

PeSNO: A Physics-embedded Stream Neural Operator for the Acceleration of Mantle Convection Simulations

Keywords: Mantle convection, Physics-embedded neural operator, Stream function formulation, Numerical acceleration

1. Research Overview

This study proposes the Physics-embedded Stream Neural Operator (PeSNO). Unlike traditional surrogate models that aim to fully replace numerical solvers, PeSNO functions as a high-fidelity warm-start mechanism integrated directly into the iterative solver loop.

The architecture features a hybrid spectral operator based on the Discrete Sine Transform (DST) and odd-symmetric extension. It is augmented by anti-symmetric reflection padding and a physical Hard Mask. This construction enforces zero Dirichlet boundary constraints at machine precision. It effectively eliminates boundary-induced Gibbs artifacts. **Additionally, the odd-symmetric design implicitly enforces a free-slip boundary condition. This permits lateral flow along the walls, satisfying the stress-free constraint of mantle convection.**¹

For optimization, we introduce a weighted Sobolev loss function. This explicitly constrains differential properties of the flow field in H^1 (velocity) and H^2 (vorticity) spaces. It mitigates the over-smoothing of high-frequency shear features in strongly variable viscosity fields.

We deploy PeSNO within a projection-based solver pipeline. It provides physics-consistent initial values. This reduces linear solver iterations by over 50%. Compared to State-of-the-Art models like NewFluidNet, PeSNO achieves a 5.06x end-to-end sequential acceleration. It strictly maintains the solver's 10^{-6} residual convergence criterion. These results demonstrate deep learning as a robust acceleration operator for high-precision numerical computation.

2. Operator Learning Task Formulation

This study reformulates the instantaneous momentum balance of mantle convection as an operator approximation task, mapping from a high-dimensional state space to the flow stream function space. We define a nonlinear operator G_θ , parameterized by neural network weights θ , to approximate the solution operator of the Stokes equations:

$$G_\theta : (T, \log \eta, \nabla T, \mathbf{x}, \mathbf{c}) \mapsto \psi \quad (1)$$

¹[Zehua 1]: Response to "Also the Neumann BCs?": Clarified that the odd-symmetric design implies a free-slip condition, allowing lateral flow rather than a no-slip wall.

The input space comprises the thermodynamic state (temperature T), rheological states (log-viscosity $\log \eta$ and **the temperature gradient's gradient**² ∇T), the spatial coordinate vector $\mathbf{x}$, and the global physical control parameter vector $\mathbf{c}$. This mapping yields a scalar stream function field ψ that implicitly satisfies physical boundary constraints. Rather than predicting observable velocities directly, the scalar output ψ is transformed into physical quantities via a deterministic differential operator D :

$$\mathbf{u} = D_u(\psi) = (\partial_y \psi, -\partial_x \psi)^\top, \quad \omega = D_\omega(\psi) = -\nabla^2 \psi \quad (2)$$

Here, $\mathbf{u}$ denotes the velocity vector field and ω the vorticity scalar field. This formulation guarantees that the velocity satisfies the incompressibility constraint $\nabla \cdot \mathbf{u} = 0$ by construction. The model yields divergence-free initial values for the numerical solver.

3. Hybrid-Augmented Data Construction

Accurate learning of fluid differential structures demands high-quality, physically complete training data. An ideal benchmark dataset must be temporally continuous to support high-order derivative learning and strictly mass-conservative to enforce hard physical constraints. However, existing public datasets primarily employ sparse sampling (approximately one snapshot every 100 steps), **which lacks the temporal resolution required for inertial blending (Eq. 3).failing to capture high-frequency evolutionary trajectories**³. Furthermore, raw outputs from baseline models often exhibit mass leakage at rigid boundaries. To address these limitations, we developed a hybrid data generation pipeline. This approach couples a pre-trained approximation model with numerical solvers to perform mass correction.

The process begins by obtaining a candidate momentum solution. At each time step t , we employ the NewFluidNet architecture (the Teacher model) [1] to predict an initial velocity field $\mathbf{u}_{\text{pred}}$ from the current temperature T_t . To suppress non-physical

²[Zehua 2]: Correction: "its" incorrectly referred to viscosity. Clarified that ∇T is the temperature gradient.

³[Zehua 3]: Response: Dense sampling is strictly required to execute the inertial blending strategy (Eq. 3). Clarified that sparse data lacks the temporal resolution for this operation.

frame-to-frame jitter and maintain temporal coherence, we introduce an inertial blending strategy:

$$\mathbf{u}^* = \alpha \mathbf{u}_{t-1} + (1 - \alpha) \mathbf{u}_{\text{pred}} \quad (3)$$

Here, the inertia coefficient α is set to 0.6. While the intermediate field $\mathbf{u}^*$ retains convective features, it temporarily violates mass conservation. This divergence acts as the source term for pressure projection. The projection restores the divergence-free constraint for the ground truth. ~~This violates the divergence-free condition and boundary constraints. This necessitates a rigorous divergence-free projection⁴.~~

We interpolate $\mathbf{u}^*$ from the collocated grid onto a staggered MAC grid. ~~We project $\mathbf{u}^*$ onto a MAC grid⁵ [2], strictly enforcing zero normal velocity at vertical boundaries. The correction potential ϕ (representing the dimensionless dynamic pressure) is resolved by solving the Poisson equation via the Conjugate Gradient method (tolerance 10^{-8}) [3]:~~

$$\nabla^2 \phi = \nabla \cdot \mathbf{u}^* \quad (4)$$

Applying the Helmholtz-Hodge decomposition $\mathbf{u}_t = \mathbf{u}^* - \nabla \phi$ [4] yields a corrected velocity field $\mathbf{u}_t$ that is divergence-free at machine precision.

Since PeSNO predicts a scalar stream function, we derive the stream potential ψ from the corrected velocity $\mathbf{u}_t$. The stream function and vorticity $\omega = \nabla \times \mathbf{u}_t$ satisfy the Poisson relationship:

$$\nabla^2 \psi = -\omega \quad (5)$$

We solve this equation using double-precision Conjugate Gradient with boundary conditions matching the PeSNO output space ($\psi = 0$ at rigid boundaries). ~~This solve regularizes the reference data. It eliminates spurious vorticity from mismatched difference schemes. vertical boundaries). This ensures the training labels are mathematically consistent with the physical velocity field, eliminating spurious vorticity introduced by mismatched difference schemes⁶.~~

Finally, temperature evolution is advanced by solving the energy equation. We employ a Finite Volume Method with a Superbee flux limiter [5] for advection and an adaptive-step fourth-order Runge-Kutta (RK4) scheme [6] for time integration. This pipeline generated 130 independent cases (1000 saved snapshots each; adaptive sub-stepping ensures fully developed regimes 1000 steps each⁷), providing complete supervisory signals including velocity, temperature, pressure, and stream functions to facilitate the learning of implicit fluid dynamics.

⁴[Zehua 4]: Response to "Mass conservation?": Clarified that blending creates a divergence source term. The subsequent projection removes this to ensure the ground truth is mass-conservative.

