https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

MSR v1.0: A High-Resolution Ocean Parameterization Approach via Multiphysics Super-Resolution

Fuhua Zhu[1], Zhan ao Huang[1], Pengfei Pan[1], Wenhao Huo[1], Fengtao Zuo[1], Xian Zhang[3], Xiaojie Li[1], Yongqiang Yu[2], and Xi Wu[1]

[1]School of Computer Science, Chengdu University of Information Technology, Chengdu, 610225, China
[2]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China
[3]The university of western Australia, Perth, 6009, Australia

Correspondence: Xiaojie Li (lixj@cuit.edu.cn) and Xi Wu (wuxi@cuit.edu.cn)

Abstract. High-resolution reconstruction of ocean dynamics is challenging because spectral bias and the neglect of cross-variable couplings in existing super-resolution (SR) methods often lead to over-smoothed, physically inconsistent outputs, limiting their utility for eddy parameterizations. To overcome these limitations, we present a Multiphysics Super-Resolution version 1.0 (MSR v1.0) framework that jointly reconstructs eight closure-relevant diagnostics—vorticity, deformation measures, stress tensor components, and subgrid momentum forcing terms—directly from low-resolution (LR) velocity fields, consistency is maintained under a filtering scale that aligns with the definition of subgrid tendencies. Our approach integrates three key components: (1) a Dynamic Enhancement Feature (DEF) module to prioritize dynamically active regions; (2) a High-Frequency Enhancement (HFE) module that fuses spatial and spectral operators via learned gating to restore suppressed fine-scale structures such as fronts and eddy rims; and (3) a Physical Consistency Loss that aligns derivative-level structures and algebraic couplings across diagnostics. Experiments on an eddy-resolving simulation dataset across multiple ocean basins and downscaling factors show that MSR consistently outperforms strong SR baselines, yielding sharper reconstructions with improved high-wavenumber spectra and cross-variable consistency. The MSR-reconstructed diagnostics are closure-ready for low-resolution ocean models and can inform or constrain eddy parameterizations, providing a practical, spectrally selective, scale-aware surrogate for high-fidelity multiphysics diagnostics.

1 Introduction

The ocean plays a crucial role in regulating the Earth's climate, carbon cycling, and weather patterns. It acts as a vast reservoir for heat and carbon, mediates air–sea exchanges, modulates storm tracks, and influences biogeochemical pathways (Mak et al., 2022). These ocean-mediated processes significantly affect global weather patterns, sea level rise, and the distribution of marine resources, making the accurate representation of ocean dynamics vital for climate science. Accurately resolving these features is crucial for improving climate reconstructions (Fox-Kemper et al., 2019), yet their impacts must, in practice, be conveyed by *ocean parameterizations* whenever eddies, fronts, and shear layers are under-resolved. Contemporary eddy parameterizations have advanced beyond purely diffusive mixing to include energy backscatter and flow-dependent terms, and data-driven or discovery-based approaches have begun to infer closures directly from high-resolution simulations (Bolton and Zanna, 2019).

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

Nevertheless, the community recognizes persistent shortcomings: parameterizations remain a source of error and uncertainty in simulations and projections, many are not robustly *scale-aware* across resolutions, and some struggle to represent up-gradient momentum transfer and realistic spectra without over-diffusion or ad hoc tuning (Zhang et al., 2023). These issues motivate work that treats parameterization not merely as smoothing, but as a structured, interpretable representation of inter-scale transfers that adapts across resolutions and stays consistent with dynamical constraints.

Due to computational constraints, many ocean and climate models are run at low resolution (Chassignet and Xu, 2021), which leaves fine scales under-resolved and makes parameterizations central to model fidelity. Implementations reveal important limitations: sensitivity to the choice of spatial filtering or spectral selectivity used to define targets and tendencies; dependence on grid spacing that necessitates explicit scale awareness; uneven skill across depths and near boundaries that invites case-specific tuning; and an offline–online mismatch whereby schemes that score well in a posteriori tests may face stability or realism challenges in long integrations (Perezhogin et al., 2023; Kitsios et al., 2023). Even gradient-informed designs can experience local energization and require safeguards, while stochastic formulations are sensitive to the spatial/temporal structure of the injected variability and to computational cost when coupled to production codes (Khani and Dawson, 2023). These realities underscore the need for parameterizations that combine physical structure with spectral selectivity and resolution awareness—capabilities our work targets explicitly.

Complementary to classical eddy parameterizations, learning-based super-resolution has emerged as a promising surrogate pathway to bridge the gap between LR and HR simulations Anwar et al. (2020); Fukami et al. (2023). These methods aim to reconstruct high-resolution fields from low-resolution inputs by learning complex mappings between the two. By leveraging the power of deep learning, SR methods can recover fine-scale ocean features and simulate small-scale processes at a lower computational cost than conventional HR simulations. While some success has been achieved in improving model accuracy, existing SR methods have limitations that hinder their ability to fully capture the ocean's dynamic features. Many of these methods focus on improving the resolution of single variables or two-dimensional (2D) slices, often neglecting the interdependencies between different physical quantities, such as vorticity, deformation, and stress. Moreover, these approaches often suffer from spectral bias, leading to the over-smoothing of important small-scale features that are critical for ocean dynamics. Another limitation is that most SR methods lack physical consistency, particularly in preserving derivative-level structures like vorticity and strain, as well as the algebraic relationships between multiphysics diagnostics. These limitations make it difficult to generate high-fidelity reconstructions of complex ocean dynamics, particularly in regions with large and spatially varying resolution gaps.

In this paper, we propose the MSR v1.0 framework to address these challenges. Our approach aims to improve the accuracy of high-resolution ocean diagnostics by reconstructing eight critical diagnostic fields (vorticity (ζ), shearing deformation (D), stretching deformation ($\tilde{D}$), stress tensor components (T_{11}, T_{22}, T_{12}), and subgrid forcing terms (S_u, S_v)) directly from low-resolution velocity fields. These fields are essential for capturing the fine-scale dynamics of ocean processes, such as eddy formation, turbulence, and shear, which are crucial for accurate climate modeling. By incorporating ocean physics directly into the model, we mitigate spectral bias and ensure that the model maintains robust performance across regions with varying downscaling factors, such as areas of strong mesoscale eddies and dynamic coastal zones. Our model improves upon existing

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

methods by restoring key small-scale features that are suppressed in low-resolution inputs, ensuring that the reconstructed high-resolution fields retain their physical integrity.

The key contributions of this work are summarized as follows:

- To improve the model's ability to generalize across ocean basins and dynamically active regions, we propose a Dynamic Enhancement Feature (DEF) module that provides more physical features and a strain-attention map derived from velocity gradients, prioritizing key areas like shear layers and frontal zones.
- To enable the effective capture of fine-scale ocean features, we propose a High-Frequency Enhancement (HFE) module that includes Sobel (edges), Laplacian (curvature), Haar wavelets (scale-localized bands), and frequency domain enhancement (FDE). These features are fused via learned gating to restore critical features like fronts and eddy rims that are suppressed in low-resolution inputs.
- To ensure the physical consistency of the model, we incorporate a physical consistency loss that aligns derivative-level structures, such as strain and vorticity, between the model's output and the true high-resolution values. This loss function leverages the relationships between the diagnostic variables and enforces the preservation of key physical dynamics during reconstruction. By aligning these structures, we ensure that fine-scale variations are accurately captured, thus maintaining the integrity of critical oceanic features across different spatial resolutions.

2 Related Work

In this section, we review two primary lines of research that are most relevant to our work: (i) data-driven super-resolution methods for geophysical fields and (ii) data-driven ocean parameterization approaches.

2.1 Super-Resolution of Geophysical Fields

Super-resolution methods have gained significant attention in the geophysical community, particularly in enhancing the resolution of oceanographic fields like sea surface temperature (SST) and sea surface height (SSH) (Meng et al., 2021). These methods, adapted from computer vision techniques, typically involve deep neural networks, such as residual and attention-based U-Nets, as well as advanced architectures like EDSR (Enhanced Deep Super-Resolution Network) and RCAN (Residual Channel Attention Network)(Buongiorno Nardelli et al., 2022; Buza and Akagic, 2025). Such models have demonstrated impressive performance in recovering small-scale features from low-resolution data, often outperforming interpolation or other baseline methods (Li et al., 2021; Sun et al., 2022; Lu et al., 2022).

However, most existing SR methods in oceanography suffer from several limitations. First, many of these approaches focus on enhancing single variables or 2D slices of data, neglecting the interdependencies between multiple ocean variables, such as vorticity, deformation, and subgrid forcings. Second, most methods rely on pixel-based losses, which fail to preserve the derivative-level structures (e.g., strain, vorticity) that govern ocean dynamics. This limitation restricts the ability of these models

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

to capture fine-scale features, such as ocean fronts and eddy rims, and hinders their portability across different regions or scale factors.

2.2 Ocean Parameterization Using Data-Driven Approaches

Ocean parameterization plays a vital role in simulating unresolved small-scale ocean processes in large-scale models (Perezhogin et al., 2025). Traditional parameterization methods rely on empirical relationships or predefined rules, which are computationally efficient but often fail to capture the full complexity of subgrid-scale dynamics. With the advent of deep learning, data-driven parameterization methods have emerged as a promising alternative, enabling more accurate representations of unresolved ocean processes, such as subgrid fluxes and stresses, without the need for high-resolution simulations.

Recent advances in data-driven ocean parameterization focus on using deep neural networks to estimate subgrid-scale stresses and fluxes directly from resolved velocity fields and gradients. Techniques like physics-informed neural networks (PINNs) have been introduced to ensure that the model estimations adhere to fundamental oceanic principles (Lawal et al., 2022; de Wolff et al., 2021; Yoon et al., 2024). Despite these advancements, a key challenge remains: most methods still focus on single, isolated variables, and they often fail to jointly reconstruct multiple coupled diagnostics, which is crucial for accurately modeling the complex interactions within the ocean.

In this paper, our MSR framework represents a significant step forward by jointly reconstructing multiple coupled diagnostics—such as vorticity, deformation, and subgrid forcings—directly from low-resolution inputs. By integrating physical knowledge with data-driven learning, our approach enhances the representation of ocean dynamics, improving both the accuracy and consistency of parameterizations in resource-constrained environments.

3 Data Preparation and Preprocessing

The parameterization dataset is constructed from **Institute of Atmospheric Physics (LASG/IAP) Climate System Ocean Model version 3 (LICOM3)** with eddy-resolving simulations at $1/10^{\circ}$ horizontal resolution and 55 vertical layers, covering the period 1901-–2010. The data files are provided in netCDF-4 classic format and is publicly archived as a curated dataset described in (Ding et al., 2022) (https://doi.org/10.1038/s41597-022-01766-9). These variables (sea water X-velocity (u) and Y-velocity (v) fields) from LICOM3 encapsulate mesoscale processes that are central to climate modeling (Zanna and Bolton, 2021), serving as the basis for deriving eddy kinetic energy and subgrid-scale stresses.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

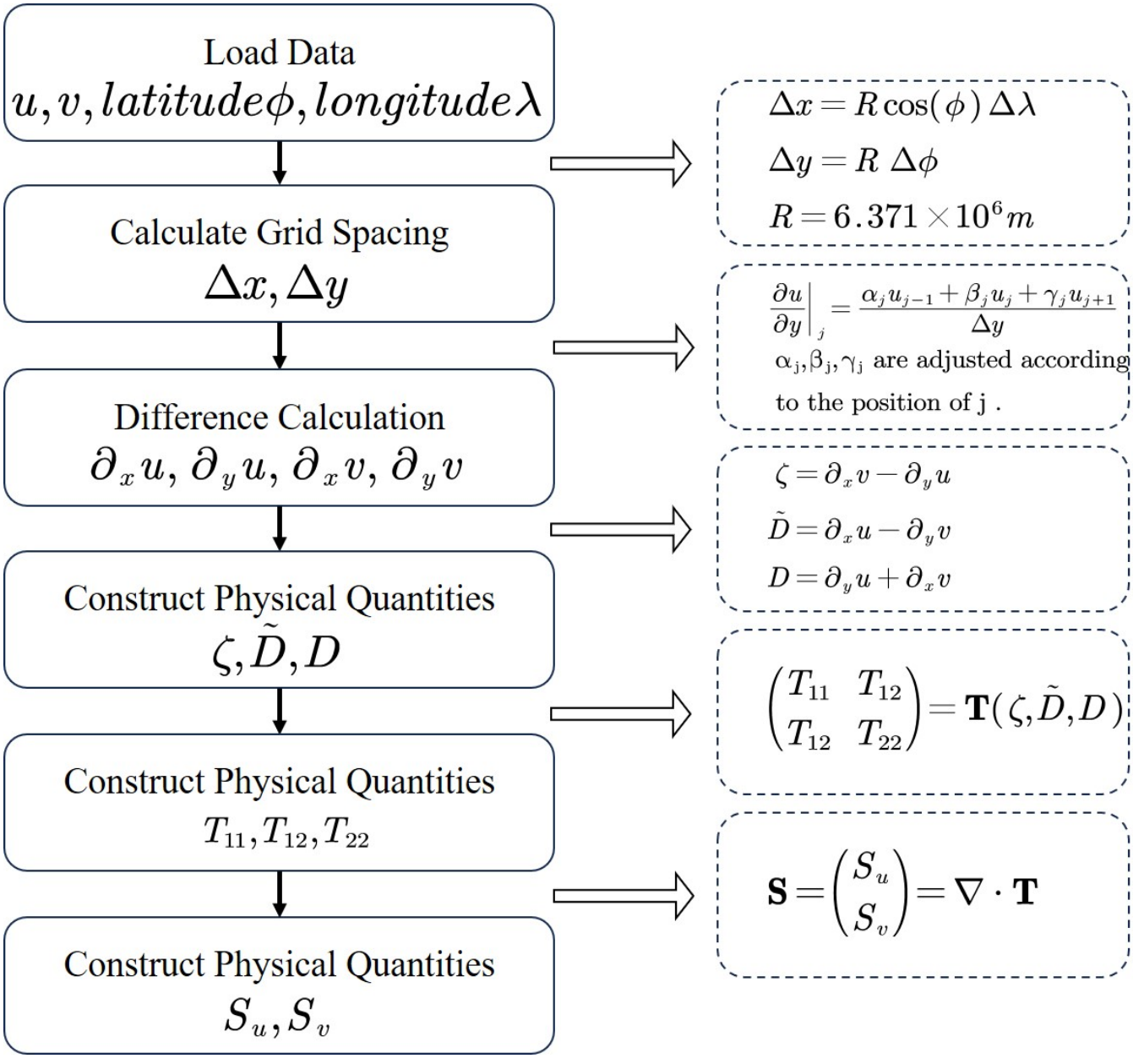

Figure 1. Calculation workflow for generating high-resolution ocean diagnostics.

