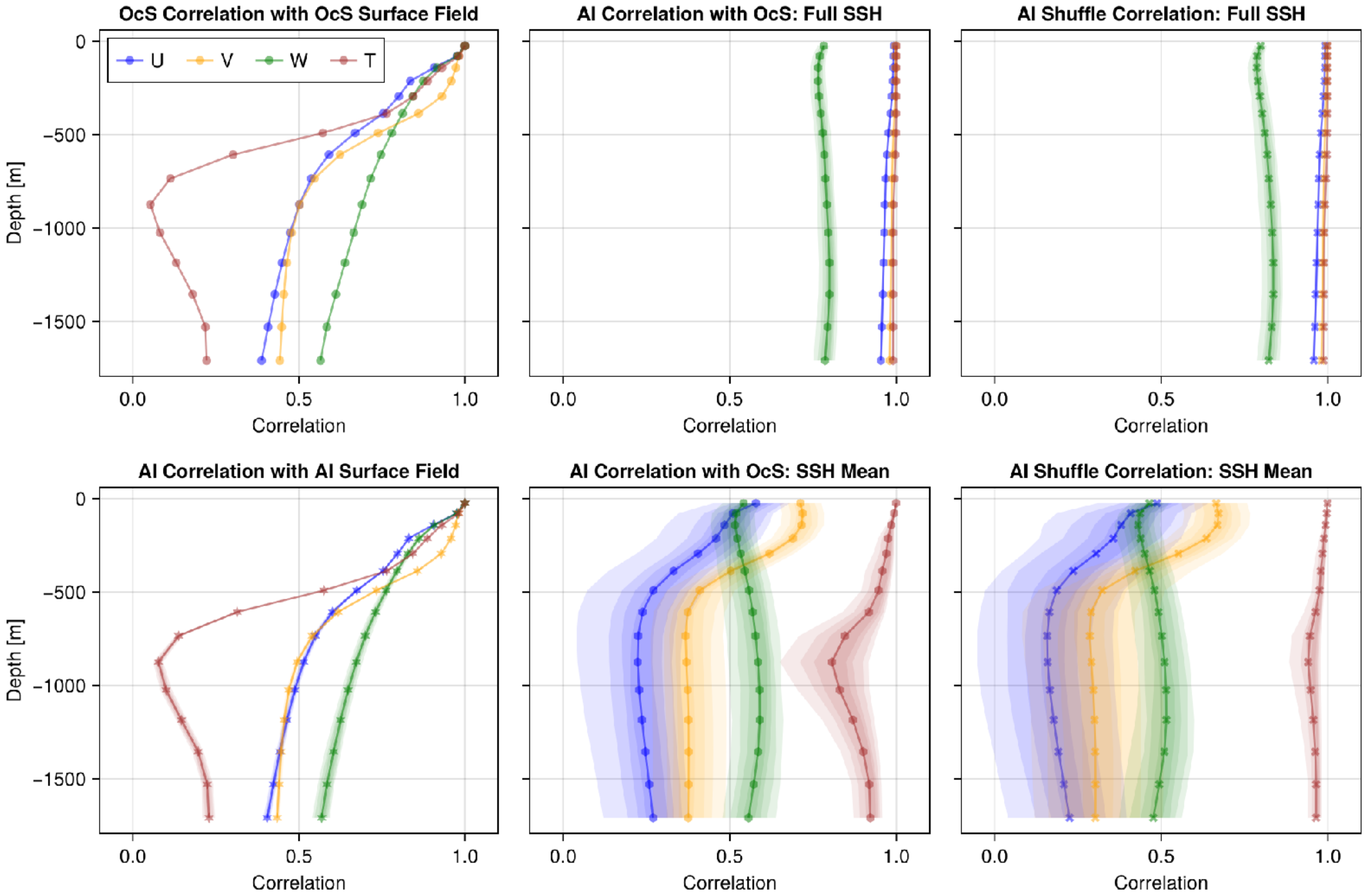

OcS Correlation with OcS Surface Field
U
V
W
T
Depth [m]
0
−500
−1000
−1500
0.0
0.5
1.0


OcS Correlation with OcS Surface Field
U
V
W
T
Depth [m]
0
−500
−1000
−1500
0.0
0.5
1.0
Correlation


AI Correlation with AI Surface Field
Depth [m]
0
−500
−1000
−1500
0.0
0.5
1.0
Correlation