⁵[Zehua 5]: Response to "GAIA/NewFluidNet are on collocated grids": Acknowledged. Explicitly stated that we interpolate the field from the collocated grid to the MAC grid to perform the projection.

⁶[Zehua 6]: Response to "Not true for NewFluidNet": Corrected "vertical" to "rigid". Clarified that this step regularizes teacher data to eliminate artifacts from mismatched schemes.

⁷[Zehua 7]: Response to "Steady-state/Step count": Clarified distinction between saved snapshots and raw CFL steps. Confirmed that adaptive sub-stepping ensures physical development.

4. Physics-Embedded Stream Neural Operator

Figure 1 illustrates the overall architecture of the proposed PeSNO framework, detailing the designs of its core spectral blocks and parameter modulation mechanisms.

4.1. Physics-Augmented Input Space & Multi-Scale Modulation

Mantle convection involves multi-physics coupling. A single temperature input fails to describe the full dynamic state. We construct a composite input space to address large parameter variations. This space incorporates thermodynamic states, rheological constraints, and spatial information.

The input tensor includes the log-viscosity field $\log(\eta)$ and the temperature gradient ∇T . The logarithmic transformation linearizes viscosity values across orders of magnitude. This facilitates the identification of the high-viscosity stagnant lid [7]. Additionally, the temperature gradient field provides derivative information on thermal fronts and slab edges. This improves model sensitivity to local rheological variations.

~~Convolutional architectures effectively extract local features. However, raw coordinate inputs lack high-frequency content. Relying on them limits the resolution of fine-scale spatial dependencies [8]. Deep convolutional networks suffer from spectral bias, causing conventional coordinate inputs to fail in resolving narrow convective channels [9].⁸~~ To mitigate this, we introduce Fourier feature mapping with positional encoding at the input stage. This projects 2D spatial coordinates $\mathbf{x} = (x, y)$ onto a basis of high-frequency sine and cosine functions [8]:

$$\gamma(\mathbf{x}) = [\sin(2^0 \pi \mathbf{x}), \cos(2^0 \pi \mathbf{x}), \dots, \sin(2^{N-1} \pi \mathbf{x}), \cos(2^{N-1} \pi \mathbf{x})] \quad (6)$$

This mapping $\gamma : \mathbb{R}^2 \rightarrow \mathbb{R}^{2N}$ expands low-dimensional positional data into high-dimensional multi-scale feature vectors, forcing the network to process information at higher spatial frequencies and thereby improving the reconstruction of fine flow structures. Finally, to ensure continuous generalization across dynamic regimes, the architecture integrates the Feature-wise Linear Modulation (FiLM) mechanism [10]. We assemble global physical control parameters into a vector $\mathbf{c}$:

$$\mathbf{c} = [H, \eta_T, \eta_z]^\top \quad (7)$$

Here, H denotes the dimensionless internal heating rate. Parameters η_T and η_z represent the temperature and depth sensitivity coefficients of viscosity. Intermediate feature maps F are dynamically modulated via affine transformations derived from $\mathbf{c}$:

$$F' = (1 + s(\mathbf{c})) \odot F + b(\mathbf{c}) \quad (8)$$

Here, $s(\mathbf{c})$ and $b(\mathbf{c})$ are learned scale and shift coefficients. The symbol $\odot$ denotes the Hadamard product. This mechanism dynamically rescales the amplitude and baseline of feature responses. It allows a single set of convolutional weights to cover

⁸[Zehua 8]: Response to "Spectral bias reference": Clarified that the limitation stems from the low-frequency nature of raw coordinate inputs rather than the CNN architecture itself.

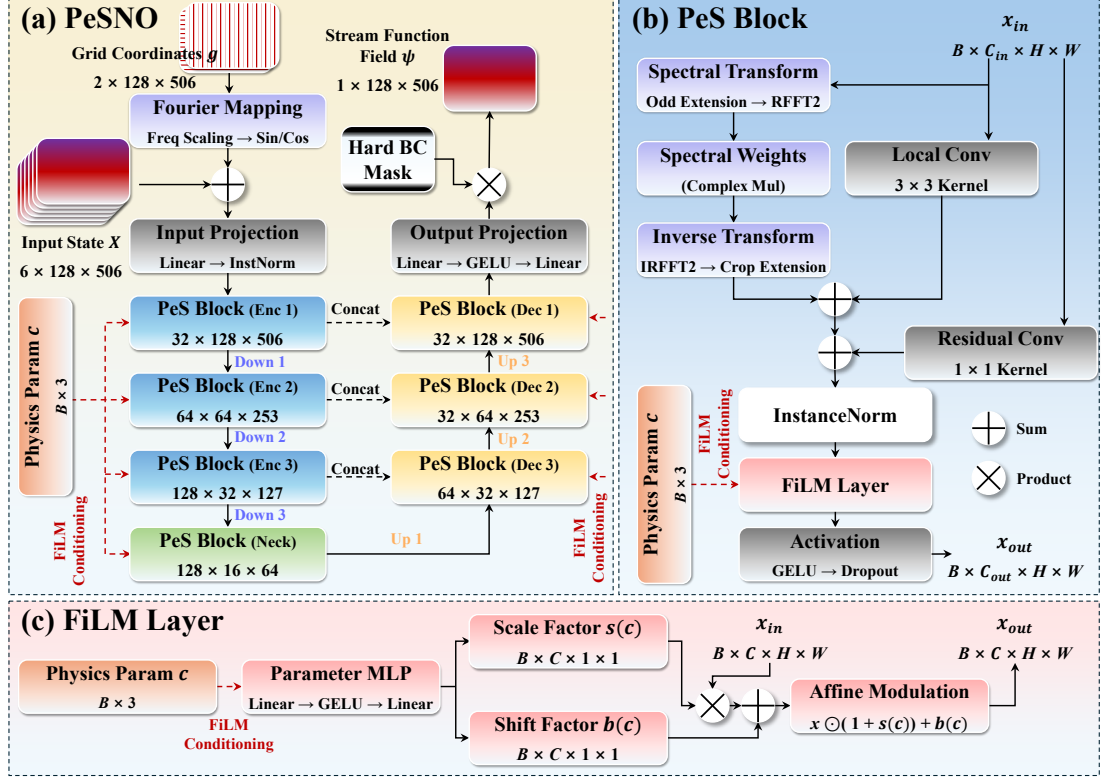


Figure 1: Overall architecture of the proposed PeSNO framework.

a broad parameter regime ranging from weak to high-intensity chaotic convection. Further validation of the model’s generalization across extreme rheological regimes is provided in [Appendix A](#).

4.2. Geometric-Adaptive Hybrid Spectral Stream Operator

The PeSNO architecture addresses the mismatch between global spectral processing and non-periodic boundary constraints. Standard neural operators rely on periodic Fast Fourier Transforms (FFT). This formulation fails for the vertical confinement of the mantle [11]. We employ a hybrid strategy to enforce zero Dirichlet conditions. This integrates odd-symmetric spectral extension with spatial anti-symmetric padding.