To generate physically consistent high-resolution data, we implement a parameterization script similar to (Perezhogin et al., 2024) that processes multi-layer velocity fields (u, v) from numerical simulations to compute key diagnostic variables across vertical levels. This approach ensures that the generated HR data accurately reflect the physical dynamics of the system, while capturing small-scale features that are crucial for mesoscale eddy dynamics. As shown in Figure 1, these diagnostic

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

variables (e.g., $\zeta, \tilde{D}, D$), generated through physical numerical calculations, serve as the labels for training the data-driven neural network.

Physics-consistent parameterization for label generation: Given latitude ϕ and longitude λ and Earth's radius $R = 6.371 \times 10^6$ m (Pacanowski and Gnanadesikan, 1998), we compute physical grid spacings Δx and Δy by Eq. (1):

$$\Delta x = R\cos(\phi)\,\Delta\lambda, \qquad \Delta y = R\,\Delta\phi, \tag{1}$$

where $\Delta\lambda$ and $\Delta\phi$ are the increments in longitude and latitude (in radians), respectively. This step ensures that the grid spacing reflects Earth's curvature, a fundamental aspect of mesoscale dynamics. The increments (Δx, Δy) directly influence the subgrid-scale interactions and energy transfer mechanisms in ocean models.

Based on the increments $(\Delta x, \Delta y)$, spatial derivatives are obtained via the hybrid differencing scheme, which is used in mesoscale ocean simulations for reducing numerical noise and improving stability (Batteen and Han, 1981). For example, the meridional derivative of the velocity component u is calculated as

$$\left.\frac{\partial u}{\partial y}\right|_j = \begin{cases} \dfrac{-3u_j + 4u_{j+1} - u_{j+2}}{2\Delta y}, & j \text{ at lower boundary}, \\ \dfrac{u_{j+1} - u_{j-1}}{2\Delta y}, & j \text{ interior}, \\ \dfrac{3u_j - 4u_{j-1} + u_{j-2}}{2\Delta y}, & j \text{ at upper boundary}. \end{cases} \tag{2}$$

This same approach is applied to compute all other velocity derivatives ($\partial u/\partial x$, $\partial v/\partial x$, $\partial v/\partial y$), ensuring both boundary accuracy and interior robustness. They are critical for accurately resolving mesoscale eddies, which are pivotal for simulating oceanic energy exchanges.

Building on the robust gradient estimates, to accurately capture both large and small-scale ocean processes, we compute the key kinematic diagnostics of ocean dynamics (Bagaeva et al., 2024)(**vorticity ζ, stretching deformation $\tilde{D}$, shearing deformation** D):

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}, \qquad \tilde{D} = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}, \qquad D = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}. \tag{3}$$

These kinematic invariants govern the interactions between eddies and large-scale currents, and are fundamental to improving the energy transfer modeling in ocean models.

Given Δx, Δy, ζ, $\tilde{D}$, D, we compute $\mathbf{T}$, $\mathbf{S}$ by Eq.(5)-Eq.(6). First, we adopt formulation from Bolton and Zanna (2019), the ZB20 coefficient is defined as

$$\kappa_{BC} = -\gamma \Delta x \Delta y \leq 0, \tag{4}$$

where γ is a dimensionless scale factor controlling backscatter strength, typically within the range of $[0.5, 2]$. This coefficient plays a critical role in controlling the strength of eddy-driven energy transfer and backscatter.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

The subgrid stress tensor incorporates nonlinear interactions among the kinematic invariants of ocean dynamics, which are essential for accurately capturing mesoscale eddy interactions. The stress tensor is formulated as:

$$\mathbf{T}(\zeta,\tilde{D},D) \;=\; \kappa_{BC}\begin{pmatrix} -\zeta D & \zeta\tilde{D} \\ \zeta\tilde{D} & \zeta D \end{pmatrix} \;+\; \frac{\kappa_{BC}}{2}\,(\zeta^2+\tilde{D}^2+D^2)\,\mathbf{I}, \tag{5}$$

whose components define (T_{11}, T_{22}, T_{12}). The subgrid momentum forcing is the divergence of $\mathbf{T}$:

$$\mathbf{S} = \begin{pmatrix} S_u \\ S_v \end{pmatrix} = \nabla\cdot\mathbf{T} = \nabla\cdot\begin{pmatrix} T_{11} & T_{12} \\ T_{12} & T_{22} \end{pmatrix} = \begin{pmatrix} \partial_x T_{11} + \partial_y T_{12} \\ \partial_x T_{12} + \partial_y T_{22} \end{pmatrix} \tag{6}$$

Eq.(5)-Eq.(6) ensure consistent units and energy transfer signatures and are used to build HR data from HR velocity fields (u, v).

Data storage and parallelization: All computed diagnostics are stored in netCDF4 format, facilitating efficient access across multiple vertical levels. Visualization routines are used to inspect multi-layer fields. The parameterization script is executed in parallel across dates using Python multiprocessing.

4 Method

Figure 2 provides our framework of multiphysics super-resolution version 1.0 (MSR v1.0). It primarily includes a Dynamic Enhancement Feature (DEF), a High-Frequency Enhancement (HFE), a Coarse module and a Multi-Branch component. The DEF module extracts low-resolution physics features from X_{LR}, and focuses on key areas like shear layers and frontal zones, laying a physics-informed foundation for accurate reconstruction. The HFE effectively extracts high-frequency information and spatial information to restore fine structures and edge features, improving sharpness and generalization. The Coarse module provides the original low-resolution information to prevent the loss of essential contextual details that high-frequency enhancements might overshadow. The Multi-Branch module enhances the feature representation for each variable, thereby ensuring tailored optimization for different diagnostics and improving consistency across variables.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

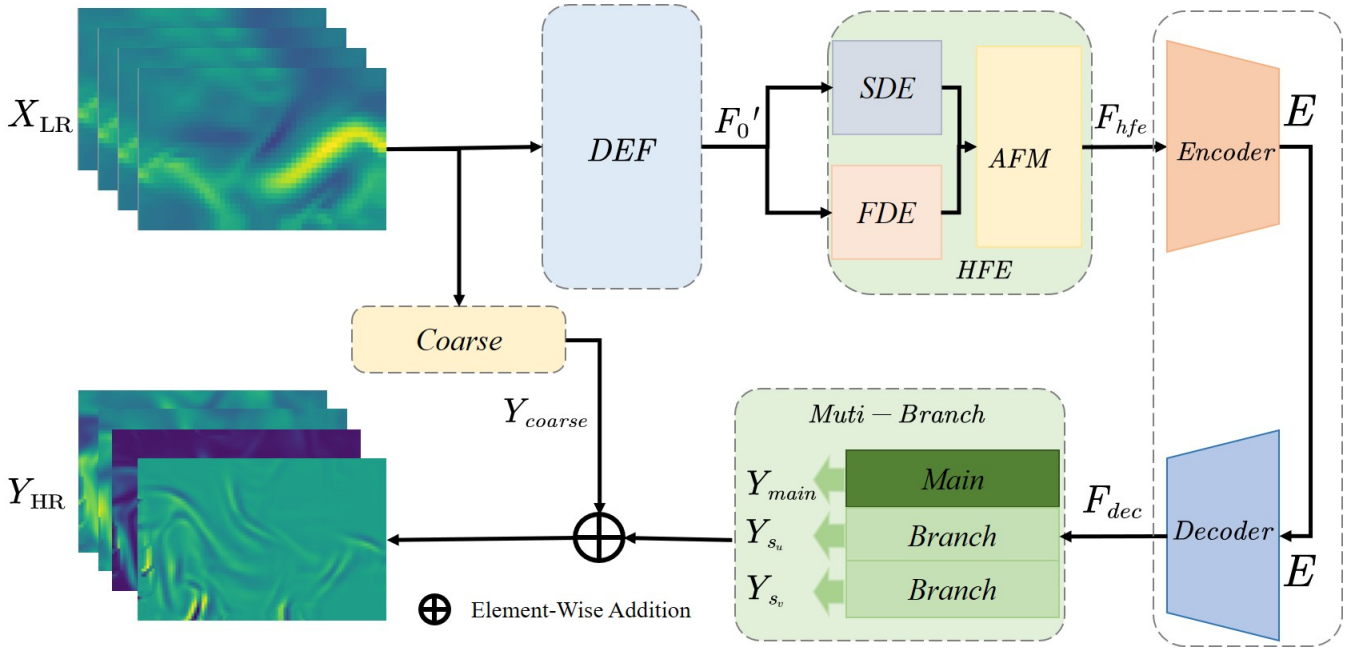

Figure 2. MSR architecture. We denote X_{LR} and Y_{HR} as the low-resolution input and high-resolution output. It includes four stages: (i) a Dynamic Enhancement Feature (DEF) stage, (ii) a High-Frequency Enhancement (HFE) stage, (iii) an encoder–decoder backbone, and (iv) a Coarse and Multi-Branch refinement stage.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

4.1 Dynamic Enhancement Feature (DEF)

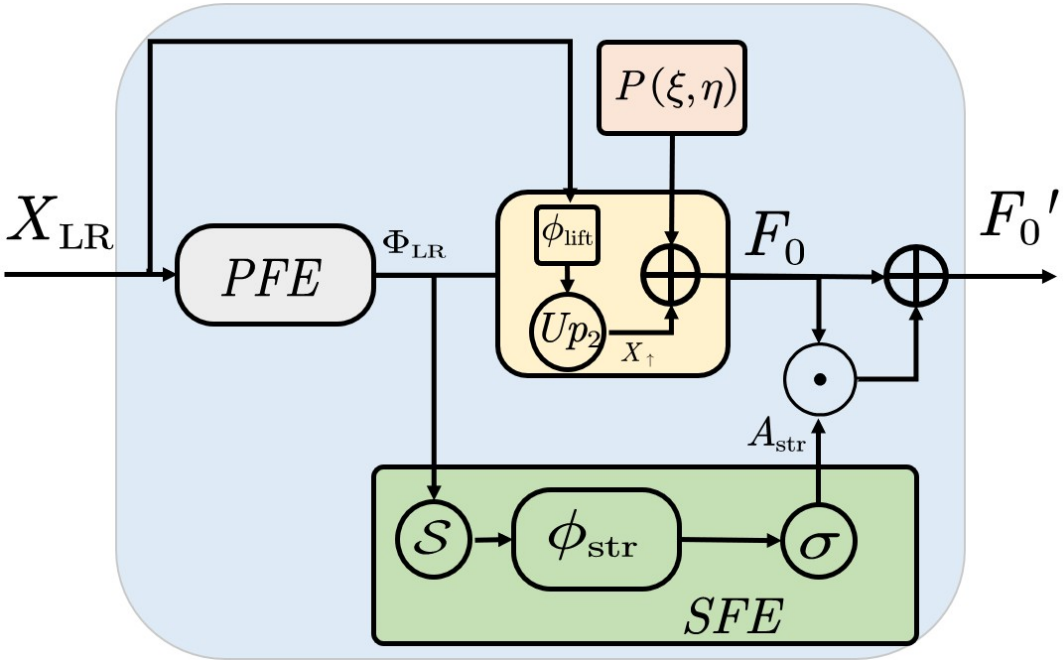

Figure 3. Dynamic Enhancement Feature (DEF) module. It extracts the physical information from X_{LR} and enhances it with SFE.

Given that the low-resolution velocity fields $X_{\mathrm{LR}} = (u, v) \in \mathbb{R}^{2\times \frac{H}{2}\times \frac{W}{2}}$ lack sufficient information to reconstruct a high-resolution result, we design a DEF module to extract more detailed information while preserving the underlying flow structure. This module consists of two key components: the Physical Feature Extractor (PFE) and the Strain Feature Enhancement (SFE).

Physical Feature Extractor (PFE): Like Gangopadhyay (2022), Hong et al. (2025), and Zhu et al. (2022), we compute internal descriptors $(\partial_x u, \partial_y u, \partial_x v, \partial_y v)$ from velocity fields (u, v). These descriptors capture important local variations in the velocity fields, which are critical for identifying fine-scale features such as fronts and eddy rims in ocean dynamics. They make the local Jacobian explicit, exposing shear, convergence/divergence, and rotational cues. Many ocean diagnostics (e.g., vorticity, divergence) are algebraic combinations of velocity gradients. Therefore, we explicitly provide the derivative fields $(\partial_x u, \partial_y u, \partial_x v, \partial_y v)$ to align the representation with the governing physics and to preserve small-scale gradients that low-resolution inputs tend to smooth out. This derivative stack reduces the need for the network to relearn finite-difference operators, improves sample efficiency and numerical stability, and yields sharper, more coherent reconstructions in dynamically active regions.

As shown in Figure 3, given low-resolution velocity fields $X_{\mathrm{LR}} = (u, v) \in \mathbb{R}^{2\times \frac{H}{2}\times \frac{W}{2}}$, we deterministically generate an auxiliary stack Φ_{LR}, it could help enrich the feature set by including additional information for the subsequent enhancement

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

process.

$$\Phi_{\mathrm{LR}}(u,v) = \big[\sigma_{\mathrm{low}}, \partial_x u, \partial_y u, \partial_x v, \partial_y v\big], \quad \sigma_{\mathrm{low}} = \partial_x u + \partial_y v, \tag{7}$$

where σ_{low} is used for establishing convenient relationships with the target variable D. We then upsample the low-resolution features X_{LR} to a high-resolution grid along with an auxiliary feature stack $\Phi_{\mathrm{LR}}(X_{\mathrm{LR}})$. Specifically, we perform an upsampling operation on the low-resolution data to boost it to the high-resolution space, generating a new feature map $X_{\uparrow}$:

$$X_{\uparrow} \;=\; \mathrm{Up}_2\Big(\phi_{\mathrm{lift}}\big([X_{\mathrm{LR}}, \Phi_{\mathrm{LR}}(X_{\mathrm{LR}})]\big)\Big) \;\in\; \mathbb{R}^{C_0\times H\times W}, \tag{8}$$

where $\mathrm{Up}_2(\cdot)$ means that the input feature will be upsampled by a factor of 2, and $\phi_{\mathrm{lift}}(\cdot)$ is the convolution operation on the input feature. This process transforms low-resolution data into features suitable for high-resolution result, ensuring enhanced model representation without introducing new input data. We inject normalized positional encodings to provide the model with the absolute position of each pixel in the spatial domain, thereby enhancing the model's ability to perceive spatial structures:

$$F_0 \;=\; X_{\uparrow} \;+\; \phi_{1\times1}\big(P(\xi,\eta)\big), \tag{9}$$

where $P(\xi,\eta)$ is a fixed coordinate mapping, ξ, η represent the height and width, and $\phi_{1\times1}(\cdot)$ denotes the projection operation performed by the 1×1 convolution. This allows the model to process spatial patterns in the image more accurately, especially with regard to specific regions such as the edges, center, and other critical areas of the image.