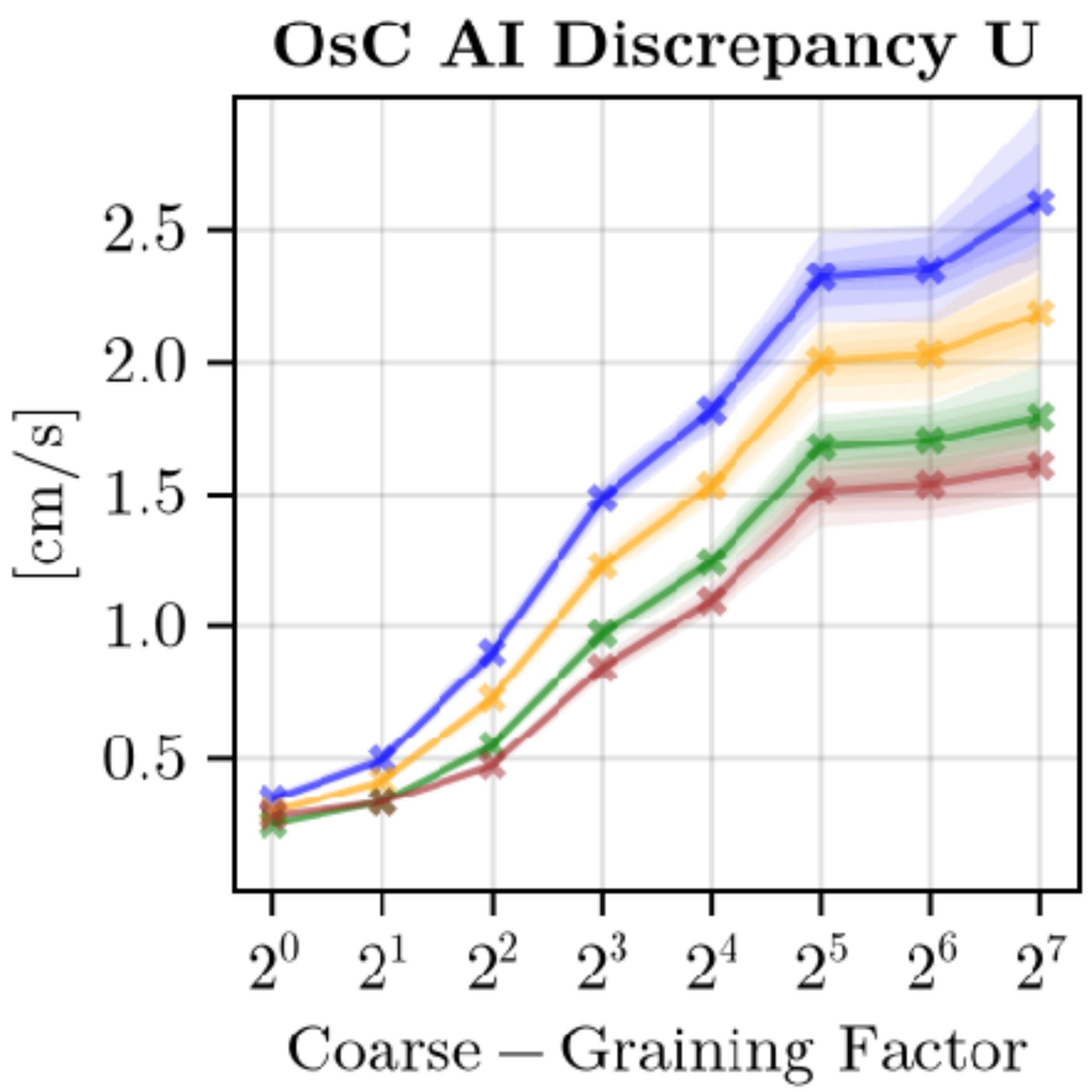


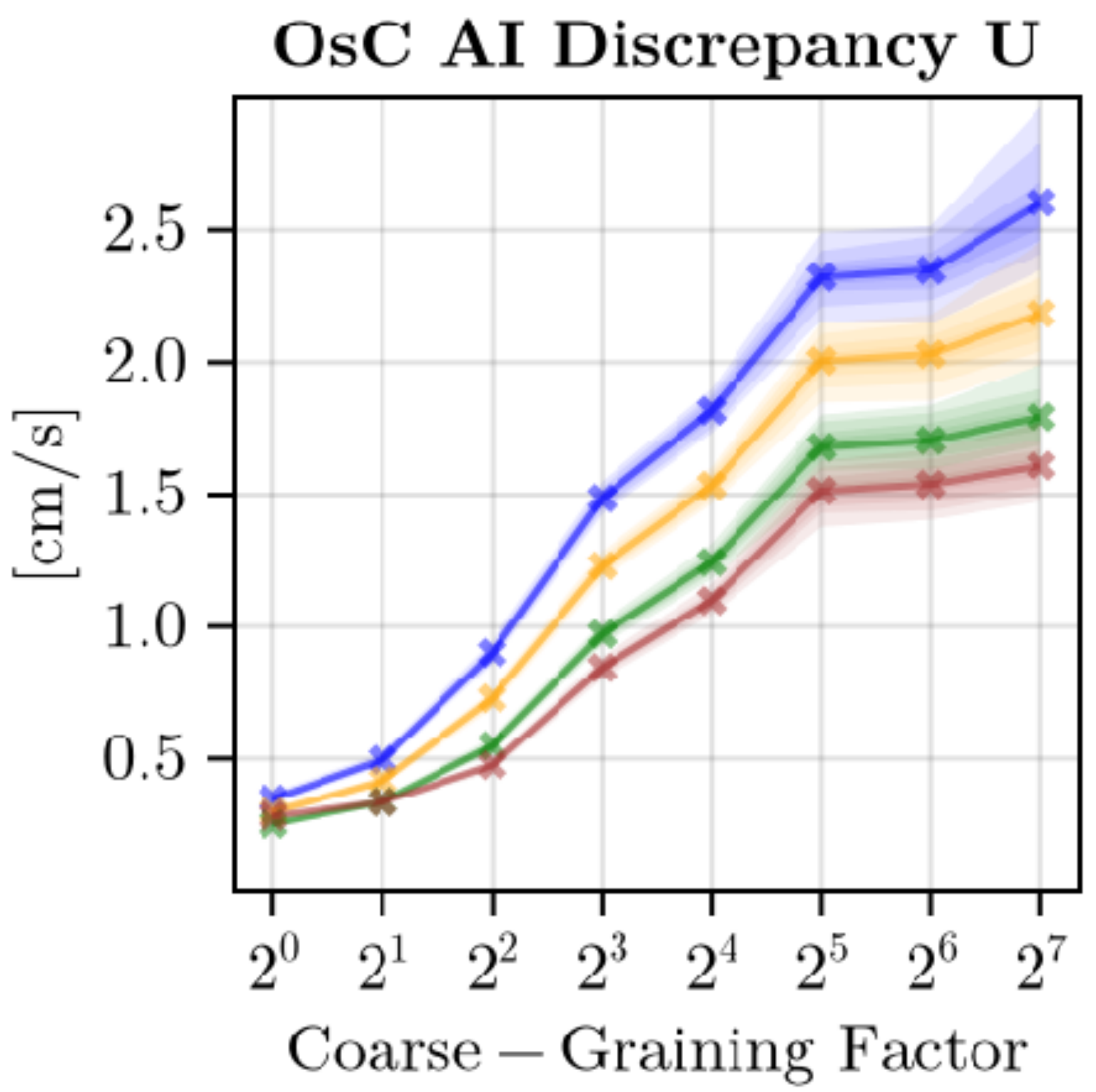
OcS Correlation with OcS Surface Field
AI Correlation with OcS: Full SSH
AI Shuffle Correlation: Full SSH
AI Correlation with AI Surface Field
U
V
W
T
Depth [m]
Correlation
0
−500
−1000
−1500
0.0
0.5
1.0