The spectral branch performs an odd-symmetric extension of the feature tensor along the vertical dimension. This is equivalent to a Discrete Sine Transform (DST) using sine basis functions $\sin(ky)$ [12]. Sine bases vanish at interval endpoints. Consequently, the reconstructed flow satisfies rigid wall constraints. This eliminates Gibbs phenomena from basis truncation [13].

We use a parallel convolution branch to capture high-frequency local information. Horizontal circular padding maintains periodicity. We apply anti-symmetric reflection padding vertically. This operation negates interior pixels ($x_{-i} = -x_i$) across the boundary. It constructs an odd function symmetry for the stream function. This provides the convolution kernel with a smooth contextual window consistent with the $\psi = 0$ constraint [14]. The computed gradient field remains continuous across the boundary. This suppresses numerical vorticity artifacts caused by truncation.

Finally, we apply a physical Hard Mask to the output. We define a binary mask $M \in \{0, 1\}^{H \times W}$. Rows corresponding to the lithosphere ($y = 1$) and core-mantle boundary ($y = 0$) are set to 0. All other rows are set to 1. The final stream function is

$$\psi = \hat{\psi} \odot M \quad (9)$$

Here, $\hat{\psi}$ denotes the raw network prediction. The symbol $\odot$ denotes the Hadamard product. The Hard Mask does not enforce abrupt truncation [15]. Rather, it filters residual machine epsilon errors [16]. Anti-symmetric padding ensures the flow field approaches **zero**. **This design guarantees kinematic compliance by construction. The derived velocity field strictly satisfies mass conservation and no-penetration boundary conditions. zero**⁹.

4.3. Gradient-Prioritized Weighted Sobolev Loss

Accurate mantle convection simulation demands the faithful reconstruction of the fluid’s dynamic state. While the stream function ψ guarantees a divergence-free field, the governing physical quantities are its spatial derivatives, namely velocity and vorticity. Conventional loss functions that minimize stream function error naively often induce high-frequency noise in the differential structure, leading to severe distortions in derived

⁹[Zehua 9]: Response to "Impressive BCs and conservation": Confirmed. Added a closing sentence to explicitly state that this architecture guarantees exact kinematic compliance and mass conservation by construction.

Internal Heating $Q \approx 6.92$; Viscosity Contrast $\Delta\eta_T \approx 2.58e+06$; $\Delta\eta_P \approx 2.29$

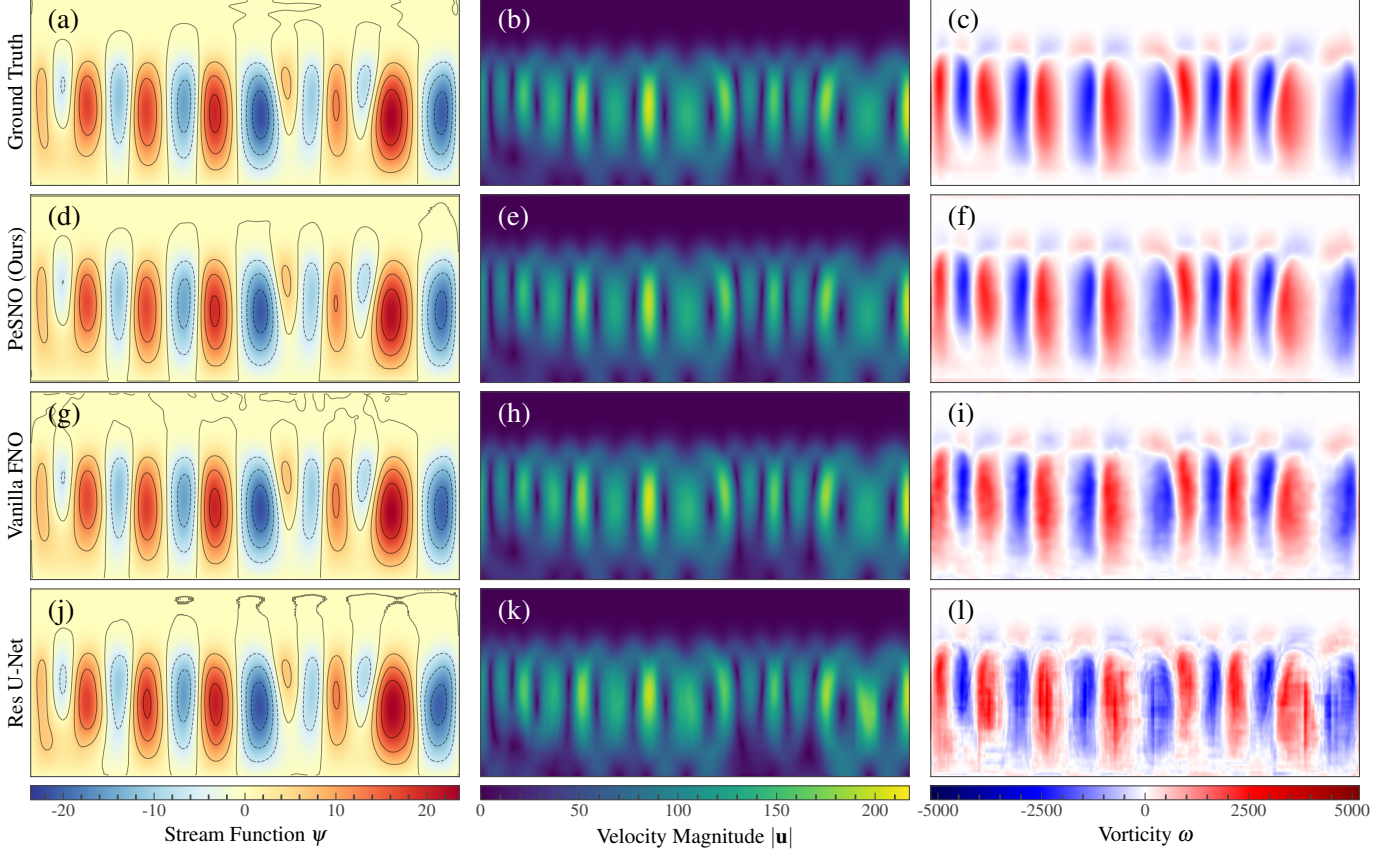


Figure 2: Comparison of instantaneous flow field reconstruction ($Q \approx 6.92$, $\Delta\eta_T \approx 2.58 \times 10^6$). Columns display the stream function ψ (left), velocity magnitude $|\mathbf{u}|$ (center), and vorticity ω (right). Rows correspond to Ground Truth, PeSNO (Ours), Vanilla FNO, and Res U-Net, respectively.

physical quantities [17]. To mitigate this, we formulate a composite loss function based on weighted Sobolev norms, explicitly constraining the multi-order differential properties of the flow field within the H^2 space [18].