Strain Feature Enhancement (SFE): With this positional encoding providing spatial context, the next crucial step is to capture sharp-gradient features of the input data, which are essential for accurately modeling dynamic ocean features such as fronts and eddy rims (Sun et al., 2024; Liu and Liao, 2023). To achieve this, we design a Strain Feature Enhancement (SFE), which uses strain-based proxies to highlight regions of high spatial variation in the velocity fields.

Because strain measures rapid spatial changes in the velocity fields, it is a natural proxy for high-frequency detail. By calculating a strain proxy from the velocity gradients, we weight the input features based on the intensity of strain, allowing the model to focus on regions of the flow with significant dynamic variation. The strain characteristic $\mathcal{S}(u,v)$ captures the magnitude of changes in the velocity fields through spatial gradients, highlighting regions with rapid variations. It is calculated as follows:

$$\mathcal{S}(u,v) \;=\; \big|\partial_x u\big| \;+\; \big|\partial_y u\big| \;+\; \big|\partial_x v\big| \;+\; \big|\partial_y v\big|, \tag{10}$$

where $\mathcal{S}$ quantifies the flow's deformation in both the u and v components. Based on the Eq. (10), we generate a weighted attention map A_{str} as follows:

$$A_{\mathrm{str}} \;=\; \sigma\big(\phi_{\mathrm{str}}(\mathcal{S})\big), \tag{11}$$

where $\sigma(\cdot)$ is the sigmoid activation function, $\phi_{\mathrm{str}}(\cdot)$ is a convolution operation. A_{str} represents the intensity of attention. Finally, to amplify features in regions of higher strain, we apply the following weight adjustment to the input feature F_0:

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

$$F_0' = F_0 \odot A_{\text{str}} + F_0, \tag{12}$$

where $\odot$ denotes element-wise multiplication, the value of each position is multiplied by the attention weight of the corresponding position. This approach emphasizes the importance of high-frequency regions by giving them higher attention weights.

4.2 High-Frequency Enhancement (HFE)

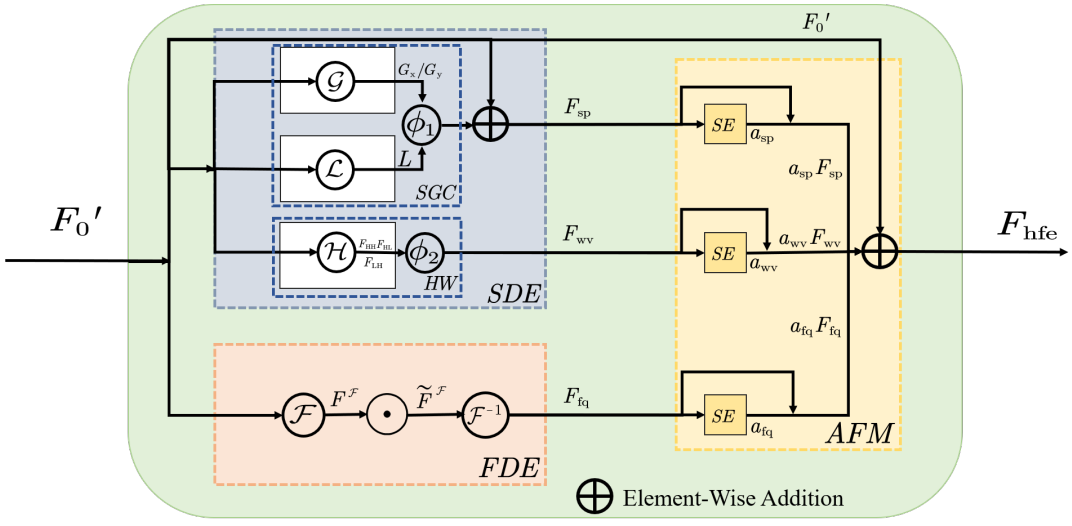

Figure 4. High-Frequency Enhancement (HFE) module. It enhances the features from spatial domain and frequency domain.

We design a HFE module to capture high-frequency information that is essential for accurately modeling dynamic ocean features such as fronts and eddy rims. It incorporates three key components: Spatial Domain Enhancement (SDE), Frequency Domain Enhancement (FDE), and Adaptive Fusion Mechanism (AFM).

Spatial Domain Enhancement (SDE): To effectively capture high-frequency details in the spatial domain, we apply three enhancement techniques: Sobel (spatial gradients), Laplacian (curvature), and Haar wavelet (multi-scale details). These techniques preserve important spatial features like edges and curvature, which are critical for identifying fine-scale structures (Huang et al., 2021).

(1) Spatial gradients and curvature (SGC): Given input F_0, we employ first-order (Sobel) and second-order (Laplacian) operators to extract edges and curvature (Wu et al., 2021; Ren et al., 2022):

$$G_x = F_0' * \mathcal{G}_{\text{Sx}}, \qquad G_y = F_0' * \mathcal{G}_{\text{Sy}}, \qquad L = F_0' * \mathcal{L}_{\Delta}, \tag{13}$$

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

where $\mathcal{G}_{\mathrm{Sx}}$ and $\mathcal{G}_{\mathrm{Sy}}$ are the convolution kernels of the Sobel operator in the x and y directions, the $\mathcal{L}_{\Delta}$ is the convolutional kernel of the Laplacian operator. Then we inject this spatial information into the network via a learnable mixer $\phi_1(\cdot)$,

$$F_{\mathrm{sp}} = F_0' + \phi_1\big([G_x, G_y, L]\big), \tag{14}$$

in this way, the gradient and curvature information are injected into the base feature map. While edge and curvature operators capture local gradients, we employ wavelet decomposition further to provide multi-scale details that complement these spatial cues.

(2) Haar wavelet (HW): It provides a way to capture local details at different scales and orientations by decomposing the image (Guo et al., 2022; Batziou et al., 2023). We first perform downsampling to decompose the features, then keep only sub-bands of details and recover them via upsampling. *Haar wavelet sub-band decomposition*: Using separable 1-D Haar low/high-pass filters along rows (first letter) and columns (second) with $2\times$ decimation, the image is decomposed by downsampling to extract four sub-bands: F_{LL} (Low-Frequency sub-band), F_{LH} (Horizontal Detail sub-band), F_{HL} (Vertical Detail sub-band), F_{HH} (Diagonal Detail sub-band):

$$\{F_{\mathrm{LL}}, F_{\mathrm{LH}}, F_{\mathrm{HL}}, F_{\mathrm{HH}}\} = \mathrm{Down}_2\big(F_0' * \{\mathrm{LL}, \mathrm{LH}, \mathrm{HL}, \mathrm{HH}\}\big), \tag{15}$$

where $\mathrm{Down}_2(\cdot)$ indicates a downsampling operation that reduces the image size by half. $\{\mathrm{LL}, \mathrm{LH}, \mathrm{HL}, \mathrm{HH}\}$ denote the four sub-bands of Haar wavelet analysis. *Haar Wavelet Sub-band Fusion*: We retain the detail sub-bands, restore the details through upsampling, and fuse them using a learnable fusion function $\phi_2(\cdot)$:

$$F_{\mathrm{wv}} = \phi_2\big(\mathrm{Up}_2(F_{\mathrm{LH}}), \mathrm{Up}_2(F_{\mathrm{HL}}), \mathrm{Up}_2(F_{\mathrm{HH}})\big), \tag{16}$$

where $\mathrm{Up}_2(\cdot)$ represents an upsampling operation that restores low-resolution detail sub-bands to the size of the original image. This spatial pathway effectively recovers localized edges and textures, yet it provides only indirect control over the image spectrum and may leave residual blur or aliasing. To complement it, we introduce a frequency-domain branch: operating on Fourier coefficients enables direct manipulation of spectral amplitudes, so we can selectively boost high-frequency components while preserving low-frequency consistency.

Frequency Domain Enhancement (FDE): In the frequency domain, we explicitly amplify the high-frequency amplitude while keeping the phase of the image constant (Liu et al., 2025). This is achieved by Fourier transform and high-pass mask. First, perform the Fourier transform on the image (F_0'), then adjust the amplitude of the high-frequency section by high-pass mask $M(\omega_x, \omega_y)$ and keep the phase constant. Finally, the image is recovered by the inverse Fourier transform.

$$F^{\mathcal{F}} = \mathcal{F}(F_0'), \qquad \widetilde{F}^{\mathcal{F}} = \big(1 + g\, M(\omega_x, \omega_y)\big) \odot F^{\mathcal{F}}, \qquad F_{\mathrm{fq}} = \mathcal{F}^{-1}(\widetilde{F}^{\mathcal{F}}), \tag{17}$$

where $\mathcal{F}(\cdot)$ represents the Fourier transform, and $F^{\mathcal{F}}$ is the frequency domain representation, $M(\omega_x, \omega_y)$ is the high-pass mask, ω_x, ω_y denote the spatial frequency components in the horizontal (x) and vertical (y) directions, x, y represent the spatial coordinates in the image domain. g is the gain factor, and $\odot$ represents element-by-element multiplication. $\mathcal{F}^{-1}$ represents the inverse Fourier transform, and F_{fq} is the gain-adjusted frequency domain image.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

Since each branch ($F_{\mathrm{sp}}, F_{\mathrm{wv}}, F_{\mathrm{fq}}$) emphasizes different aspects of high-frequency information, we employ an adaptive fusion mechanism to integrate them into a unified representation. As shown in Figure 5, we show feature maps with and without the Sobel, Laplacian, Haar, and FDE models. Each column corresponds to one operator, while the top row displays the initial feature maps, the middle row shows the enhanced feature maps, and the bottom row highlights the differences between the two. These comparisons illustrate the distinct contributions of each operator in enhancing fine-scale features such as edges, curvature, and high-frequency details, all of which are crucial for accurately capturing dynamic ocean features like fronts and eddy rims.

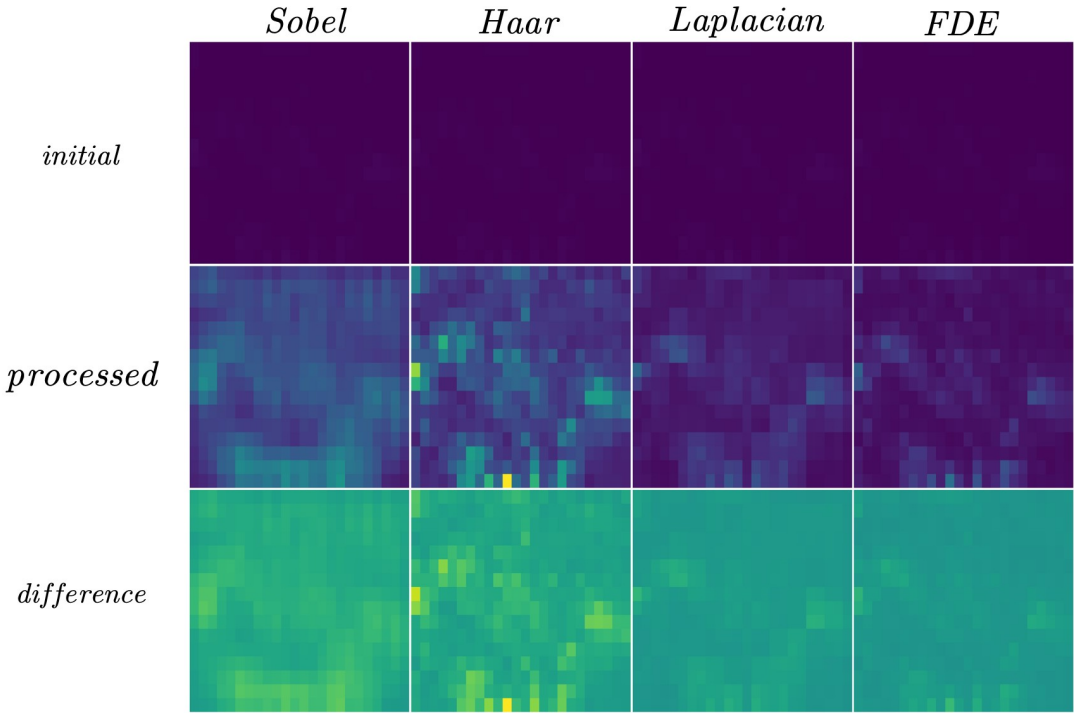

Figure 5. Feature maps with and without the Sobel, Haar, Laplacian, and FDE operators. Rows represent initial maps, processed maps, and their differences respectively. It illustrates how each operator enhances fine-scale features to support the capture of ocean fronts and eddy rims.

Adaptive Fusion Mechanism (AFM): To effectively fuse this high-frequency enhancement information, we employ a channel-level squeeze-excitation (SE) module to adaptively reweight different feature branches based on the global context (Hu et al., 2018). In each SE module: we apply global average pooling (GAP) to each branch (spatial branch (F_{sp}), waveform branch (F_{wv}), frequency branch (F_{fq})) and compute branch weights by a two-layer fully connected network:

$$a_b = \sigma\Big(W_b^{(2)}\,\delta\big(W_b^{(1)}\,\mathrm{GAP}(F_b)\big)\Big), \quad b \in \{\mathrm{sp}, \mathrm{wv}, \mathrm{fq}\}, \tag{18}$$

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

where $\sigma(\cdot)$ is the sigmoid activation function, $\delta(\cdot)$ is the ReLU activation, and $W_b^{(1)}$ and $W_b^{(2)}$ are the learnable weight matrices. To adaptively fuse the reweighted branches, we perform a gated residual aggregation:

$$F_{\mathrm{hfe}} = F_0' + a_{\mathrm{sp}} F_{\mathrm{sp}} + a_{\mathrm{wv}} F_{\mathrm{wv}} + a_{\mathrm{fq}} F_{\mathrm{fq}}. \tag{19}$$

where a_{sp}, a_{wv}, and a_{fq} are the weighting coefficients learned through the SE module.