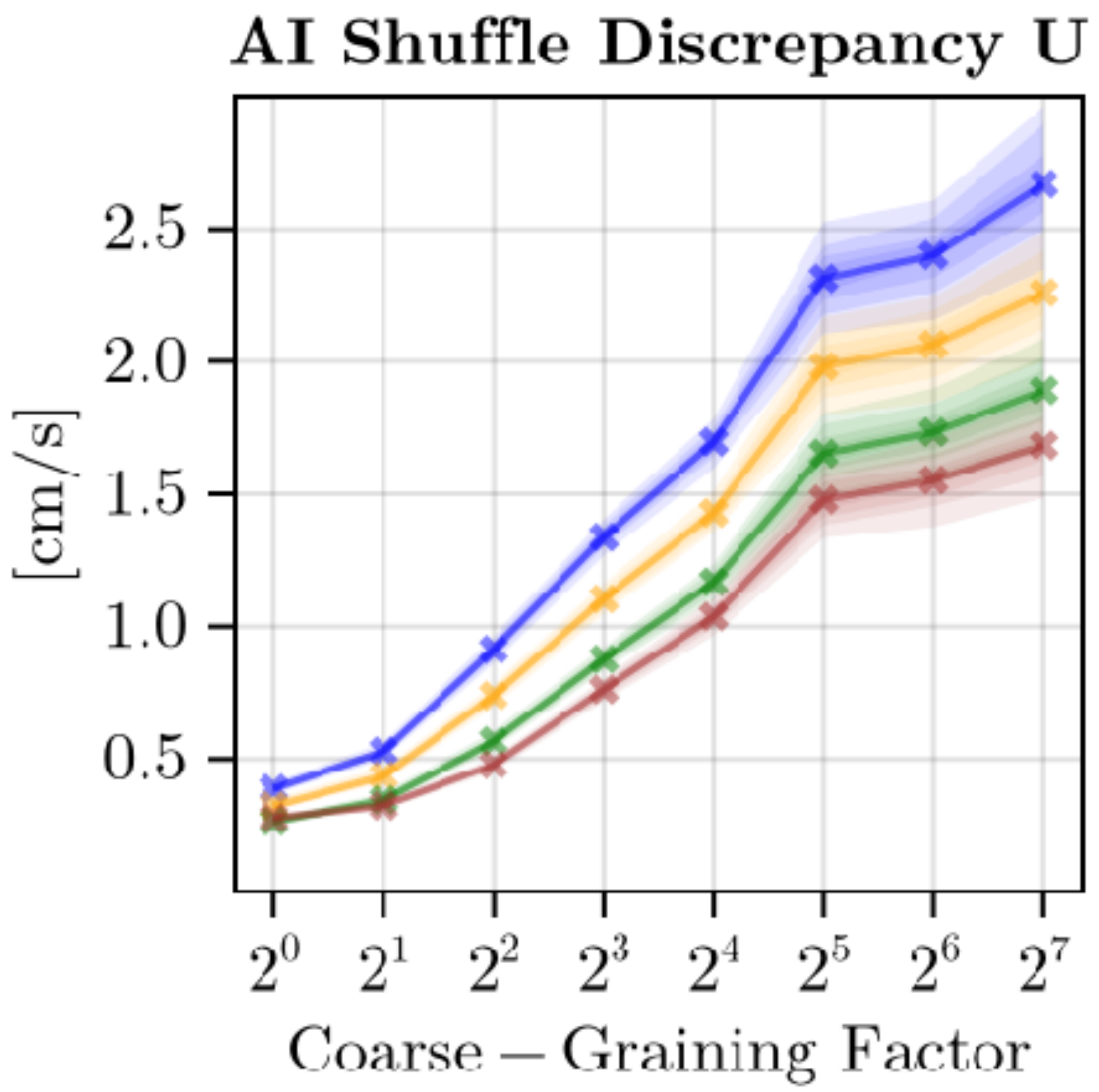
OcS Correlation with OcS Surface Field
AI Correlation with OcS: Full SSH
AI Shuffle Correlation: Full SSH
AI Correlation with AI Surface Field
AI Correlation with OcS: SSH Mean
AI Shuffle Correlation: SSH Mean
U
V
W
T
Depth [m]
0
−500
−1000
−1500
0.0
0.5
1.0
Correlation