Our optimization objective targets the zeroth-order potential, first-order kinetic energy (velocity), and second-order rotation (vorticity). To ensure consistency between the neural network and the numerical solver, we calculate derivatives using the exact staggered grid finite difference scheme employed during data generation [19]. The total loss function $\mathcal{L}$ and associated physical definitions are

$$\begin{cases} \mathcal{L} &= \|\psi - \hat{\psi}\|_{1,W} + \lambda_1(t)\|\mathbf{u} - \hat{\mathbf{u}}\|_{1,W} + \lambda_2(t)\|\omega - \hat{\omega}\|_{1,W} \\ \mathbf{u} &= (\partial_y\psi, -\partial_x\psi)^\top \\ \omega &= -\nabla^2\psi \end{cases} \quad (10)$$

Here, $\psi, \mathbf{u}, \omega$ denote the ground truth stream function, velocity, and vorticity. Variables $\hat{\psi}, \hat{\mathbf{u}}, \hat{\omega}$ denote network predictions. We define the spatially weighted L_1 distance as

$$\|f\|_{1,W} = \frac{1}{N} \sum_{\mathbf{x}} W(\mathbf{x})|f(\mathbf{x})| \quad (11)$$

Here, N represents the total number of grid points. Hyperparameters $\lambda_i(t)$ adjust dynamically during training. Physically, the velocity term constrains mass and momentum fluxes. The

vorticity term enforces local rotational and shear structures. This allows the model to resolve steep gradients at thermal anomaly edges.

To address the spatial heterogeneity caused by high-velocity plumes amidst a stagnant background, we introduce an adaptive spatial weighting mechanism $W(\mathbf{x})$ based on velocity magnitude. This assigns higher penalty weights to high-velocity regions, preventing the extensive low-velocity background from dominating the optimization [20]. Furthermore, given the sensitivity of second-order derivatives to noise, we implement a Curriculum Learning strategy [21]. We initialize λ_1, λ_2 with low values to prioritize large-scale circulation features, then progressively increase them to refine small-scale shear structures [22].

5. Experiments and Analysis

5.1. Experimental Setup

We evaluate PeSNO using the constructed Hybrid Physics-Augmented dataset (130 independent cases), randomly partitioned into training (100), validation (15), and testing (15) sets. We benchmark against two representative architectures: Residual U-Net (Res U-Net) [23, 24, 25], representing local convolutional methods, and the Fourier Neural Operator (FNO) [11],

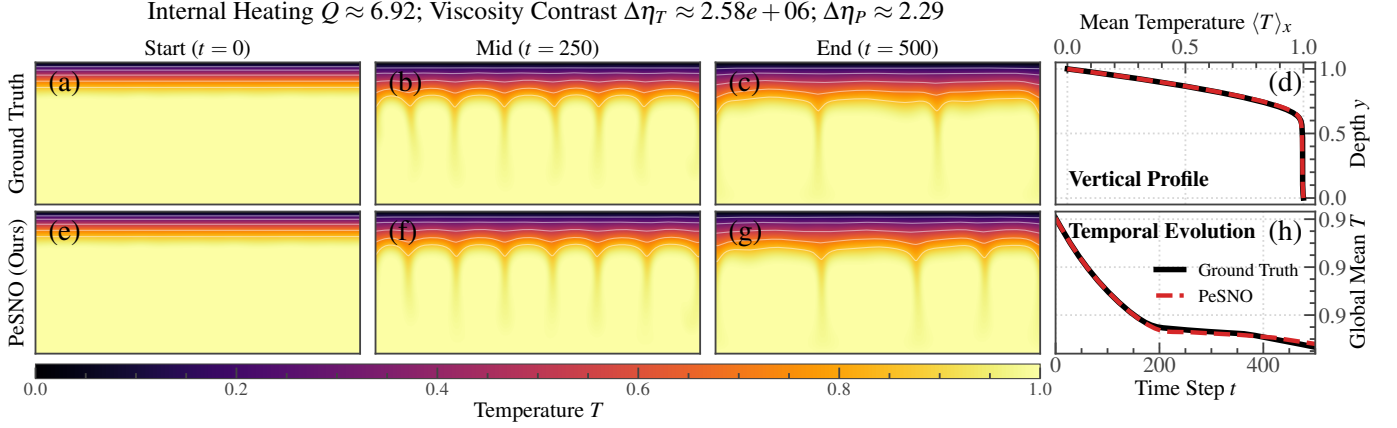


Figure 3: Long-term spatiotemporal evolution and conservation statistics. (a-c, e-g) Temperature snapshots at $t = 0, 250, 500$ (Top: Ground Truth; Bottom: PeSNO). White lines denote isotherms. (d) Horizontally averaged temperature profile $\langle T \rangle_x$ at the final step. (h) Global mean temperature time series. [video]

Table 1: Evaluation of physical conservation over a 500-step auto-regressive evolution. The table presents the final error metrics for the temperature field T , with key results highlighted in bold.

Metric	MAE	RMSE	L_∞ (Max Error)
Profile (Structure)	2.53×10^{-3}	3.80×10^{-3}	1.01×10^{-2}
Time Series (Energy)	5.13×10^{-4}	6.83×10^{-4}	1.49×10^{-3}

representing spectral operator learning. To ensure fair comparison, all baselines are configured to directly predict the stream function, with parameter counts strictly aligned across models.

Beyond numerical approximation, our evaluation framework prioritizes the dynamical fidelity of derived physical quantities. We define four hierarchical error metrics corresponding to the Sobolev spaces of varying orders [26].

First, the relative L_2 norm of the stream function measures zeroth-order structural consistency:

$$E_\psi = \frac{\|\hat{\psi} - \psi\|_2}{\|\psi\|_2} \quad (12)$$

Here, E_ψ quantifies the overall reconstruction accuracy of the large-scale convective circulation geometry.

Second, since velocity governs mass and momentum transport, we evaluate the kinematic accuracy via the relative vector L_2 error:

$$\begin{cases} \mathbf{u} &= (\partial_y \psi, -\partial_x \psi)^\top \\ E_u &= \frac{\|\hat{\mathbf{u}} - \mathbf{u}\|_2}{\|\mathbf{u}\|_2} \end{cases} \quad (13)$$

This metric, mathematically equivalent to the H^1 semi-norm error, assesses the model’s ability to resolve first-order differential properties and momentum fluxes [27, 28].

Third, we introduce vorticity ω as a second-order physical constraint to test the resolution of local fluid structures, particularly shear bands at thermal anomaly edges and slab deformation zones:

$$\begin{cases} \omega &= -\nabla^2 \psi \\ E_\omega &= \frac{\|\hat{\omega} - \omega\|_2}{\|\omega\|_2} \end{cases} \quad (14)$$

As second-order derivatives are highly sensitive to high-frequency noise, E_ω serves as a critical proxy for operator smoothness and microstructural fidelity [29].