As shown in Figure 6, we present a comparison between the features before and after the HFE process. The top-left, top-right and bottom-left panels show the magnitude of the input features ("F_0'"), the enhanced output features ("F_{hfe}"), and their difference ("$F_{\mathrm{hfe}} - F_0'$"), respectively. This highlights the changes made during the enhancement process. On the bottom-right, the radial power spectrum illustrates the boost in medium to high spatial frequencies in the enhanced features (orange curve) compared to the original features (blue curve). hp_gain (here 2.49) indicates a $2.49\times$ increase in average high-frequency power, demonstrating the significant enhancement in fine-scale details. To produce the right-hand curve and quantify the boost, we compute a *radial power spectrum* from the features before and after HFE. Specifically, each multi-channel feature Z is first L2-aggregated into a single map $X(i,j) = \sqrt{\sum_c Z_c(i,j)^2}$. We then take its 2-D DFT $F = \mathcal{F}\{X\}$ and define the spectral power as $\Phi(u,v) = |F(u,v)|^2$ (using Φ to avoid symbol clashes). After shifting the zero frequency to the center, we form radial bins by the distance

$$r(u,v) = \sqrt{(u-u_0)^2 + (v-v_0)^2}, \tag{20}$$

and perform azimuthal averaging over $N = 40$ concentric annuli $\Omega_b = \{(u,v) : r_b \le r(u,v) < r_{b+1}\}$ to obtain

$$\bar{\Phi}(r_b^*) = \frac{1}{|\Omega_b|} \sum_{(u,v)\in\Omega_b} \Phi(u,v), \qquad r_b^* = \tfrac{1}{2}(r_b + r_{b+1}). \tag{21}$$

We plot the radial spatial frequency r_b^* (pixel units) on the abscissa and $\bar{\Phi}(r_b^*)$ on a logarithmic ordinate (adding a small $\varepsilon = 10^{-12}$ for numerical stability). A dashed line marks the cutoff $r_0 = 0.25\, r_{\max}$ that separates low and high bands, and we report

$$\mathrm{hp\_gain} = \frac{\langle \bar{\Phi}_{\mathrm{after}}(r) \rangle_{r>r_0}}{\langle \bar{\Phi}_{\mathrm{before}}(r) \rangle_{r>r_0}}. \tag{22}$$

so that the orange curve's uplift at medium–high r directly translates into the reported gain values.

This frequency spectrum analysis provides a deeper understanding of the enhancement process. The low-frequency segment (left side) of the spectrum shows the typical smoothness of natural ocean features, where the power decreases as frequency increases. However, the enhanced spectrum (orange curve) clearly exhibits increased power in the higher-frequency range (right side), corresponding to more detailed, sharp features like edges and textures. The After curve in the high-frequency region shows a noticeable increase compared to the Before curve, which signifies that the enhancement process has indeed successfully emphasized these fine-scale details. The hp_gain metric quantifies this enhancement and validates that the network has effectively boosted high-frequency components, as intended. This aligns with the goal of improving fine details such as eddy rims and fronts while maintaining overall structural stability in the output.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

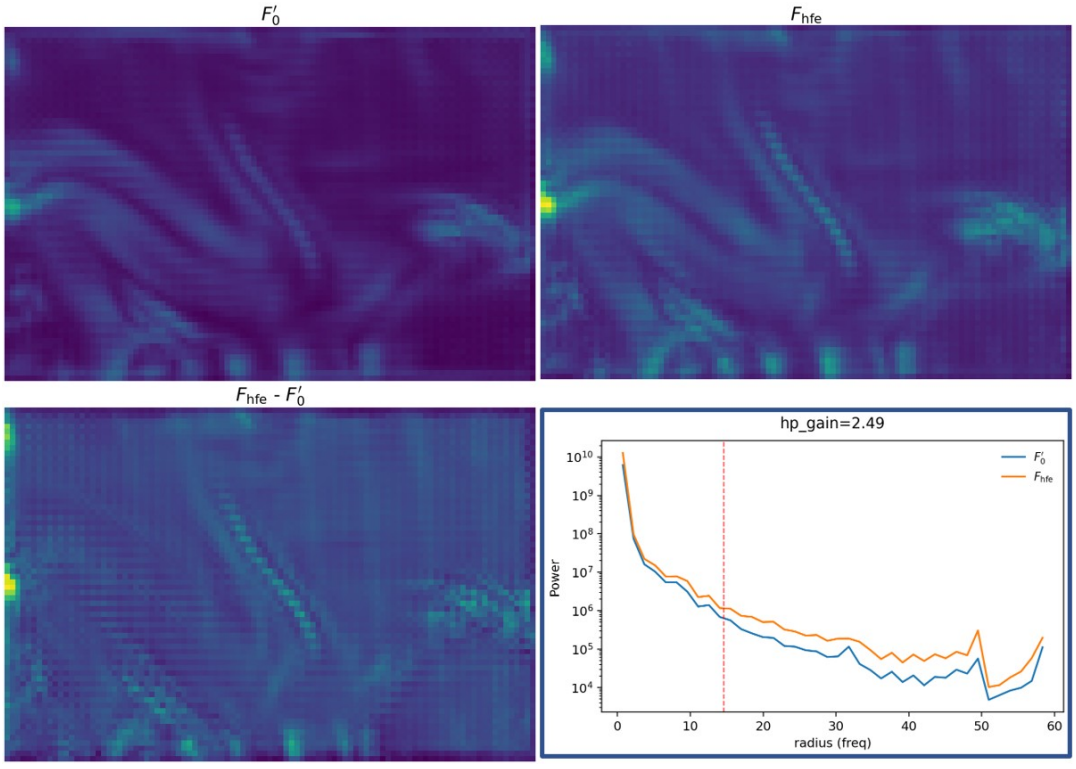

Figure 6. Fused feature F_{hfe} with enhancement, F_0' without enhancement and their radial power spectrum hp_gain. $F_{\mathrm{hfe}} - F_0'$ is their difference. It shows that the orange curve has greater power at medium–high spatial frequencies than the blue curve. The vertical dashed line marks the cutoff separating low and high frequencies used to compute hp_gain.

4.3 Encoder-Decoder Backbone

High-resolution ocean diagnostics require the ability to capture both large-scale dynamics, such as mesoscale coherence, and fine-scale features, such as fronts and filaments. To achieve this, we design an encoder-decoder module that is specifically tailored to preserve important spatial details while handling multi-scale ocean processes. As shown in Figure 2, MSR is further enhanced with attention mechanisms and residual-dense blocks, which improve the model's stability and the quality of feature aggregation during both encoding and decoding stages.

A succinct representation of the encoder-decoder process is as follows:

$$E = \mathrm{Enc}(F_{\mathrm{hfe}}), \tag{23}$$

$$F_{\mathrm{dec}} = \mathrm{Dec}(E), \tag{24}$$

where F_{hfe} refers to the high-frequency enhanced features. In the encoder, the ASPP module expands the receptive fields, the channel attention module selectively focuses on important features, and the RRDB stabilizes the residual aggregation process.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

The **encoder** (Enc) consists of stacked residual convolution blocks, followed by an Atrous Spatial Pyramid Pooling (ASPP) module with dilations $\{1, 3, 6\}$ (Liu et al., 2022). This allows the network to capture a broad range of spatial dependencies, effectively expanding the receptive fields without losing resolution. The ASPP module is crucial for resolving large-scale dynamics in ocean fields, such as mesoscale eddies or currents, where different spatial scales need to be captured simultaneously. In addition, channel attention is applied to highlight the most relevant features across different spatial scales, allowing the model to focus on dynamic regions and avoid unnecessary computational overhead in less informative areas.

In the **decoder** (Dec), the encoded feature maps are upsampled and refined through DetailEnhanced blocks and Residual in Residual Dense Blocks (RRDB) (Ibrahim et al., 2022). The DetailEnhanced blocks are particularly effective for reconstructing sharp boundaries and fine-scale gradients, which are important in oceanographic applications such as front detection or eddy analysis. The RRDBs help to stabilize the aggregation of residual features, ensuring that fine-scale structures are preserved while mitigating potential issues like gradient explosion or overfitting. Finally, spatial attention is applied at the high-resolution stage to ensure that the network focuses on key spatial regions that are most critical for the diagnostic task, such as regions of high vorticity or significant deformation.

This encoder-decoder design is carefully crafted to capture both global dynamics and local features in high-resolution ocean diagnostics, offering a robust solution for reconstructing complex ocean processes from low-resolution inputs.

4.4 Multi-branch

Ocean variables inherently exhibit varying spectral characteristics, which poses a challenge when reconstructing high-resolution ocean diagnostics (Tian et al., 2024; Gao et al., 2024). For smooth-spectrum variables, which feature gradual transitions and lower-frequency components (such as vorticity, deformation, and stretching deformation), shared spatial representations are highly effective. These variables can be processed efficiently using a common spatial head, which preserves low-frequency components without sacrificing the detail necessary for accurate reconstruction. However, high-frequency features—such as stress divergences, which involve rapid spatial variations—are more challenging. For these, standard spatial domain convolution operations fail to capture the fine-grained details adequately. These high-frequency regions require frequency-domain enhancements, which allow for the preservation of critical features without introducing artifacts like over-smoothing. To address these challenges, we design a **Multi-Branch** module, it separates the processing of smooth-spectrum and high-frequency variables. This design efficiently handles the diverse nature of the target variables, improving model accuracy across scales. The module comprises two main components: the **Main** for smooth-spectrum variables and the **Branches** for high-frequency variables.

Main for Smooth-Spectrum Variables: To handle smooth-spectrum variables like vorticity (ζ), deformation ($\tilde{D}, D$), and stress tensor ($\mathbf{T}$), which benefit from shared spatial features, we design a main head. This component applies a spatial convolution operation to synthesize these variables while ensuring that the high-frequency details are not overmodulated and low-frequency structures are effectively preserved. Specifically, the Main head outputs six key diagnostic variables, which are mapped directly from the decoder's feature map. The purpose is to reconstruct the variables accurately in the low-frequency spectrum, where shared spatial representations excel. Importantly, this approach maintains the spectral integrity of smooth-

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

spectrum variables, preventing any significant loss of detail while emphasizing both global coherence and local accuracy.

$$\widehat{Y}_{\text{main}} = \{Y_\zeta, Y_{\tilde{D}}, Y_D, Y_{T_{11}}, Y_{T_{22}}, Y_{T_{12}}\} \tag{25}$$

$$\widehat{Y}_{\text{main}} = \phi_{\Omega_{\text{main}}}(F_{\text{dec}}) \in \mathbb{R}^{6\times H\times W}, \quad \Omega_{\text{main}} = \{\zeta, \tilde{D}, D, T_{11}, T_{22}, T_{12}\}, \tag{26}$$

where $\phi_{\Omega_{\text{main}}}(\cdot)$ is a spatial convolution operation, and F_{dec} is the feature map output by the decoder. The head generates mapping results for six target variables through processing of decoder outputs. The main goal of the head is to synthesize the smooth-spectral variables, ensuring that the high frequencies are not overmodulated and that the low frequencies are properly preserved.

Although the main head works well for smooth-spectrum variables, S_u and S_v involve high-frequency details, we need to better capture using frequency-based enhancements.

Branches: For subgrid forcing terms S_u and S_v, predominantly containing high-frequency details, conventional spatial-domain convolutions are insufficient to recover these small-scale variations. These features require frequency-domain processing for accurate reconstruction. To address this, we design a dedicated head in the **Branches** that performs high-frequency enhancement directly in the frequency domain. This operation involves computing the pre-activation value (Q_k) via a frequency-specific convolution, followed by high-frequency gain in the Fourier domain, denoted by $\mathcal{F}$ (the Fourier transform operator). The high-frequency section is amplified using a gain factor (g_k) and a thresholding mechanism based on the frequency radius (ρ), which ensures that only the relevant high-frequency components are enhanced. Finally, the inverse Fourier transform ($\mathcal{F}^{-1}$) is applied to recover the final output $\widehat{Y}_k$ ($\widehat{Y}_{S_u}, \widehat{Y}_{S_v}$) in the spatial domain, maintaining both the structure and integrity of the high-frequency details.

$$Q_k = \varphi_k(F_{\text{dec}}), \quad k \in \{S_u, S_v\}, \tag{27}$$

$$\widehat{Y}_k = \mathcal{F}^{-1}\Big(\big(1 + g_k\, \mathbf{1}_{\{\rho\geq\tau_k\}}\big) \odot \mathcal{F}(Q_k)\Big), \quad \rho = \sqrt{\omega_x^2 + \omega_y^2}, \tag{28}$$

where φ_k denotes the convolution operation of the frequency-specific header, g_k is the gain factor, $\rho = \sqrt{\omega_x^2 + \omega_y^2}$ is the radial distance of the frequency space, τ_k is the frequency threshold, and $\mathbf{1}(\cdot)$ is the indicator function which has a value of 1 when $\rho \geq \tau_k$ and 0 otherwise. This operation enhances the high-frequency section within the frequency domain and ensures that no additional phase changes are introduced, maintaining the structural integrity of the image. In this way, the high-frequency parts of S_u and S_v are effectively enhanced to better reconstruct strain information.

4.5 Coarse Module

To preserve the essential low-frequency information from X_{LR} and to make sure it is faithfully expressed in the outputs (Y_{HR}), we introduce a **Coarse module**. In order to maintain spatially accurate large-scale structure, a low-pass, spatially aligned copy

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

of the input is combined with the decoded result at the last stage. Concretely, we downsample the input and then upsample it back to its original resolution to obtain a coarse surrogate; this upsampled surrogate is added to the network output as a residual. In this way, the coarse path acts as a shallow, structure-preserving skip: it conveys basin-scale and scene-scale patterns that are often under-represented by purely high-frequency enhancement, while keeping the alignment with the original input.

This approach serves as a safeguard against the loss of low-frequency dynamics that might otherwise be overshadowed by high-frequency processing. By anchoring the reconstruction with a filtered, spatially consistent coarse component, the model preserves global context and prevents structural drift, while leaving the specialized enhancement branches free to recover edges, fronts, and other fine-scale details. As a result, the final maps retain the broad spatial organization of the scene, and the high-frequency modules can focus on sharpening without distorting large-scale features.

$$Y_{\text{coarse}} = \text{Up}\Big(\phi_{\text{coarse}}(\text{Pool}(X_{\text{LR}}))\Big) \in \mathbb{R}^{8\times H\times W}, \tag{29}$$

where $\text{Pool}(\cdot)$ represents the pooling operation, $\text{Up}(\cdot)$ represents the upsampling operation, and $\phi_{\text{coarse}}(\cdot)$ is the convolution that processes of low-frequency features. The final result concatenates the six main variables with $(\widehat{S}_u, \widehat{S}_v)$ and adds the coarse scaffold:

$$Y_{\text{HR}} = \Pi\big(\widehat{Y}_{\text{main}}, Y_{\widehat{S}_u}, Y_{\widehat{S}_v}\big) + Y_{\text{coarse}} \in \mathbb{R}^{8\times H\times W}, \tag{30}$$

where $\Pi(\cdot)$ represents the concatenation of the six main variables with the high-resolution versions of S_u and S_v, followed by the addition of the coarse Y_{coarse}. This ensures that the large-scale low-frequency structures from the original input are preserved and effectively integrated into the final high-resolution output, offering a solid foundation for the high-frequency features to build upon. By including this coarse scaffold, we prevent the loss of crucial contextual information while still enhancing fine details.