Results: Discrepancies

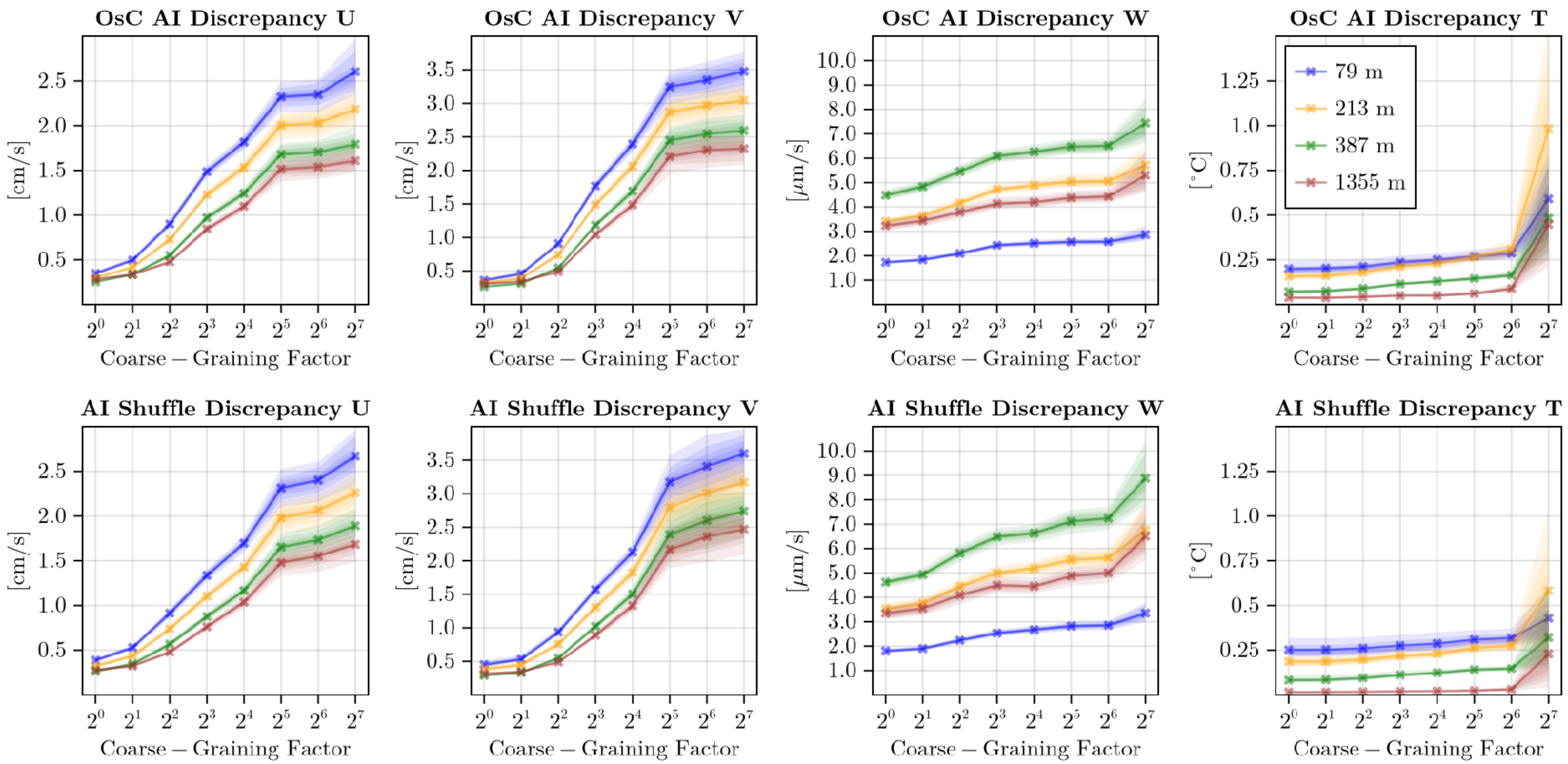

OsC AI Discrepancy U
[cm/s]
2.5
2.0
1.5
1.0
0.5
2^0 2^1 2^2 2^3 2^4 2^5 2^6 2^7
Coarse – Graining Factor



OsC AI Discrepancy U
[cm/s]
2.5
2.0
1.5
1.0
0.5
2^0 2^1 2^2 2^3 2^4 2^5 2^6 2^7
Coarse − Graining Factor