Finally, to prevent the averaging effect of L_2 norms from masking local extremes, we employ the relative L_∞ norm:

$$E_\infty = \frac{\|\hat{\mathbf{u}} - \mathbf{u}\|_\infty}{\|\mathbf{u}\|_\infty} \quad (15)$$

Here, $\|\cdot\|_\infty$ denotes the pointwise supremum over the domain. This metric detects over-smoothing by quantifying accuracy in local anomalies, such as high-velocity cores and downwelling tips [30].

5.2. Instantaneous Flow Reconstruction and Structural Fidelity Analysis

We quantitatively evaluate PeSNO generalization using a test set of 15 independent evolution sequences. These span weak to strongly chaotic convection. Table 2 summarizes the results. PeSNO outperforms baseline models across physical metrics including flow morphology, velocity magnitude, and local vorticity structure. We observe distinct error reductions in vorticity (E_ω) and peak velocity (E_∞). These metrics characterize high-frequency features and extreme-value capture. PeSNO reduces these errors by 24.5% and 19.8% compared to Res U-Net. This demonstrates a specific advantage. The model resolves high-gradient local features beyond large-scale mean flows.

Table 2: Quantitative evaluation of flow field reconstruction accuracy on the test set. Metrics represent relative errors for the stream function (E_ψ), velocity field (E_u), vorticity field (E_ω), and peak velocity (E_∞).

Model	$E_\psi(L_2)$	$E_u(H^1)$	$E_\omega(H^2)$	$E_\infty(L_\infty)$
PeSNO (Ours)	0.2319	0.2909	0.5754	0.3709
Vanilla FNO	0.2857	0.3654	0.6270	0.4368
Res U-Net	0.2462	0.3160	0.7619	0.4627

Figure 2 visualizes flow field structures for a representative case. This case features strong internal heating and high-viscosity contrast. The dynamics involve horizontal flow beneath a stagnant lid and localized cold downwellings.

We first examine the global stream function geometry (Fig. 2, first column). PeSNO predictions (Fig. 2(d)) remain parallel to rigid boundaries. This avoids non-physical penetration and aligns with the Ground Truth (Fig. 2(a)). In contrast, Vanilla FNO (Fig. 2(g)) exhibits Gibbs phenomena. This results from relying on periodic Fourier bases for non-periodic vertical boundaries. It causes numerical artifacts near walls.

Visual differences in global velocity magnitude are subtle. However, the quantitative E_∞ metric reveals the performance gap in capturing extremes. Res U-Net yields higher peak error. This results from the over-smoothing tendency of standard convolutions. In contrast, PeSNO achieves lower E_∞ . This confirms that physics-augmented inputs and derivative constraints preserve sharp gradients at plume cores. In the velocity field (Figure 2, second column), the Res U-Net prediction (Figure 2(k)) captures the overall contour but underestimates velocity extremes in high-velocity cores, reflected in its high E_∞ error. This confirms that standard convolutional networks tend toward over-smoothing, failing to reconstruct sharp velocity peaks without explicit derivative constraints. Conversely, PeSNO (Figure 2(e)) successfully preserves sharp velocity gradients, facilitated by its physics-augmented inputs and the FiLM modulation mechanism.¹⁰

We also analyze the vorticity field (Fig. 2, third column). PeSNO separates positive and negative shear bands with continuous boundaries and minimal numerical noise (Fig. 2(f)). This fidelity results from the weighted Sobolev loss. This constraint ensures the network conforms to differential physics. In contrast, Res U-Net (Fig. 2(l)) produces discontinuous shear layers. This confirms that optimizing numerical values alone fails to capture high-frequency derivative features.

5.3. Long-term Spatiotemporal Consistency

We performed a 500-step auto-regressive evolution to evaluate error accumulation. The test case involves an unsteady system with strong internal heating and high-viscosity contrast. This regime exhibits high sensitivity to initial conditions.

Figure 3 compares PeSNO predictions with the numerical ground truth. PeSNO maintains phase consistency through step 250. The predicted cold drips in Fig. 3(f) match the numerical results in Fig. 3(b). A secondary downwelling appears in the PeSNO prediction at step 500. This feature is absent from the ground truth. Such trajectory divergence is expected in nonlinear convection. Numerical perturbations trigger instability paths once the evolution exceeds the system characteristic time. However, this emergent structure adheres to fluid mechanics laws regarding geometry and sinking velocity. This suggests PeSNO captures underlying physical operators rather than memorizing trajectories.

¹⁰[Zehua 10]: Response to "Hard to see": Acknowledged that visual differences are subtle. Pivoted to the quantitative E_∞ metric (Table 2) as the definitive evidence for peak capture accuracy, avoiding reliance on visual inspection.

We assessed macroscopic conservation by calculating statistical errors. Table 1 summarizes the results. The maximum error of the temperature profile remains under 1.01% after 500 steps. As shown in Fig. 3(d), PeSNO preserves the thermal boundary layer structure. Errors arise primarily from transient phase differences in downwelling fronts. The geothermal gradient and stagnant lid thickness remain accurate. This confirms that the secondary downwelling does not destabilize thermodynamic stratification.

Fig. 3(h) illustrates energy conservation. PeSNO exhibits minimal energy deviation ($L_\infty \approx 0.15\%$). The slight decay in mean temperature indicates numerical dissipation. This characteristic resembles high-precision numerical schemes. It suppresses non-physical energy accumulation at the grid scale and ensures long-term stability.

5.4. Operator Robustness to Input Perturbations

We evaluate whether PeSNO has learned physical operators with inherent regularization by performing a sensitivity analysis against input noise. In practical mantle convection simulations, temperature fields are frequently subjected to numerical truncation or observational errors. A robust operator must suppress high-frequency non-physical noise rather than amplifying it into spurious physical gradients.

We quantify output deviation relative to a noise-free baseline using the Relative Drift Error (E_{drift}):

$$E_{\text{drift}} = \frac{\|\hat{\mathbf{u}}_\sigma - \hat{\mathbf{u}}_{\text{clean}}\|_2}{\|\hat{\mathbf{u}}_{\text{clean}}\|_2} \quad (16)$$

Here, $\hat{\mathbf{u}}_\sigma$ and $\hat{\mathbf{u}}_{\text{clean}}$ are the velocity fields predicted from noisy and noise-free inputs, respectively. Furthermore, we employ the anisotropic Total Variation (TV) norm to detect non-physical oscillations [31]:

$$\text{TV}(\mathbf{u}) = \frac{1}{N} \sum_{(i,j)} (\|(\nabla u)_{i,j}\|_1 + \|(\nabla v)_{i,j}\|_1) \quad (17)$$

Here, N is the total number of grid points. The TV norm measures spatial smoothness; a sharp increase indicates the emergence of high-frequency artifacts.

We perturbed the input temperature field with additive Gaussian white noise ($\sigma \in [0.00, 0.05]$). Table 3 summarizes the quantitative results across varying noise levels.

The results show that Res U-Net fails even at $\sigma = 0.01$. The drift error E_{drift} rises to 70.27. The TV norm reaches 2533.60. This confirms that standard convolutional networks amplify stochastic noise. They lack physical constraints. In contrast, PeSNO maintains stability at $\sigma = 0.05$. The error E_{drift} remains 0.87 and the TV norm stays at 1029.95.