4.6 Training losses and optimization

Denote $\widehat{Y} = \{\widehat{Y}_{main}, \widehat{Y}_{\widehat{S}_u}, \widehat{Y}_{\widehat{S}_v}\} = \{\widehat{\zeta}, \widehat{\tilde{D}}, \widehat{D}, \widehat{T}_{11}, \widehat{T}_{22}, \widehat{T}_{12}, \widehat{S}_u, \widehat{S}_v\}$ as outputs and Y represents HR labels.

Pixel loss (multi-target)

To achieve accurate inversion of multiple physical quantities (Cai et al., 2022), we combine ℓ_1 and ℓ_2 losses to calculate the error between $\widehat{Y}$ and Y. For all eight physical quantities:

$$\mathcal{L}_{\text{pix}} = \sum_{c\in\Omega} \big(\alpha_1 \|\widehat{Y}_c - Y_c\|_1 + \alpha_2 \|\widehat{Y}_c - Y_c\|_2^2\big), \quad \Omega = \{\zeta, \tilde{D}, D, T_{11}, T_{22}, T_{12}, S_u, S_v\}, \tag{31}$$

where α_1 and α_2 are hyperparameters that regulate the loss weights of ℓ_1 and ℓ_2.

Spectral amplitude loss

Standard pixel-wise losses (e.g., L1 or L2) often lead to overly smooth results, suffering from spectral bias—specifically, a deficit in high-frequency energy. To mitigate this, we introduce a spectral amplitude loss that constrains the model in the frequency domain. This loss forces the reconstructed output to match the spectral statistics of the ground truth, thereby ensuring

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

realistic energy distribution across scales:

$$\mathcal{L}_{\text{spec}} = \left\| \left|\mathcal{F}(\widehat{Y})\right| - \left|\mathcal{F}(Y)\right| \right\|_2^2, \tag{32}$$

where $\mathcal{F}(\cdot)$ denotes the Fast Fourier Transform (FFT), and $|\cdot|$ represents the magnitude (amplitude) of the complex spectrum. By minimizing the discrepancy in spectral amplitudes, the model is encouraged to generate finer details that are often lost in the spatial domain.

Gradient Loss

Since high-frequency information (such as fronts and eddies) is primarily encoded in the rapid spatial variations of the velocity fields, pixel-level losses alone may fail to capture sharp transitions. To address this, we incorporate a gradient loss that explicitly penalizes the discrepancies in first- and second-order spatial derivatives:

$$\mathcal{L}_{\text{grad}} = \|\nabla_x \widehat{Y} - \nabla_x Y\|_1 + \|\nabla_y \widehat{Y} - \nabla_y Y\|_1 + \|\Delta \widehat{Y} - \Delta Y\|_1, \tag{33}$$

where ∇_x and ∇_y denote the horizontal and vertical gradients, respectively, and Δ denotes the Laplacian operator (second-order derivative). We utilize the L1 norm here to promote sharpness in the reconstructed edges and to reduce the smoothing effect typically associated with L2 minimization.

Physical Consistency Loss

In order to ensure the physical consistency of the model output, we construct the stress tensor based on the inverted strain invariants ($\widehat{\zeta}$, $\widehat{\tilde{D}}$ and $\widehat{D}$) and compare them with the inverted stress tensor. By calculating the difference in stress tensor, we can guarantee that the inversion results are physically consistent. The formula for calculating the stress tensor $\mathbf{T}$ is as follows:

$$\mathbf{T}(\widehat{\zeta}, \widehat{\tilde{D}}, \widehat{D}) = \kappa_{BC} \begin{pmatrix} -\widehat{\zeta}\widehat{D} & \widehat{\zeta}\widehat{\tilde{D}} \\ \widehat{\zeta}\widehat{\tilde{D}} & \widehat{\zeta}\widehat{D} \end{pmatrix} + \frac{\kappa_{BC}}{2}\left(\widehat{\zeta}^2 + \widehat{\tilde{D}}^2 + \widehat{D}^2\right)\mathbf{I},$$

where κ_{BC} is a physical constant, and $\mathbf{I}$ is an identity matrix. The physical consistency loss function is:

$$\mathcal{L}_{\text{phy}} = \beta_T \left\| \widehat{\mathbf{T}} - \mathbf{T}(\widehat{\zeta}, \widehat{\tilde{D}}, \widehat{D}) \right\|_1 + \beta_S \left\| \widehat{\mathbf{S}} - \mathbf{S} \right\|_1,$$

where $\widehat{\mathbf{T}} = \begin{pmatrix} \widehat{T_{11}} & \widehat{T_{12}} \\ \widehat{T_{12}} & \widehat{T_{22}} \end{pmatrix}$ is the mapping stress tensor, $\widehat{\mathbf{S}} = (\widehat{S_u}, \widehat{S_v})$ is the mapping subgrid momentum forcing. β_T and β_S are hyperparameters that adjust the loss weights of stress tensor. These loss terms ensure that the model's outputs are consistent with the theoretical stress relationships, which are essential for maintaining the physical validity of the reconstructed variables.

Total loss and optimization

The final objective is

$$\mathcal{L} = \mathcal{L}_{\text{pix}} + \lambda_{\text{spec}}\mathcal{L}_{\text{spec}} + \lambda_{\text{grad}}\mathcal{L}_{\text{grad}} + \lambda_{\text{phy}}\mathcal{L}_{\text{phy}} \tag{34}$$

where λ terms are the weight hyperparameters of each loss term, they are obtained by tuning on the validation set.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

5 Experiments

To rigorously validate the proposed multi-physical super-resolution framework, we conducted a series of seven targeted experiments based on the LICOM3 eddy-resolving dataset. These experiments systematically assess the model's performance in terms of accuracy, generalization, physical consistency, and sensitivity to architectural components and downscaling factors. The experimental setup utilized two NVIDIA GeForce GTX 1080Ti GPUs. We evaluated the effectiveness of our method using Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R^2) scores as evaluation metrics.

5.1 Experiment 1: Baseline Model Comparisons in a Single Region

The purpose of this experiment is to compare the performance of the proposed MSR with four strong baselines—LowResUNet (Low), HighResUNet (trained on HR inputs), Enhanced Deep Residual Networks (EDSR) (Abidin et al., 2025), and Fourier Neural Operator (FNO) (Li et al., 2020) on a fixed North Pacific subdomain. This comparison helps assess if MSR can outperform these models in reconstructing high-resolution ocean diagnostics from low-resolution inputs. The models (Low, EDSR, FNO) use low-resolution velocities as input to reconstruct eight high-resolution (HR) diagnostics: $\{\zeta, D, \tilde{D}, T_{11}, T_{22}, T_{12}, S_u, S_v\}$. Metrics (RMSE, R^2) are reported on normalized fields over valid ocean points.

The results, shown in Table 1 and Table 2, demonstrate that our method significantly outperforms the best baseline, HighResUNet, reducing the average RMSE from 0.126 to 0.113 (-10.5%) and achieving the highest mean R^2 of 0.947. The most notable improvements in RMSE occur for the variables T_{22} (-20.3%), T_{12} (-13.3%), ζ (-13.2%), and D (-12.3%). While improvements for S_u (-4.7%) and S_v (-1.5%) are more modest, they remain consistent across all tests. For $\tilde{D}$, the results are nearly identical to HighResUNet (-0.5%). In terms of R^2 scores, our method excels for rotation and tensor diagnostics (ζ, T_{11}, T_{22}, T_{12}), with values ranging from 0.960 to 0.994, and it remains competitive for subgrid forcings ($S_u = 0.847$, $S_v = 0.887$).

Table 1. RMSE of the diagnostic variables obtained from the different models in Experiment 1. Lower RMSE values indicate better performance. Boldface indicates the best value in each row.

Variable	LowResUNet	HighResUNet	EDSR	FNO	Ours
ζ	0.259	0.096	0.091	0.270	**0.079**
D	0.625	0.183	0.351	0.736	**0.179**
$\tilde{D}$	0.274	0.111	0.123	0.308	**0.110**
T_{11}	0.253	0.123	0.120	0.264	**0.109**
T_{22}	0.235	0.109	0.107	0.237	**0.085**
T_{12}	0.224	0.113	0.107	0.237	**0.093**
S_u	0.268	0.157	0.160	0.278	**0.150**
S_v	0.243	0.116	0.143	0.247	**0.114**
Mean	0.297	0.126	0.150	0.322	**0.113**

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

Table 2. R^2 of the diagnostic variables obtained from the different models in Experiment 1. Higher R^2 values indicate better performance. Boldface indicates the best value in each row.

Variable	LowResUNet	HighResUNet	EDSR	FNO	Ours
ζ	0.934	0.991	0.992	0.928	**0.994**
D	0.506	0.958	0.844	0.314	**0.958**
$\tilde{D}$	0.913	0.986	0.982	0.890	**0.986**
T_{11}	0.787	0.950	0.952	0.768	**0.960**
T_{22}	0.808	0.959	0.960	0.804	**0.975**
T_{12}	0.822	0.955	0.959	0.799	**0.969**
S_u	0.513	0.832	0.825	0.476	**0.847**
S_v	0.488	0.883	0.822	0.469	**0.887**
Mean	0.721	0.939	0.917	0.681	**0.947**

In qualitative comparisons, shown in Figure 7, the proposed model preserves sharper eddy rims and frontal filaments compared to LowResUNet and FNO, and avoids the mild blurring observed in EDSR for tensor components. The error maps in Figure 8 further highlight that the residuals produced by our model are weaker and less spatially coherent along shear bands, especially for T_{11} and ζ. The improvements for S_v are subtle but consistent with the small RMSE/R^2 gains observed in the quantitative results (Table 1–Table 2).

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

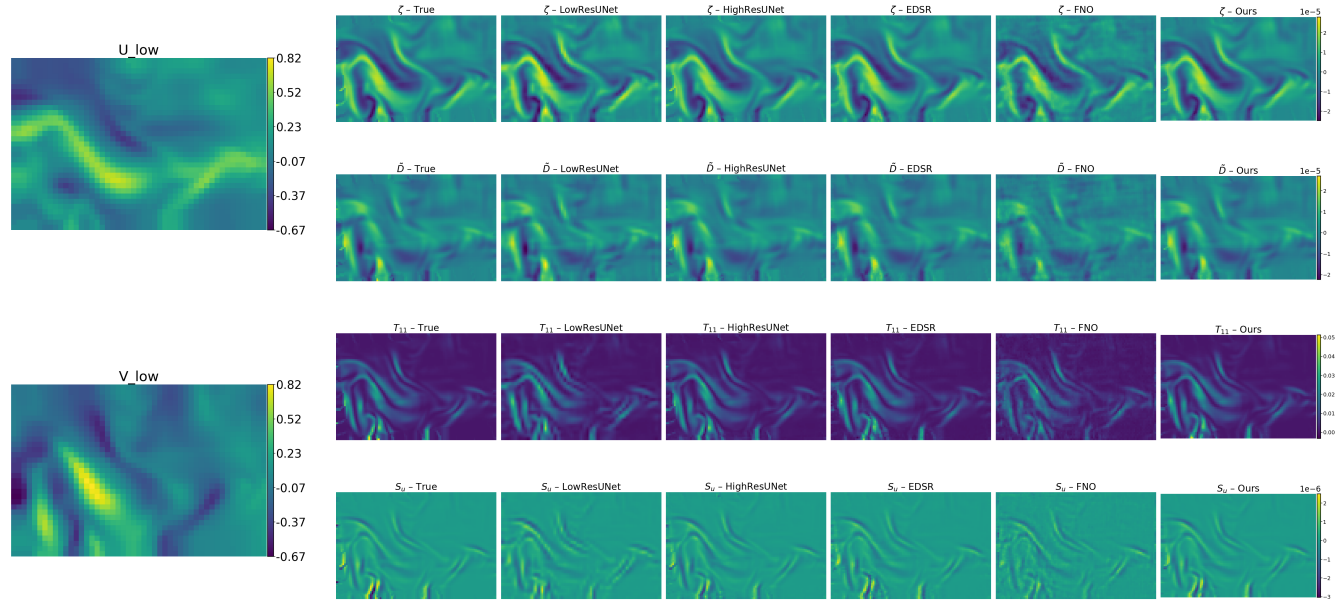

Figure 7. Qualitative comparison. U_low and V_low represent the low-resolution inputs, the other columns represent the labels and models' outputs. The rows represent the ocean diagnostics.

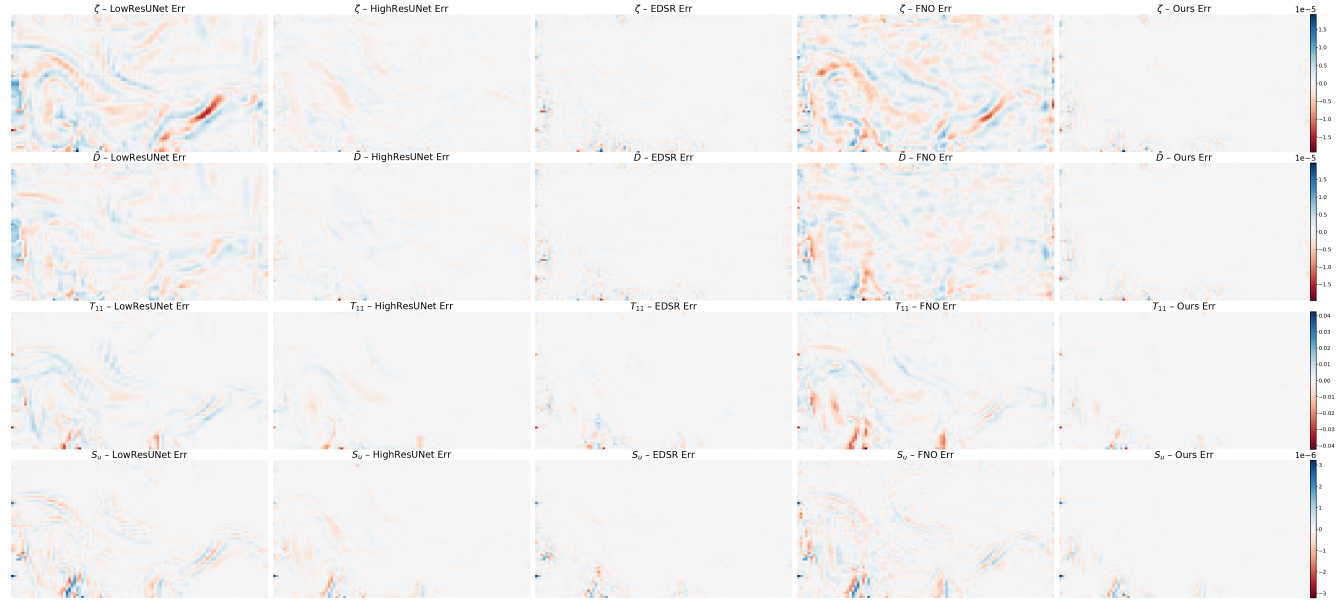

Figure 8. Absolute error maps corresponding to the snapshot in Fig. 7. This figure visualizes the reconstruction error magnitude by calculating the absolute deviation between the truth and the output of each model.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

The key takeaways from this experiment are as follows: (1) The HFE and DEF of our model lead to consistent improvements, particularly for tensor and rotational diagnostics such as T_{22}, T_{12}, ζ, and D. (2) While subgrid forcings like S_u and S_v are more challenging to map, our model still offers systematic improvements in these areas. (3) The high R^2 values and decorrelated error patterns along fronts suggest that our method recovers derivative-level structures better than models using only low-resolution inputs, where such structures are typically suppressed.