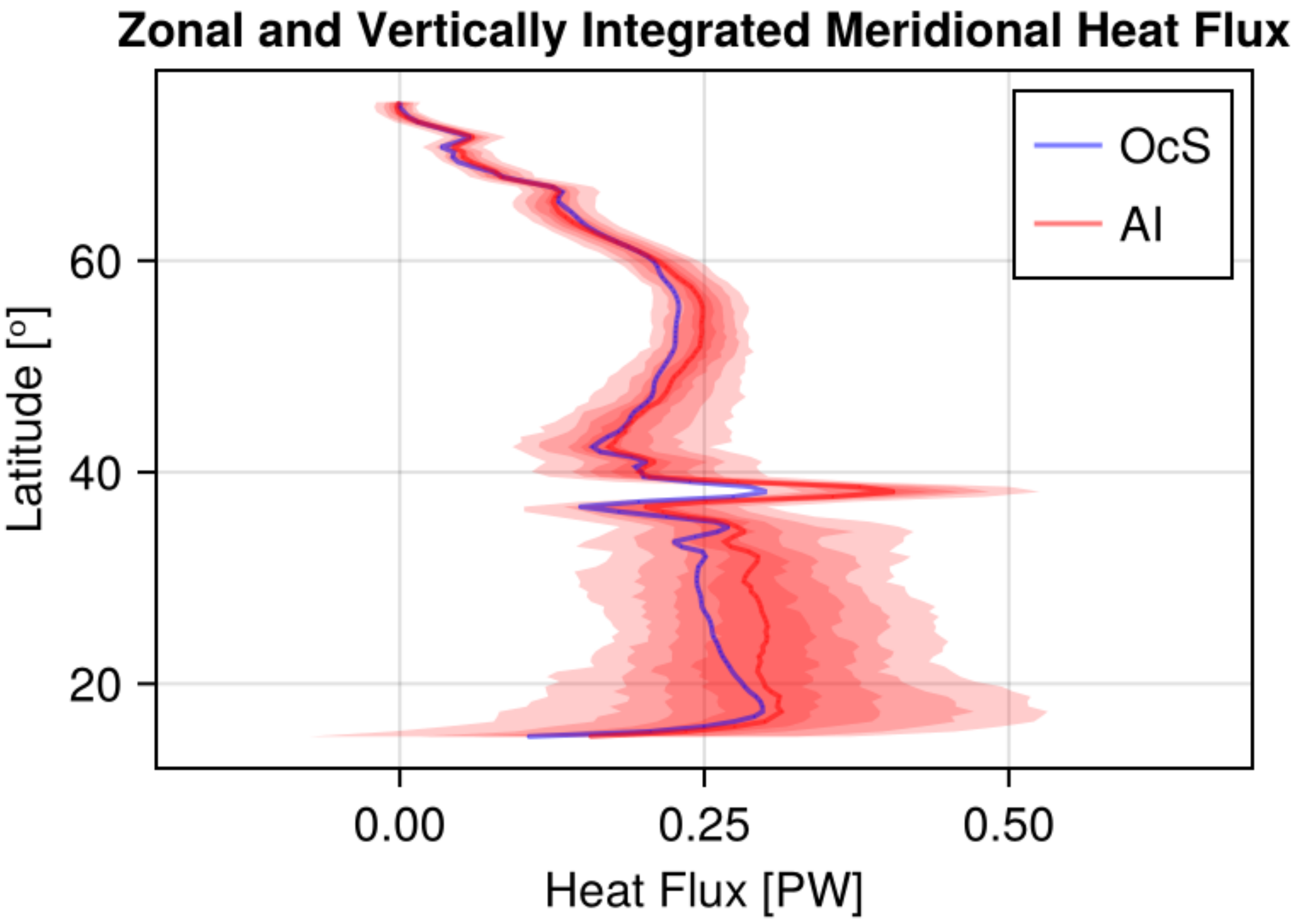


AI Shuffle Discrepancy U
[cm/s]
2.5
2.0
1.5
1.0
0.5
2^0 2^1 2^2 2^3 2^4 2^5 2^6 2^7
Coarse − Graining Factor


OsC AI Discrepancy U
OsC AI Discrepancy V
OsC AI Discrepancy W
OsC AI Discrepancy T
AI Shuffle Discrepancy U
AI Shuffle Discrepancy V
AI Shuffle Discrepancy W
AI Shuffle Discrepancy T
[cm/s]
[μm/s]
[°C]
Coarse – Graining Factor
79 m
213 m
387 m
1355 m

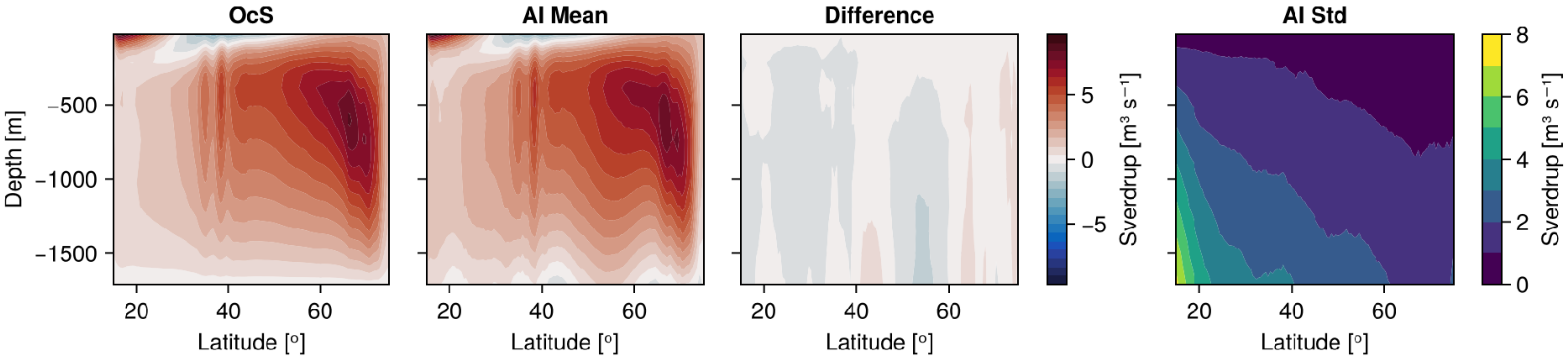
Results: Heat Transport

Application: Ocean Inference Combining Depths



Results: Overturning

Application: Ocean Inference Combining Depths



Take aways

- Questions in Geophysics should be reformulated as probabilistic inference questions