Figure 4 illustrates this disparity. Under $\sigma = 0.05$ noise (Fig. 4(m)), the Res U-Net velocity field (Fig. 4(o)) shows a noisy distribution. Artifacts obscure primary convective structures. This highlights the absence of regularization. Conversely, the PeSNO output (Fig. 4(n)) remains clear. This stability stems from the combined spectral architecture and Sobolev loss. Spectral layers act as low-pass filters to suppress stochastic noise. The Sobolev loss enforces differential

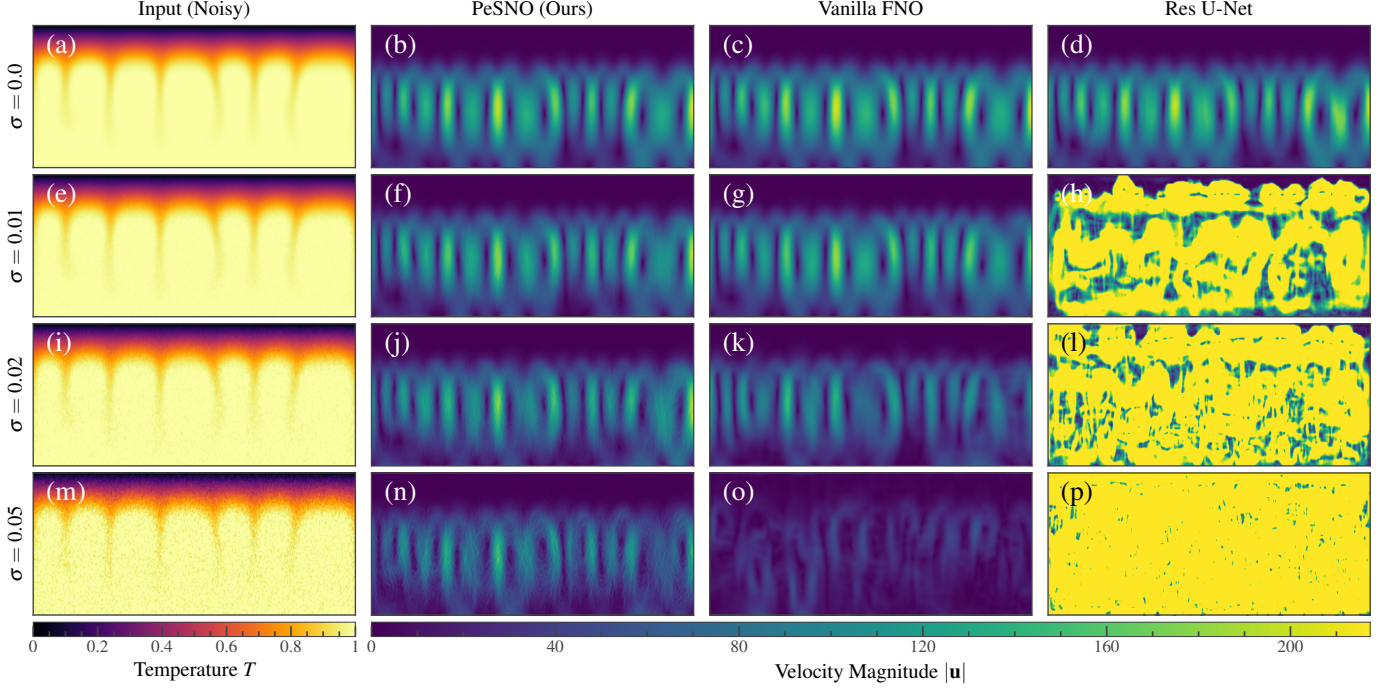


Figure 4: Operator robustness against input perturbations ($\sigma \in [0.00, 0.05]$). The leftmost column displays the noisy input temperature T . Subsequent columns compare predicted velocity magnitude $|u|$ for PeSNO, Vanilla FNO, and Res U-Net. Rows correspond to increasing noise intensity σ .

Table 3: Quantitative robustness analysis under varying input noise levels. E_{drift} denotes deviation from the clean baseline; TV Norm measures spatial smoothness.

Noise Level (σ)	Model	Drift Error (E_{drift})	Smoothness (TV Norm)
0.00 (Clean)	PeSNO (Ours)	0.2965	753.36
	Vanilla FNO	0.3686	744.87
	Res U-Net	0.3527	1218.68
0.01	PeSNO (Ours)	0.4457	765.95
	Vanilla FNO	0.5355	745.13
	Res U-Net	70.2719	2533.60
0.05	PeSNO (Ours)	0.8734	1029.95
	Vanilla FNO	2.6124	410.65
	Res U-Net	161.2299	9632.94

constraints to preserve physical shear features. While Vanilla FNO outperforms Res U-Net, its error rises to 2.61 at high noise levels. Its TV norm drops to 410.65. This indicates over-smoothing (Fig. 4(p)). Furthermore, our operator formulation supports mesh-independent inference. We demonstrate zero-shot super-resolution in Section 5.5. This stability originates from the hybrid spectral architecture and the Sobolev loss: the spectral layers perform implicit low-pass filtering via wavenumber truncation, while the vorticity constraint enforces physically consistent differential structures. While Vanilla FNO outperforms Res U-Net, its error increases significantly to 2.61 at high noise levels, accompanied by an abnormal drop in the TV norm (410.65). This suggests over-smoothing (Fig. 4(p)), which blurs physical edges. These findings verify that PeSNO captures a continuous physical operator rather than a discrete pixel-wise mapping.¹¹

5.5. Zero-Shot Spectral Super-Resolution Analysis

Neural operators differ from standard convolutional networks by learning mappings between continuous function spaces. This property allows for discretization invariance. We evaluate this capability through a zero-shot super-resolution test. The models were trained on a $1\times$ coarse grid. We then queried them directly at $2\times$ and $4\times$ resolutions. The high-resolution inputs were generated via bilinear interpolation of the source fields.^{Zehua}

Figure 5 presents the spectral behaviors of the different architectures. The Res U-Net exhibits limited generalization. The $2\times$ predictions (Fig. 5(i)–(j)) retain basic contours. However, the $4\times$ outputs (Fig. 5(k)–(l)) contain high-frequency pixelated noise. This indicates that standard convolutional networks learn discrete local correlations tied to the training grid. They lack

spectral layers filter noise while Sobolev loss preserves shear features. 2. Retained FNO data to demonstrate over-smoothing. 3. Referenced Section 5.5 for the mesh-independence test.