5.2 Experiment 2: Multi-Region Generalization

The purpose of this experiment is to evaluate the generalization ability of MSR across different ocean basins, assessing if it maintains performance despite geographical and dynamical variability. By repeating Experiment 1 in three dynamically distinct basins—Region 1 (North Pacific, 35°–40°N, 150°–160°E), Region 2 (South Pacific, 40°–35°S, 160°–150°W), and Region 3 (North Atlantic, 35°–40°N, 40°–30°W)—we test whether MSR can consistently outperform baseline models across regions with different oceanic characteristics. All methods operate on the same low-resolution velocity inputs (the proposed *MSR* additionally uses gradient features) to reconstruct the eight high-resolution diagnostics $\{\zeta, D, \tilde{D}, T_{11}, T_{22}, T_{12}, S_u, S_v\}$; evaluation metrics (MSE, RMSE, R^2) are computed on normalized fields restricted to valid ocean points via land masks. Figures 9–11 report per-variable MSE, RMSE, and R^2 for each basin. Across all three regions, *MSR* consistently achieves the lowest errors and the highest R^2 for most variables, with the clearest margins for deformation/forcing terms (D, S_u, S_v) where high-wavenumber content is critical; tensor/rotational diagnostics (ζ, T_{11}, T_{22}, T_{12}) maintain high R^2 with only modest basin-to-basin spread, indicating that the physics-guided attention and dual spatial–frequency branches transfer effectively across regimes. The North Atlantic (Region 3) is the most energetic among the three, with sharper fronts and stronger shear that amplify baseline errors—especially for D and S_u—yet our model preserves a clear lead in both RMSE and R^2, evidencing improved resilience to mesoscale variability. Overall, the method ranking is stable across basins (Ours > EDSR > FNO on average), and the inter-basin variability of our scores is modest, suggesting robust generalization without region-specific tuning.

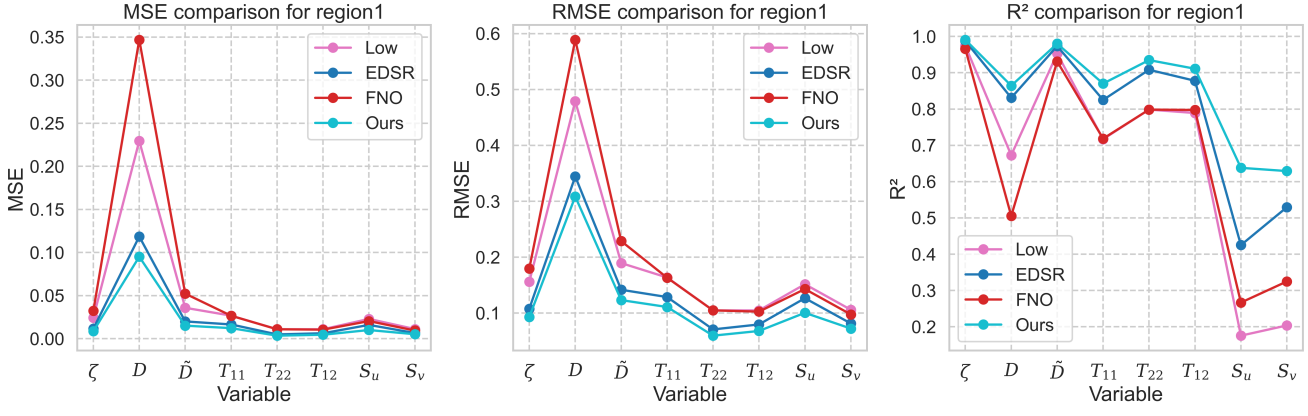

Figure 9. Region 1 (North Pacific). Per-variable MSE (left), RMSE (middle), and R^2 (right) for Low (baseline), EDSR, FNO, and the proposed method (*MSR*). Our method consistently obtains the lowest errors, the highest R^2 with the clearest gains on D, S_u, and S_v.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

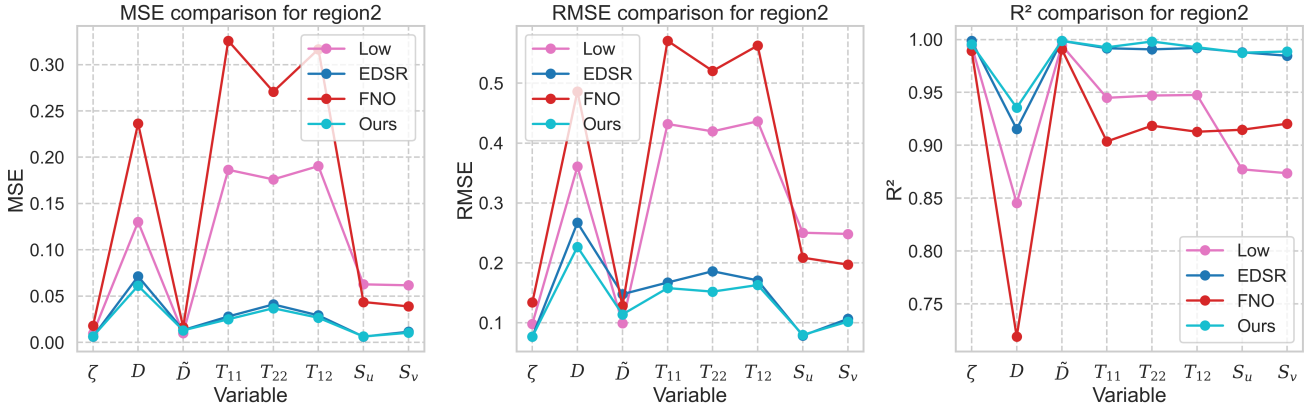

Figure 10. Region 2 (South Pacific). Performance remains strong across variables; note the robust R^2 for tensor components (T_{11}, T_{22}, T_{12}) and improved RMSE on D, S_u, S_v relative to baselines.

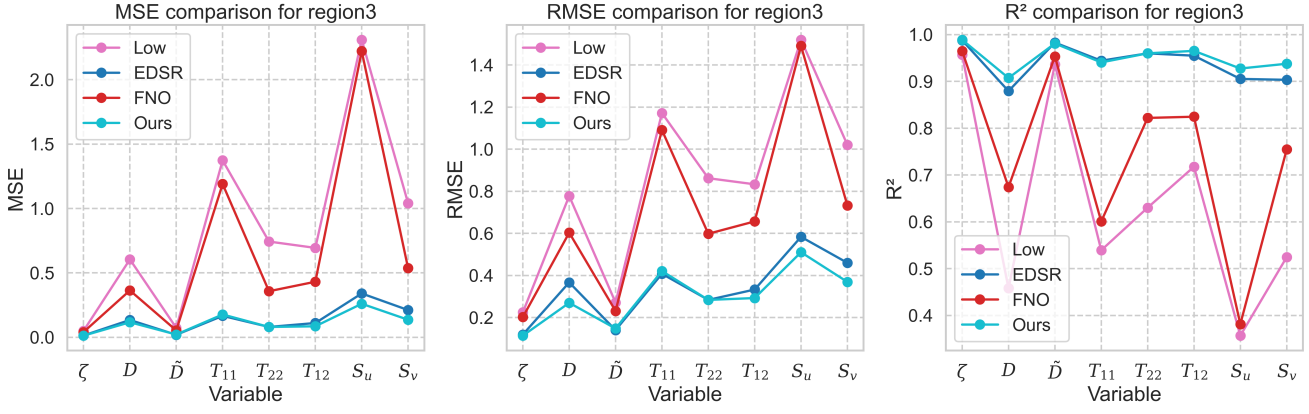

Figure 11. Region 3 (North Atlantic). In this energetic region, baseline errors increase notably for D and S_u. The proposed model consistently preserves the performance lead, demonstrating robust handling of sharp fronts and strong shear.

5.3 Experiment 3: Sensitivity to Downscaling Factors

The purpose of this experiment is to assess MSR's robustness to various downscaling factors, where $\times 2$, $\times 4$, and $\times 8$ represent reductions in input resolution. By evaluating the model's performance at these different scales, we can determine if MSR remains effective as input resolution decreases, thereby testing its capacity to recover fine-scale ocean features despite losing high-frequency details in the low-resolution inputs. The data split, preprocessing, and masking strategy are identical to Experiments 1–2. Metrics are computed over the eight target diagnostics $\{\zeta, D, \tilde{D}, T_{11}, T_{22}, T_{12}, S_u, S_v\}$ and then averaged to report mean MSE, RMSE, and R^2. Figure 12 summarizes the results.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

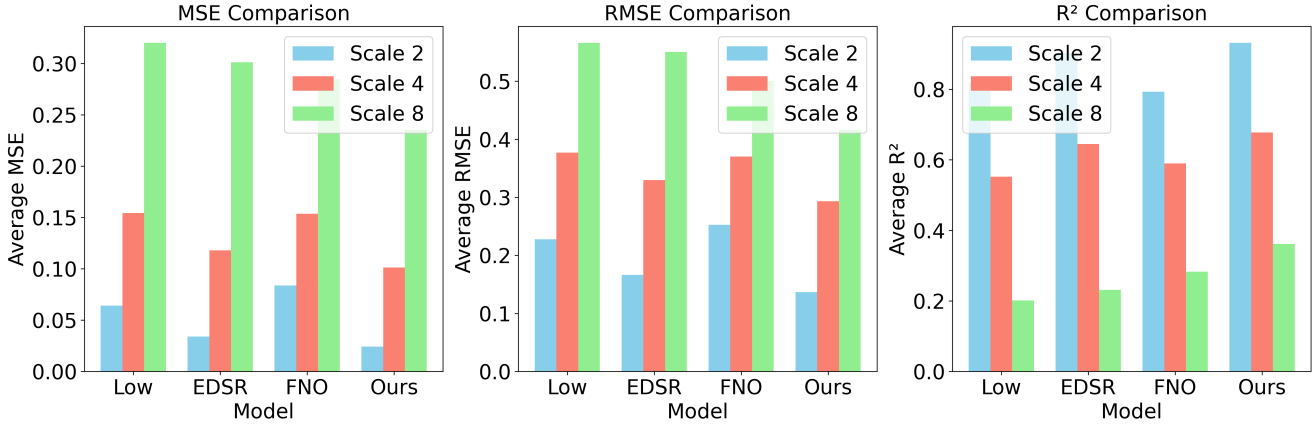

Figure 12. Performance across scale factors. Mean MSE (*left*), RMSE (*middle*), and R^2 (*right*) averaged over the eight diagnostics for Interpolation, EDSR, FNO, and *Ours* at $\times 2$, $\times 4$, and $\times 8$. Our method achieves the lowest errors and highest R^2 at all scales, with the performance gap widening as the scale factor increases.

The key findings from this experiment are as follows: First, as the downscaling factor increases from $\times 2$ to $\times 8$, MSE and RMSE increase, and R^2 decreases for all models, reflecting the growing information gap introduced by coarser inputs. This demonstrates the difficulty in maintaining mapping accuracy as the resolution decreases.

Second, at all scales, *ours* consistently achieves the best performance in terms of MSE/RMSE and the highest R^2, indicating that DEF and HFE provide reliable gains across different resolution levels. This consistent superior performance of our method demonstrates the robustness of the proposed model.

Finally, as the downscaling factor increases, the performance gap between our method and the baselines widens, indicating that our method is more robust to lower resolution inputs. The error bars for the baseline methods grow more steeply with higher downscaling factors, while *ours* exhibits a more gradual degradation, resulting in the largest margins in both error and R^2. This indicates that *ours* is more scalable, especially when high-wavenumber content is severely diminished at higher downscaling factors.

The observed robustness can be attributed to three key design choices in our model. First, the HFE (Sobel, Laplacian, Haar wavelets, FDE) restores important features such as fronts, eddy rims, and filaments that are suppressed in low-resolution inputs, thereby reducing spectral bias as the downscaling factor increases. Second, the *DEF and physical consistency loss* align derivative-level structures (e.g., strain, vorticity, tensor relations), stabilizing the mapping of derivative-dominated targets across scales. Third, the *targeted HF heads* for (S_u, S_v) mitigate the loss of fine-scale gradients, which would otherwise dominate the error at large downscaling factors.

In conclusion, across the range from $\times 2$ to $\times 8$, our method consistently outperforms the baselines and exhibits the slowest performance decay, confirming its cross-scale generalization and making it highly practical for mapping low-resolution velocities to near-eddy-resolving multi-physics diagnostics in resource-constrained environments.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

5.4 Experiment 4: Ablation Study

The purpose of this experiment is to analyze the impact of different architectural components on MSR's performance. By conducting an ablation study using a baseline model—without any enhancement or attention modules—we can isolate and evaluate the specific contributions of components like DEF, Sobel enhancement, etc., and understand which parts of the module lead to performance improvements. The baseline model contains only basic convolutional layers without any additional enhancement or attention modules. On this basis, we successively incorporate or remove individual modules to examine their influence on model performance. The modules considered include DEF, Sobel Enhancement, Laplacian Enhancement, Wavelet Enhancement, FDE, and the Coarse Branch. We report results in terms of R^2 across the eight diagnostic variables $(\zeta, D, \tilde{D}, T_{11}, T_{22}, T_{12}, S_u, S_v)$.

To complement the analysis, we additionally report per-variable normalized RMSE across ablation configurations (Fig. 13). Each panel corresponds to one diagnostic, and bars within each panel are sorted in ascending order (lower is better). The Full Model ranks best or near-best across all variables. The largest reductions appear for deformation/forcing-related quantities—D, T_{12}, T_{22}, S_u, and S_v—indicating that DEF and HFE effectively mitigate spectral bias and better preserve derivative-level structures that are typically suppressed in low-resolution inputs. These RMSE results are consistent with the R^2 findings and provide a complementary error-magnitude perspective.