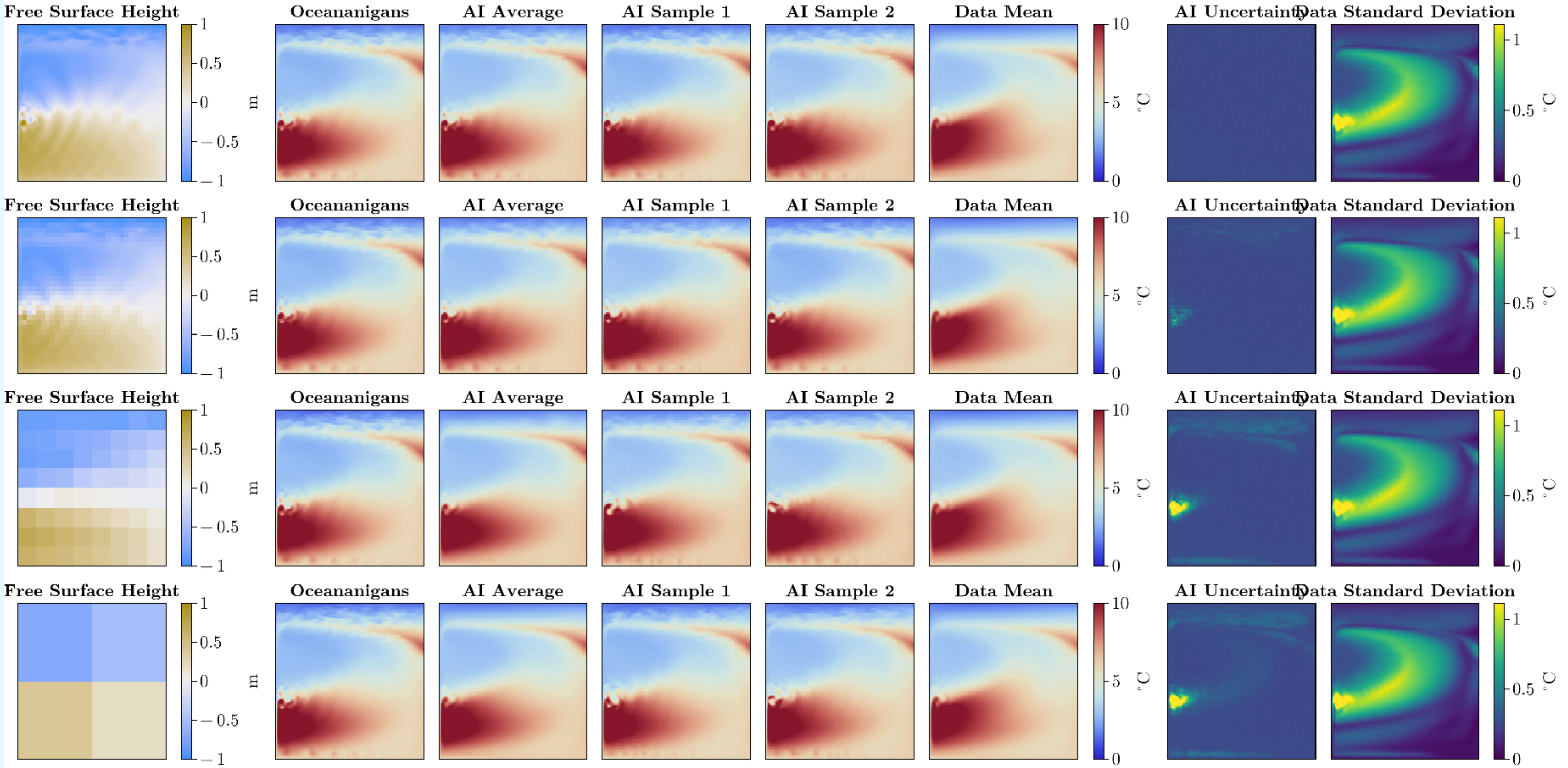
- Gen AI Provides Meaningful Uncertainty Estimates

Questions?

Drawbacks

- Expensive to Evaluate
- Expensive to Train
- Harder to understand the theory of why it works
- Harder for interpretability
- Harder to evaluate
- Requires modifying old / designing new architectures

Application: Ocean Inference Result (400m depth)



Brief Overview of Score-Matching

Brief Overview of Score-Matching

Empirical Distribution

$$\rho(x) = \frac{1}{N}\sum_{n=1}^{N}\delta(x - x^n)$$

Brief Overview of Score-Matching

Empirical Distribution

$$\rho(x) = \frac{1}{N} \sum_{n=1}^{N} \delta(x - x^n)$$

Regularize

$$\rho_\sigma(x) = \frac{1}{N\sqrt{2\pi\sigma^2}} \sum_{n=1}^{N} \exp(-(x - x^n)^2/(2\sigma^2))$$

Brief Overview of Score-Matching

Empirical Distribution

$$\rho(x) = \frac{1}{N} \sum_{n=1}^{N} \delta(x - x^n)$$

Regularize

$$\rho_\sigma(x) = \frac{1}{N\sqrt{2\pi\sigma^2}} \sum_{n=1}^{N} \exp(-(x - x^n)^2/(2\sigma^2))$$

Derivative of Regularized

$$\nabla \ln \rho_\sigma(x) = \frac{\sum_{n=1}^{N} (x^n - x) \exp(-(x - x^n)^2/(2\sigma^2))}{\sum_{m=1}^{N} \exp(-(x - x^m)^2/(2\sigma^2))}$$

Brief Overview of Score-Matching

Empirical Distribution

$$\rho(x) = \frac{1}{N} \sum_{n=1}^{N} \delta(x - x^n)$$

Regularize

$$\rho_\sigma(x) = \frac{1}{N\sqrt{2\pi\sigma^2}} \sum_{n=1}^{N} \exp(-(x - x^n)^2/(2\sigma^2))$$