¹¹[Zehua 11]: Response to "Mechanism and Resolution": 1. Explained that

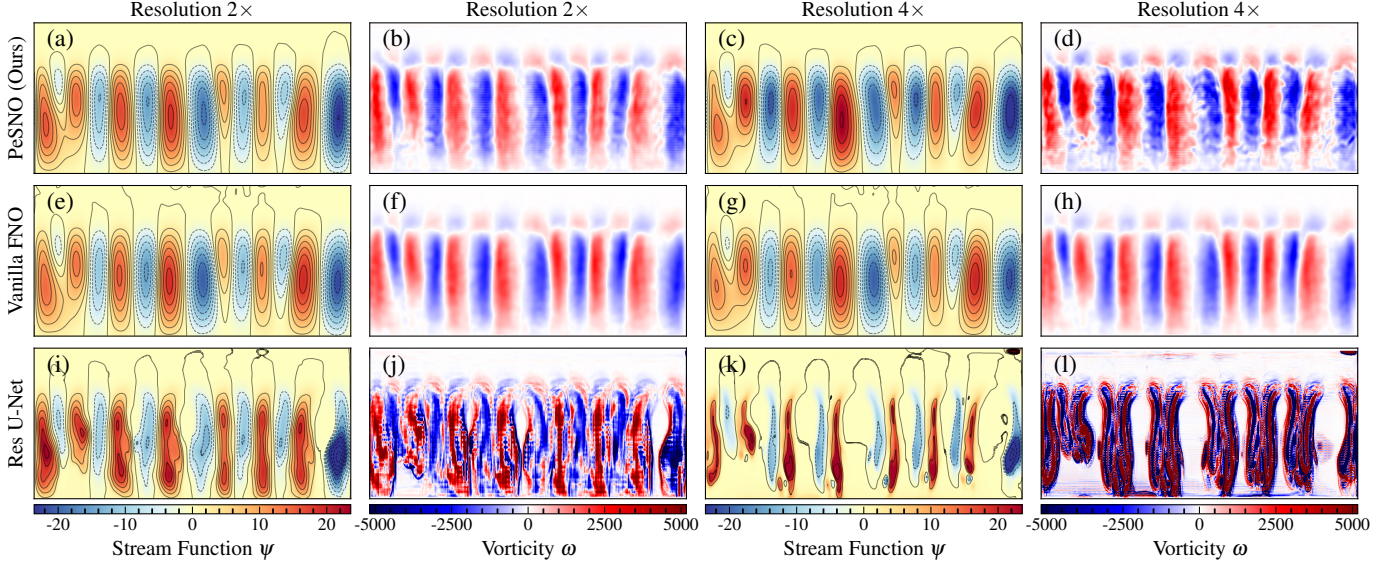


Figure 5: Zero-shot spectral super-resolution analysis. Rows compare PeSNO, Vanilla FNO, and Res U-Net. The first two columns display the Stream Function ψ and Vorticity ω at $2\times$ resolution. Subsequent columns show the corresponding fields at $4\times$ resolution.

the continuous functional basis required for mesh-independent inference.^{Zehua}

The comparison between Vanilla FNO and PeSNO indicates a trade-off between spectral smoothness and structural acuity. The Vanilla FNO (Fig. 5(e)–(h)) produces smooth fields. However, the high-shear bands in the $4\times$ vorticity field (Fig. 5(h)) appear diffused. This suggests that the FNO acts as a low-pass filter. It truncates the high-wavenumber energy required for sharp gradients.^{Zehua}

In contrast, PeSNO (Fig. 5(a)–(d)) retains high-frequency features. The vorticity shear bands (Fig. 5(d)) remain localized at $4\times$ resolution. Faint horizontal oscillating stripes appear in the prediction. These are structured Gibbs-like phenomena. The Fourier feature mapping introduces high-frequency sinusoidal bases into the input space. The Sobolev loss penalizes derivative errors and assigns weights to these bases to fit shear gradients. The subsequent hard truncation in the spectral layers, combined with boundary effects from local convolutions, produces the observed spatial ringing. This characteristic preserves the spatial structure of the gradient field. The model approximates the differential operator and maintains kinematic consistency over visual smoothness.^{Zehua}

5.6. Accelerating CFD via Warm-Start

We evaluate the computational efficiency gained by using PeSNO as a high-fidelity warm-start initializer for traditional numerical solvers. In projection-based incompressible flow simulations, the pressure projection step eliminates velocity field divergence and consumes the majority of computational resources. This is especially evident in mantle convection problems where viscosity varies by nine orders of magnitude, leading to extreme numerical stiffness in the pressure Poisson equation and slow solver convergence. We benchmark the warm-start efficacy by comparing solver performance using the

Teacher model (NewFluidNet) versus the PeSNO model for initial guesses. We conducted both single-step cold start tests and 100-step sequential rollout tests, with results summarized in Table 4.

As shown in Table 4, the Teacher model exhibits an initial divergence peak of 9.91×10^2 , indicating a significant deviation from the incompressibility constraint. In contrast, PeSNO yields an initial divergence of only 3.50×10^{-3} , representing a five-order-of-magnitude reduction. Crucially, both methods ultimately converge to the same residual level (10^{-6}), ensuring a consistent baseline for performance benchmarking. This confirms that PeSNO’s acceleration does not stem from relaxing convergence standards but from an initial guess that aligns closely with the divergence-free manifold.

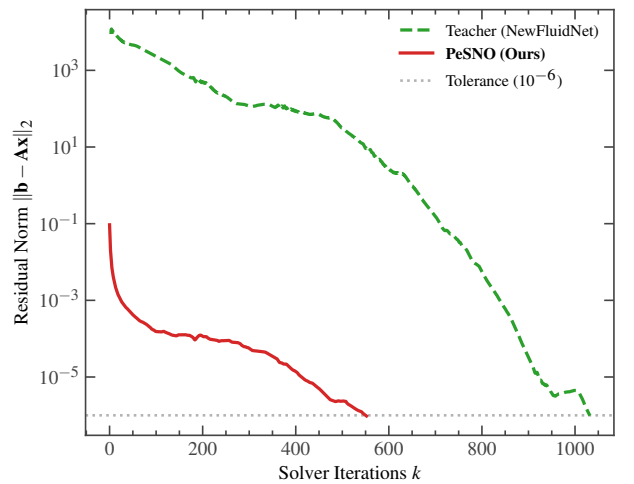


Figure 6: Residual convergence history of the pressure Poisson solver. The solver iterations start from initial guesses provided by PeSNO and the Teacher model, respectively. The horizontal dotted line marks the convergence tolerance ($\epsilon = 10^{-6}$).

The residual convergence curves in Figure 6 illustrate the dy-

Table 4: Comparison of solver acceleration performance. Init Div (L_∞) and Final Res (L_2) verify that both methods converge to the same physical standard (10^{-6}). Sequential Efficiency reflects the amortized cost per step over a 100-step rollout.