Overall, the ablation results demonstrate that the combination of multiple complementary modules is crucial for achieving the best performance. While the baseline model alone already provides reasonable reconstructions, integrating targeted modules markedly boosts accuracy, and the full model achieves the most balanced and consistent improvements across all diagnostics.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

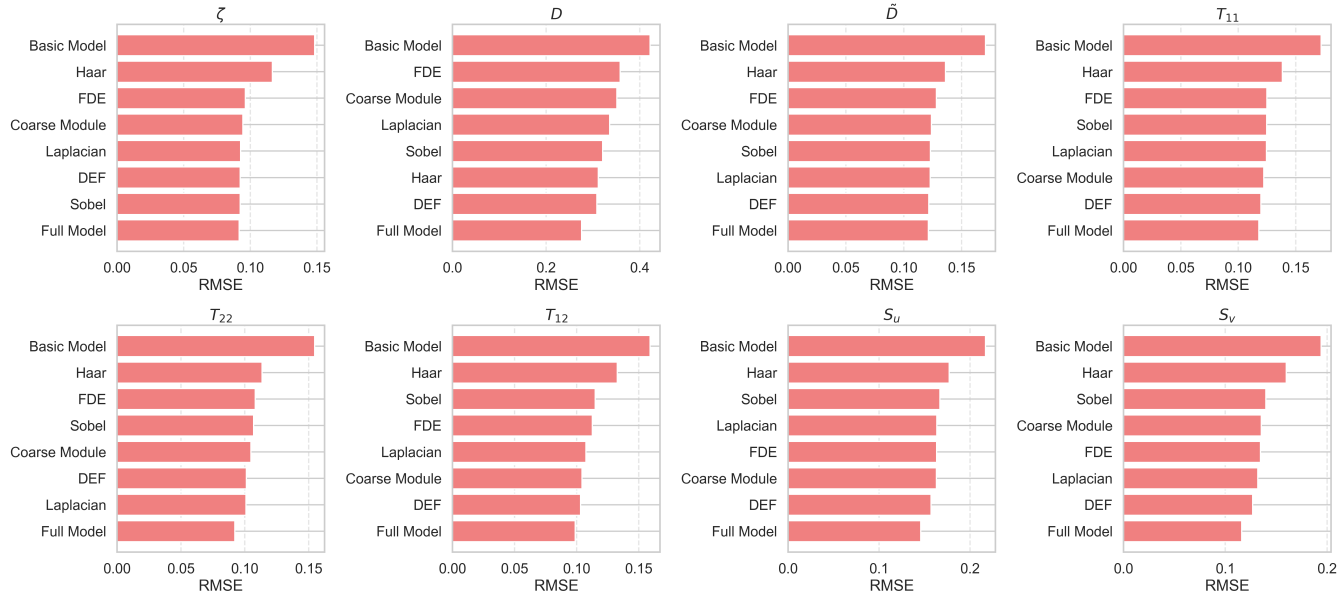

Figure 13. Per-variable normalized RMSE (lower is better) across ablation configurations. Each panel corresponds to one diagnostic—ζ, D, $\tilde{D}$, T_{11}, T_{22}, T_{12}, S_u, S_v—with bars sorted in ascending order within each panel. Configurations: Basic Model (baseline), DEF, Sobel, Laplacian, Haar (wavelet), FDE, Coarse Module, Full Model. The Full Model is consistently best or near-best; the largest RMSE gains are observed for D, T_{12}, T_{22}, S_u, and S_v.

5.5 Experiment 5: Loss Component Ablation

The purpose of this experiment is to evaluate the contribution of each loss component: pixel loss (Pix), spectral amplitude loss (Spec), gradient loss (Grad), and physical consistency loss (Phy) to the performance of MSR. By conducting a loss component ablation, we can identify which terms are most crucial for achieving accurate high-resolution diagnostics in ocean modeling. All splits, inputs, masks, and training schedules match previous experiments. We compare five variants: **Full** (pixel loss + spectral amplitude loss + gradient loss + physical consistency loss), **no_Grad** (remove gradient loss terms), **no_Spec** (remove Spectral amplitude loss terms), **no_Phy** (remove physical consistency loss), and **No_Pix** (remove pixel loss).

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

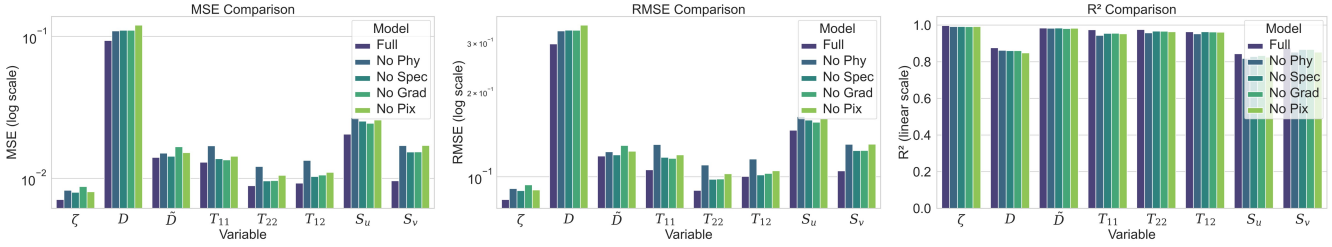

Figure 14. Loss ablation across eight diagnostics. Bars compare `Full`, `No_Grad`, `No_Spec`, and `No_Phy`, `No_Pix`. The full composite loss consistently attains the best trade-off: lowest MSE/RMSE and highest R^2, with the largest margin on high-wavenumber targets (D, S_u, S_v).

The results are summarized in Figure 14, which presents the MSE, RMSE, and R^2 scores for each of the variants. The **Full** model consistently achieves the best balance between error reduction and model performance, showing the lowest MSE/RMSE and the highest R^2, with the most significant improvements seen on high-wavenumber targets like D, S_u, and S_v.

Three key findings emerge from this ablation study: First, the **Full** loss function is consistently superior, achieving uniformly lower errors and higher R^2 for all target variables. This suggests that the pixel loss, spectral amplitude loss, gradient loss and physical consistency loss in the loss function are complementary, rather than redundant. Second, the Phy terms are crucial for small-scale feature recovery. When Phy constraints are removed (**No_Phy**), performance degrades significantly for variables dominated by sharp gradients and filaments, such as ζ, D, S_u, and S_v. This demonstrates that explicit high-frequency guidance helps mitigate spectral bias. Third, the inclusion of the physics-relation term improves mechanical coherence in the model. Dropping this term (**No_Phy**) consistently reduces performance, particularly for the stress components (T_{11}, T_{22}, T_{12}) and deformation (D), indicating that aligning cross-variable derivative relations helps stabilize the model and improves generalization.

5.6 Experiment 6: Hyperparameter Sensitivity

The purpose of this experiment is to quantify how MSR responds to frequency-domain controls and loss weights in order to recover band-limited small-scale dynamics in the presence of spectral attenuation. Because low-resolution inputs suppress shear- and divergence-related (high-wavenumber) content, we hypothesize that explicitly strengthening the high-pass pathway via a cutoff–gain pair (τ, g) reduces the high-frequency spectral mismatch and, in turn, improves RMSE and R^2 for derivative-dominated diagnostics.

To test this, we perform a coarse-to-fine grid search over (τ, g) for the branch-specific Fourier heads of S_u and S_v, and tune a *global* HFE mask for D; in parallel, we vary the loss weights λ_{spec} (spectral amplitude), λ_{grad} (spatial gradients), and λ_{phy} (physical consistency) by $\pm 50\%$ around the baseline. Each configuration is trained with three random seeds under the same split and schedule as Sec. 4.6 and evaluated on held-out days across all levels.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

We report per-sample (date×level) RMSE and R^2 for

$$\{\zeta, D, \tilde{D}, T_{11}, T_{22}, T_{12}, S_u, S_v\},$$

together with a high-frequency spectral-amplitude error

$$E_{\mathrm{HF}} = \left\| M_\tau \odot \left(|\mathcal{F}(\widehat{y})| - |\mathcal{F}(y)| \right) \right\|_2^2, \qquad M_\tau(u,v) = \mathbf{1}\{ r(u,v) \geq \tau \}, \tag{35}$$

where $\mathcal{F}$ denotes the 2-D discrete Fourier transform (DFT), $r(u,v)$ is the normalized radial spatial frequency, and $\odot$ is element-wise multiplication. Influence is summarized by a normalized finite-difference sensitivity

$$S_\theta = \frac{\left| \mathrm{RMSE}(\theta + \delta) - \mathrm{RMSE}(\theta - \delta) \right|}{2\,\delta\,\mathrm{RMSE}_{\mathrm{base}}}, \tag{36}$$

with δ equal to the local sweep step.

The sweeps show that increasing g within a moderate τ window consistently decreases E_{HF} and yields concomitant gains in RMSE/R^2 for S_u and S_v; applying a comparable mask globally in HFE benefits D. Among loss weights, λ_{grad} is most influential for S_u/S_v, whereas λ_{spec} has the largest impact on D. The selected settings and test-set improvements are summarized in Table 3; they provide stable defaults that deliver modest but consistent gains across seeds without degrading large-scale agreement.

Table 3. Best hyperparameter settings and corresponding performance gains in Experiment 6. For each target diagnostic, the selected cut-off–gain pair (τ, g) and the most influential loss weight are reported, together with improvements on the test set. For D, (τ, g) refers to the global HFE mask; for S_u and S_v they refer to the branch-specific Fourier heads.

Variable	Best (τ, g)/ key λ	ΔRMSE	95% CI	ΔR^2
S_u	$(0.15, 2.5)$ / $\lambda_{\mathrm{grad}} = 0.10$	-0.007	$[-0.008, -0.006]$	$+0.01$
S_v	$(0.12, 2.6)$ / $\lambda_{\mathrm{grad}} = 0.10$	-0.006	$[-0.007, -0.005]$	$+0.01$
$D^\dagger$	$(0.10, 2.5)$ / $\lambda_{\mathrm{spec}} = 0.50$	-0.003	$[-0.004, -0.002]$	$+0.02$

† For D, (τ, g) are applied in the global HFE module (not a diagnostic-specific branch).

5.7 Experiment 7: Statistical Significance

The purpose of this experiment is to verify that the observed improvements are not due to sampling noise, we perform paired hypothesis tests on the same per-sample RMSEs (date×level), comparing MSR against strong baselines and against the untuned MSR configuration. For each diagnostic, we compute the paired difference $\Delta\mathrm{RMSE} = \mathrm{RMSE\ (MSR)} - \mathrm{RMSE\ (HighResUnet)}$ and report a paired t-test with Cohen's d as effect size; Wilcoxon signed-rank is provided as a non-parametric corroboration. Because multiple variables and comparisons are assessed, we adjust p-values using Holm–Bonferroni (we also report BH–FDR, which yields consistent conclusions), and we quantify uncertainty with a bootstrap 95% confidence interval for ΔRMSE (2,000 resamples). The statistical significance results are summarized in the table below, where we observe statistically significant reductions in RMSE for S_u, S_v, and D, with corresponding effect sizes and confidence intervals.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

Across the eight diagnostics, MSR achieves statistically significant reductions on each variable versus the primary baseline after Holm correction (Table 4). For example, S_u improves by Δ = -0.0031 (95% CI [-0.0034, -0.0028]), $t = 19.97$, $p <$0.001, $d =$-1.26, S_v by $\Delta =$-0.0014, and D by $\Delta =$-0.0028. Effect sizes lie in the moderate-to-large regime, indicating practical as well as statistical relevance.

Table 4. Paired t-test results in Experiment 7. Paired RMSE difference (ΔRMSE = RMSE(MSR) − RMSE(HighResUnet)) for all eight diagnostics based on $N = 252$ paired samples per variable. Reported are the mean ΔRMSE, 95% confidence interval, effect size (Cohen's d), and Holm–Bonferroni corrected p-values.

Variable	ΔRMSE mean	95% CI	Cohen's d	Holm p
S_u	-0.0031	$[-0.0034, -0.0028]$	-1.26	< 0.001
S_v	-0.0014	$[-0.0017, -0.0011]$	-0.63	< 0.001
T_{22}	-0.0033	$[-0.0041, -0.0026]$	-0.54	< 0.001
D	-0.0028	$[-0.0035, -0.0021]$	-0.52	< 0.001
T_{11}	-0.0028	$[-0.0035, -0.0021]$	-0.52	< 0.001
$\tilde{D}$	-0.0025	$[-0.0032, -0.0018]$	-0.49	< 0.01
T_{12}	-0.0022	$[-0.0028, -0.0016]$	-0.45	< 0.01
ζ	-0.0018	$[-0.0024, -0.0012]$	-0.41	< 0.01

6 Conclusions

This work introduces a *multiphysics super-resolution* (MSR) framework that reconstructs eight dynamically coupled diagnostics ($\zeta, D, \tilde{D}, T_{11}, T_{22}, T_{12}, S_u, S_v$) directly from low-resolution velocity fields. The framework combines (i) Dynamic Enhancement Feature, (ii) High-Frequency Enhancement, (iii) a Multi-Branch module with a shared encoder and dedicated heads for each diagnostic, and (iv) a Physical Consistency Loss that aligns derivative-level structures across variables. These elements aim to mitigate spectral bias while ensuring physical consistency across the reconstructed fields.

Comprehensive experiments validate the effectiveness and robustness of the approach. In single-region comparisons against strong baselines (U-Net variants, EDSR, and FNO), the proposed model consistently achieves the lowest errors and the highest R^2 for most diagnostics, producing notably sharper fronts, eddy rims, and filaments. Cross-basin tests demonstrate consistent performance across regions with modest inter-region variance, indicating excellent portability without region-specific tuning. Sensitivity to downscaling factors demonstrates graceful degradation from $\times 2$ to $\times 8$, with the performance gap widening at larger downscaling factors, underscoring the advantage of explicit high-frequency control when information loss is most severe. Ablation studies isolate the contributions of key components: removing the high-frequency pathway primarily impacts ζ, D, S_u, and S_v, while omitting the physics-relation term degrades the stress components and deformation. These findings demonstrate that spectral recovery and cross-variable alignment are complementary and essential.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

The practical implication is that this data-driven surrogate can generate near-eddy-resolving multi-physics diagnostics from inexpensive low-resolution inputs. Such a capability can accelerate climate diagnostics, facilitate closure of energy and momentum budgets, and provide high-fidelity targets for parameterization development and data assimilation in operational settings where high-resolution simulations are computationally prohibitive.