Derivative of Regularized

$$\nabla \ln \rho_\sigma(x) = \frac{\sum_{n=1}^{N}(x^n - x)\exp(-(x - x^n)^2/(2\sigma^2))}{\sum_{m=1}^{N}\exp(-(x - x^m)^2/(2\sigma^2))}$$

Neural-Network implicit regularization

$$\mathbb{E}|\nabla \ln \rho_\sigma(x) - N_\theta(x)|^2$$

Brief Overview of Score-Matching

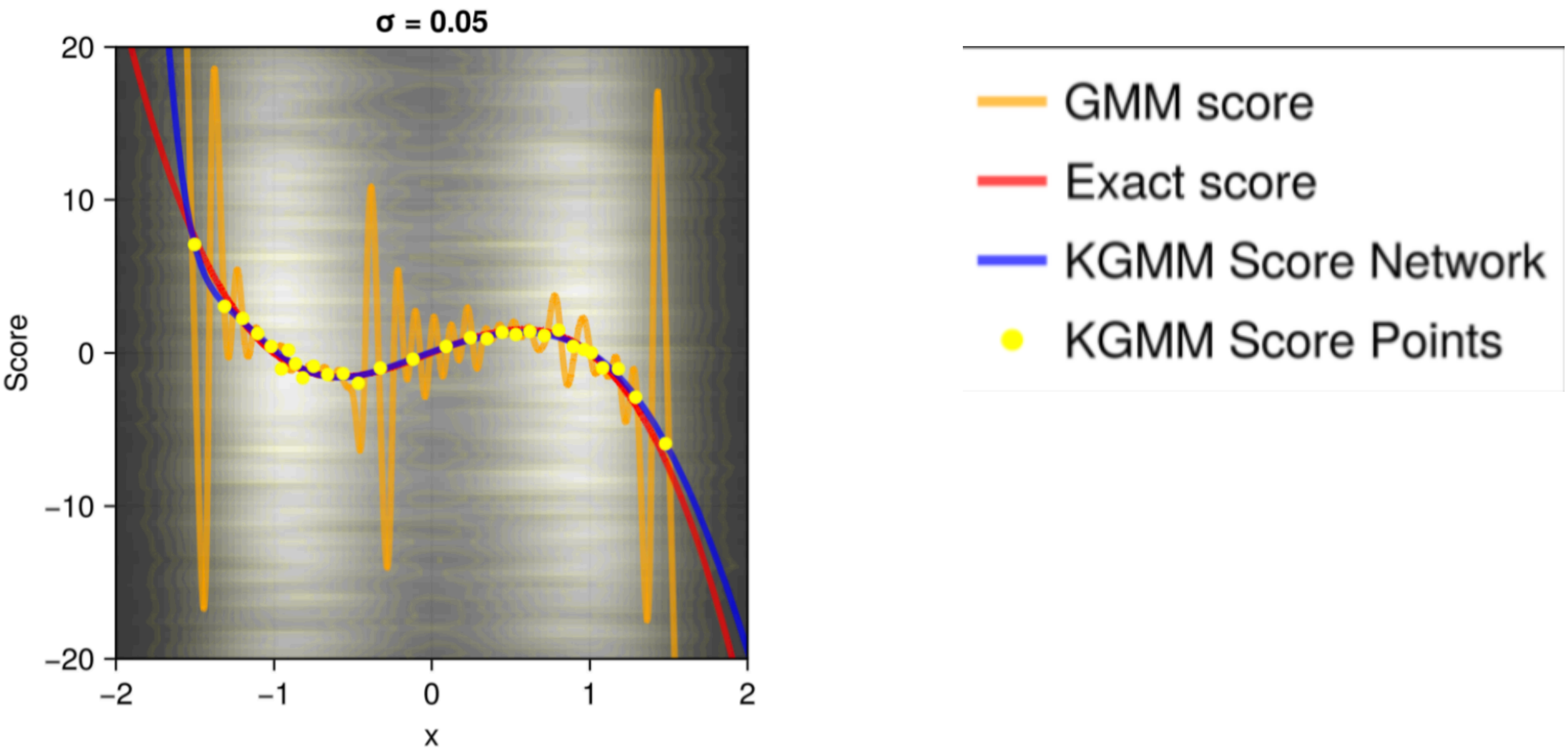
Instead of minimizing

$$\mathbb{E}|\nabla \ln \rho_\sigma(x) - N_\theta(x)|^2$$

We minimize

$$\mathbb{E}_x \mathbb{E}_Z |Z/\sigma + N_\theta(x + \sigma Z)|^2$$

Brief Overview of Score-Matching



Arxiv, Giorgini, Bischoff, Souza:
KGMM: A K-means Clustering Approach to GaussianMixture Modeling for Score Function Estimation