Limitations and Future Work: Our models were trained on LICOM3 outputs and thus inherit biases from the simulation; evaluating transferability to other models and to observation-constrained products remains important. Currently, we enforce physics via soft constraints; integrating hard conservation (e.g., mass/energy) through differentiable partial differential equation (PDE) solvers or constrained decoding is a natural extension. Temporal consistency and true 3-D volumetric coupling were not the focus of this work, but future neural-operator variants that learn spatio-temporal mappings could further stabilize derivative-dominated quantities. Additionally, uncertainty quantification (using ensembles or Bayesian objectives) and domain adaptation across regimes and resolutions would further enhance the reliability of the framework for decision support.

In summary, the proposed framework delivers cross-scale, cross-region gains for multi-physics ocean diagnostics by combining frequency-aware design with physics-guided learning. We view it as a significant step toward fast, physically consistent surrogates that bridge low-cost low-resolution inputs and the information needed for climate science and operational modeling.

Code and data availability

The LICOM3 eddy-resolving simulation velocity fields (u, v) used in this study are available from the curated dataset described in (Ding et al., 2022) (https://doi.org/10.1038/s41597-022-01766-9). The exact version of the Multiphysics Super-Resolution framework used in this paper, MSR v1.0, is archived on Zenodo (https://doi.org/10.5281/zenodo.18203835). The Zenodo archive contains the source code for MSR v1.0, including the method source code used for the main experiments, the experiment scripts corresponding to Experiments 1–7, and Jupyter notebooks for data processing and result visualization. The archive is organized into directories for method source code and data-processing/result-drawing code. A README file is included in the Zenodo archive to describe the directory structure, software dependencies, basic usage, and the contents of the archive. A LICENSE file is also included. The code is released under the MIT licence.

Author contributions

FZhu, ZH, XW, and XL designed the study and the methodology. FZhu and ZH developed the MSR framework, wrote the code, and carried out the experiments and analyses. PP, WH, FZuo, and XZ assisted with data preparation, diagnostic computation, quality control, and evaluation of the results. YY provided the LICOM3 simulation data and related technical support. XL and XW supervised the project. All authors contributed to interpreting the results and to writing and revising the manuscript.

Competing interests

The authors declare there are no conflicts of interest for this manuscript.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

Acknowledgements. This work was supported by the National Natural Science Foundation of China (Grant No. 42130608) and the Sichuan Science and Technology program (Grant No. 2024YFG0001).

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

References

Abidin, M., Hussin, U., Taufik, A., Ishadi, N., Jalil, M., and Khan, S.: Implementing the Enhanced Deep Super Resolution (EDSR) Model for Single Image Super-Resolution in a Web Application, in: Data-Driven Decision Making for Sustainable Business Growth, pp. 87–98, Springer, https://doi.org/10.1007/978-3-031-96530-2_9, 2025.

Anwar, S., Khan, S., and Barnes, N.: A deep journey into super-resolution: A survey, ACM computing surveys (CSUR), 53, 1–34, https://doi.org/10.1145/3390462, 2020.

Bagaeva, E., Danilov, S., Oliver, M., and Juricke, S.: Advancing eddy parameterizations: Dynamic energy backscatter and the role of subgrid energy advection and stochastic forcing, Journal of Advances in Modeling Earth Systems, 16, e2023MS003 972, https://doi.org/10.1029/2023MS003972, 2024.

Batteen, M. L. and Han, Y.-J.: On the computational noise of finite-difference schemes used in ocean models, Tellus, 33, 387–396, https://doi.org/10.1111/j.2153-3490.1981.tb01761.x, 1981.

Batziou, E., Ioannidis, K., Patras, I., Vrochidis, S., and Kompatsiaris, I.: Low-light image enhancement based on U-Net and Haar wavelet pooling, in: International Conference on Multimedia Modeling, pp. 510–522, Springer, https://doi.org/10.1007/978-3-031-27818-1_42, 2023.

Bolton, T. and Zanna, L.: Applications of deep learning to ocean data inference and subgrid parameterization, Journal of Advances in Modeling Earth Systems, 11, 376–399, https://doi.org/10.1029/2018MS001472, 2019.

Buongiorno Nardelli, B., Cavaliere, D., Charles, E., and Ciani, D.: Super-resolving ocean dynamics from space with computer vision algorithms, Remote Sensing, 14, 1159, https://doi.org/10.3390/rs14051159, 2022.

Buza, M. K. I. E. and Akagic, A.: Super-Resolution in Medical Imaging: Evaluating SRCNN, EDSR, ESPCN, and SRGAN on Dermatological Images, in: Advanced Technologies, Systems, and Applications X: Proceedings of the International Symposium on Innovative and Interdisciplinary Applications of Advanced Technologies (IAT) 2025, p. 202, Springer Nature, https://doi.org/10.1007/978-3-032-05159-2_13, 2025.

Cai, Z., Xin, J., Shi, P., Zhou, S., Wu, J., and Zheng, N.: Meta pixel loss correction for medical image segmentation with noisy labels, in: Workshop on Medical Image Learning with Limited and Noisy Data, pp. 32–41, Springer, https://doi.org/10.1007/978-3-031-16760-7_4, 2022.

Chassignet, E. P. and Xu, X.: On the importance of high-resolution in large-scale ocean models, Advances in Atmospheric Sciences, 38, 1621–1634, https://doi.org/10.1007/s00376-021-0385-7, 2021.

de Wolff, T., Carrillo, H., Martí, L., and Sanchez-Pi, N.: Assessing physics informed neural networks in ocean modelling and climate change applications, in: AI: Modeling Oceans and Climate Change Workshop at ICLR 2021, https://inria.hal.science/hal-03262684, 2021.

Ding, M., Liu, H., Lin, P., Meng, Y., Zheng, W., An, B., Luan, Y., Yu, Y., Yu, Z., Li, Y., et al.: A century-long eddy-resolving simulation of global oceanic large-and mesoscale state, Scientific Data, 9, 691, https://doi.org/10.1038/s41597-022-01766-9, 2022.

Fox-Kemper, B., Adcroft, A., Böning, C. W., Chassignet, E. P., Curchitser, E., Danabasoglu, G., Eden, C., England, M. H., Gerdes, R., Greatbatch, R. J., et al.: Challenges and prospects in ocean circulation models, Frontiers in Marine Science, 6, 65, https://doi.org/10.3389/fmars.2019.00065, 2019.

Fukami, K., Fukagata, K., and Taira, K.: Super-resolution analysis via machine learning: a survey for fluid flows, Theoretical and Computational Fluid Dynamics, 37, 421–444, https://doi.org/10.1007/s00162-023-00663-0, 2023.

Gangopadhyay, A.: Introduction to ocean circulation and modeling, CRC Press, https://doi.org/10.1201/9780429347221, 2022.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

Gao, X., Wu, S., Zhou, Y., Wu, X., Wang, F., and Hu, X.: Lightweight image super-resolution via multi-branch aware CNN and efficient transformer, Neural Computing and Applications, 36, 5285–5303, https://doi.org/10.1007/s00521-023-09353-8, 2024.

Guo, W., Xu, G., Liu, B., and Wang, Y.: Hyperspectral image classification using CNN-enhanced multi-level Haar wavelet features fusion network, IEEE Geoscience and Remote Sensing Letters, 19, 1–5, https://doi.org/10.1109/LGRS.2022.3167535, 2022.

Hong, Y., Wang, X., Wang, B., Li, W., and Han, G.: A Fusion Method Based on Physical Modes and Satellite Remote Sensing for 3D Ocean State Reconstruction, Remote Sensing, 17, 1468, https://doi.org/10.3390/rs17081468, 2025.

Hu, J., Shen, L., and Sun, G.: Squeeze-and-excitation networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7132–7141, https://doi.org/10.48550/arXiv.1709.01507, 2018.

Huang, J., Guan, D., Xiao, A., and Lu, S.: Fsdr: Frequency space domain randomization for domain generalization, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 6891–6902, https://doi.org/10.48550/arXiv.2103.02370, 2021.

Ibrahim, M. R., Benavente, R., Lumbreras, F., and Ponsa, D.: 3DRRDB: Super resolution of multiple remote sensing images using 3D residual in residual dense blocks, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 323–332, 2022.

Khani, S. and Dawson, C. N.: A gradient based subgrid-scale parameterization for ocean mesoscale eddies, Journal of Advances in Modeling Earth Systems, 15, e2022MS003 356, https://doi.org/10.1029/2022MS003356, 2023.

Kitsios, V., Frederiksen, J., and O'Kane, T.: Subgrid parameterization of eddy, meanfield and topographic interactions in simulations of an idealized Antarctic Circumpolar Current, Journal of Advances in Modeling Earth Systems, 15, e2022MS003 412, https://doi.org/10.1029/2022MS003412, 2023.

Lawal, Z. K., Yassin, H., Lai, D. T. C., and Che Idris, A.: Physics-informed neural network (PINN) evolution and beyond: A systematic literature review and bibliometric analysis, Big Data and Cognitive Computing, 6, 140, https://doi.org/10.3390/bdcc6040140, 2022.

Li, Y., Sixou, B., and Peyrin, F.: A review of the deep learning methods for medical images super resolution problems, Irbm, 42, 120–133, https://doi.org/10.1016/j.irbm.2020.08.004, 2021.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., and Anandkumar, A.: Fourier neural operator for parametric partial differential equations, arXiv preprint arXiv:2010.08895, https://doi.org/10.48550/arXiv.2010.08895, 2020.

Liu, G., Li, Q., Wang, Y., Zhou, X., Li, Y., Liu, Y., Li, X., and Liu, Z.: Resolution enhancement and target segmentation of medical images based on the frequency-domain information in deep learning, Applied Optics, 64, 7083–7092, https://doi.org/10.1364/AO.557903, 2025.

Liu, R., Tao, F., Liu, X., Na, J., Leng, H., Wu, J., and Zhou, T.: RAANet: A residual ASPP with attention framework for semantic segmentation of high-resolution remote sensing images, Remote Sensing, 14, 3109, https://doi.org/10.3390/rs14133109, 2022.

Liu, Z. and Liao, G.: Relationship between global ocean mixing and coherent mesoscale eddies, Deep Sea Research Part I: Oceanographic Research Papers, 197, 104 067, https://doi.org/doi.org/10.1016/j.dsr.2023.104067, 2023.

Lu, Z., Li, J., Liu, H., Huang, C., Zhang, L., and Zeng, T.: Transformer for single image super-resolution, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 457–466, https://doi.org/10.48550/arXiv.2108.11084, 2022.

Mak, J., Marshall, D. P., Madec, G., and Maddison, J. R.: Acute sensitivity of global ocean circulation and heat content to eddy energy dissipation timescale, Geophysical Research Letters, 49, e2021GL097 259, https://doi.org/10.1029/2021GL097259, 2022.

Meng, L., Yan, C., Zhuang, W., Zhang, W., Geng, X., and Yan, X.-H.: Reconstructing high-resolution ocean subsurface and interior temperature and salinity anomalies from satellite observations, IEEE Transactions on Geoscience and Remote Sensing, 60, 1–14, https://doi.org/10.1109/TGRS.2021.3109979, 2021.

Pacanowski, R. C. and Gnanadesikan, A.: Transient response in a z-level ocean model that resolves topography with partial cells, Monthly Weather Review, 126, 3248–3270, https://doi.org/10.1175/1520-0493(1998)126<3248:TRIAZL>2.0.CO;2, 1998.

https://doi.org/10.5194/egusphere-2026-691
Preprint. Discussion started: 6 May 2026

Perezhogin, P., Zanna, L., and Fernandez-Granda, C.: Generative data-driven approaches for stochastic subgrid parameterizations in an idealized ocean model, Journal of Advances in Modeling Earth Systems, 15, e2023MS003 681, https://doi.org/10.1029/2023MS003681, 2023.

Perezhogin, P., Zhang, C., Adcroft, A., Fernandez-Granda, C., and Zanna, L.: A stable implementation of a data-driven scale-aware mesoscale parameterization, Journal of Advances in Modeling Earth Systems, 16, e2023MS004 104, https://doi.org/10.1029/2023MS004104, 2024.

Perezhogin, P., Balakrishna, A., and Agrawal, R.: Large eddy simulation of ocean mesoscale eddies, arXiv preprint arXiv:2501.05357, https://doi.org/10.48550/arXiv.2501.05357, 2025.

Ren, X., Lai, S., et al.: Medical image enhancement based on laplace transform, sobel operator and histogram equalization, Academic Journal of Computing & Information Science, 5, 48–54, https://doi.org/10.25236/AJCIS.2022.050608, 2022.

Sun, H., Yang, Q., Li, J., Zhao, W., and Tian, J.: Parameterization of shear-to-strain ratio used in finescale parameterization, Journal of Geophysical Research: Oceans, 129, e2023JC020 393, https://doi.org/10.1029/2023JC020393, 2024.

Sun, L., Pan, J., and Tang, J.: Shufflemixer: An efficient convnet for image super-resolution, Advances in Neural Information Processing Systems, 35, 17 314–17 326, https://proceedings.neurips.cc/paper_files/paper/2022/file/6e60a9023d2c63f7f0856910129ae753-Paper-Conference.pdf, 2022.

Tian, H., Zhang, L., Li, S., Yao, M., and Pan, G.: Multi-depth branch network for efficient image super-resolution, Image and Vision Computing, 144, 104 949, https://doi.org/10.1016/j.imavis.2024.104949, 2024.

Wu, T., Wang, L., and Zhu, J.: Image Edge Detection Based on Sobel with Morphology, in: 2021 IEEE 5th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), vol. 5, pp. 1216–1220, IEEE, https://doi.org/10.1109/ITNEC52019.2021.9586895, 2021.

Yoon, S., Park, Y., Gerstoft, P., and Seong, W.: Predicting ocean pressure field with a physics-informed neural network, The Journal of the Acoustical Society of America, 155, 2037–2049, https://doi.org/10.1121/10.0025235, 2024.

Zhang, C., Perezhogin, P., Gultekin, C., Adcroft, A., Fernandez-Granda, C., and Zanna, L.: Implementation and evaluation of a machine learned mesoscale eddy parameterization into a numerical ocean circulation model, Journal of Advances in Modeling Earth Systems, 15, e2023MS003 697, https://doi.org/10.1029/2023MS003697, 2023.

Zhu, Y., Zhang, R.-H., Moum, J. N., Wang, F., Li, X., and Li, D.: Physics-informed deep-learning parameterization of ocean vertical mixing improves climate simulations, National Science Review, 9, nwac044, https://doi.org/10.1093/nsr/nwac044, 2022.