Model	Physics & Numerics				Single-Step Efficiency			Sequential Efficiency
	Init Div (L_∞)	Final Res (L_2)	Vel Corr.	Iter.	Inf. Time	Total	Speedup	Avg. Time
Teacher	9.91×10^2	9.78×10^{-7}	5.53×10^4	1033	0.49 s	1.05 s	1.00×	0.75 s
PeSNO	3.50×10^{-3}	9.62×10^{-7}	1.04×10^4	553	0.01 s	0.38 s	2.73×	0.15 s

namic advantages of the PeSNO initialization. The Teacher model exhibits a prolonged stagnation period, requiring 1033 iterations to reach the convergence threshold. Conversely, PeSNO begins with a significantly lower residual and requires only 553 iterations. When accounting for both model inference and numerical solution time, PeSNO achieves a 2.73x end-to-end acceleration in single-step tests. **Part of this efficiency stems from the distribution shift in training data. The baseline targets raw collocated fields. In contrast, PeSNO learns the projected MAC-grid state. This target explicitly satisfies mass conservation. Consequently, the solver bypasses initial correction steps required for raw predictions.**¹²

In the 100-step sequential rollout, the acceleration ratio for PeSNO increases to 5.06x, significantly higher than the single-step gain. This non-linear boost results from amortizing fixed overheads. **The Teacher model relies on a heavy multi-scale U-Net architecture with learned boundary layers. This limits its inference speed to 0.49s. In contrast, PeSNO utilizes a lightweight spectral design. It achieves a negligible 0.01s latency while retaining full parametric coverage via FiLM. This efficiency allows the solver to operate at near-numerical speeds with a 50% reduction in iterations. While the Teacher model is limited by its 0.49s inference cost, PeSNO's negligible 0.01s latency allows the solver to operate at near-numerical speeds while benefiting from a 50% reduction in iterations.**¹³ These results demonstrate that PeSNO is an efficient, physics-consistent initializer that mitigates the computational burden of long-term simulations. For an evaluation of acceleration stability under different rheological conditions, refer to [Appendix B](#).

6. Conclusion

This study introduces the Physics-embedded Stream Neural Operator (PeSNO), a framework designed to accelerate high-precision numerical solvers rather than replace them. Unlike black-box surrogate models, PeSNO serves as a robust warm-start mechanism for the pressure projection step in incompressible flow simulations. **The framework also integrates a data correction strategy. This pipeline regularizes baseline predictions**

into mass-conservative training targets. Consequently, the operator learns consistent physical dynamics from noisy references. The resulting physical fidelity exceeds the approximate source data.¹⁴

Our results demonstrate that mathematical compatibility between the neural operator and boundary conditions ensures conservation laws. PeSNO integrates hybrid spectral layers with anti-symmetric reflection padding. This reconciles zero-flux constraints with differential continuity on discrete grids. It eliminates Gibbs artifacts found in conventional spectral methods. Furthermore, the Sobolev loss penalizes high-order derivatives. This constraint prevents over-smoothing. It allows the model to reproduce nonlinear flow instabilities during long-term evolution.

We quantified the utility of PeSNO as a solver preconditioner. The model maps initial guesses close to the divergence-free manifold. This bypasses convergence stagnation associated with traditional cold starts. Experimental data show that PeSNO achieves a 5.06x improvement in end-to-end sequential throughput. The solver strictly maintains the 10^{-6} residual standard. This confirms that deep learning models function as high-performance acceleration components in numerical workflows.

Future work will integrate PeSNO into mature mantle convection codes. This will establish a production tool capable of handling planetary-scale evolution across a broad parameter space.

References

- [1] S. Agarwal, A. C. Bekar, C. Hüttig, et al., Physics-based machine learning for mantle convection simulations, *Physics of Fluids* 37 (8) (2025).
- [2] F. H. Harlow, J. E. Welch, Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface, *Physics of fluids* 8 (12) (1965) 2182.
- [3] J. R. Shewchuk, An introduction to the conjugate gradient method without the agonizing pain (1994).
- [4] A. J. Chorin, Numerical solution of the navier-stokes equations, *Mathematics of computation* 22 (104) (1968) 745–762.

¹²[Zehua 12]: Response to "Data distribution shift": Acknowledged. Clarified that the efficiency gain stems from the distinct training target: PeSNO learns the projected MAC-grid state (mass-conservative), whereas the baseline targets raw collocated fields. This allows the solver to bypass initial correction steps.

¹³[Zehua 13]: Response to "Why newfluidnet is slow / Capacity": 1. Explained latency difference: Heavy U-Net (Teacher) vs. Lightweight Spectral (PeSNO). 2. Confirmed capacity: PeSNO retains full parametric coverage via FiLM.

¹⁴[Zehua 14]: Response to "Data correction strategy": Acknowledged. Highlighted the data correction pipeline. Clarified that regularizing baseline predictions into mass-conservative targets allows the model to learn physical dynamics exceeding the fidelity of the approximate source data.

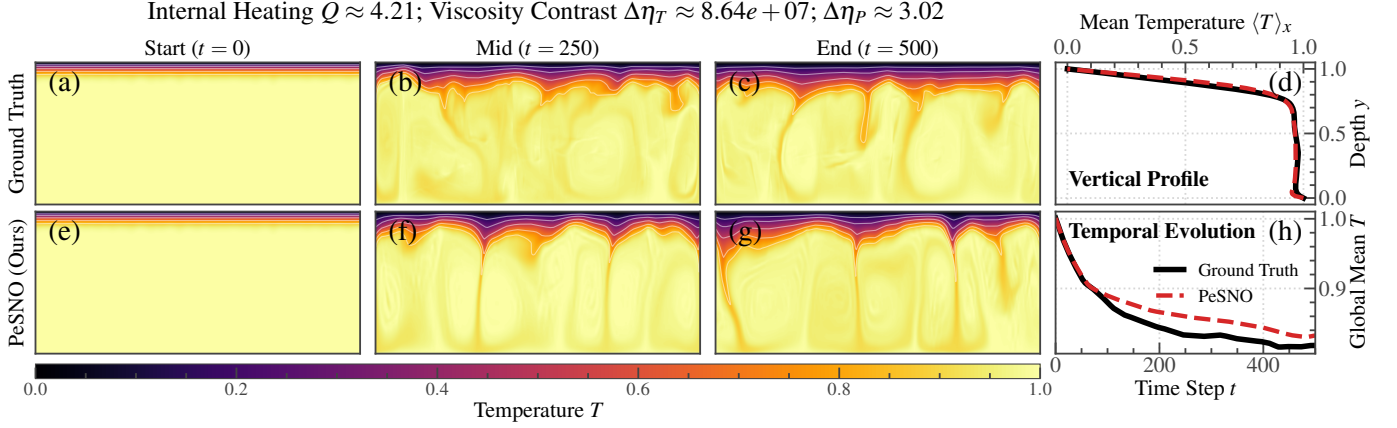


Figure A1: Long-term spatiotemporal evolution in the high-energy vigorous convection regime ($Q \approx 4.21$, $\Delta\eta_T \approx 8.64 \times 10^7$). Note that despite high viscosity contrast, strong internal heating reduces internal viscosity and increases flow velocity. (a–c, e–g) Temperature snapshots at $t = 0, 250, 500$. White lines denote isotherms. (d) Horizontally averaged temperature profile $\langle T \rangle_x$ at the final step. (h) Global mean temperature time series. [\[video\]](#)

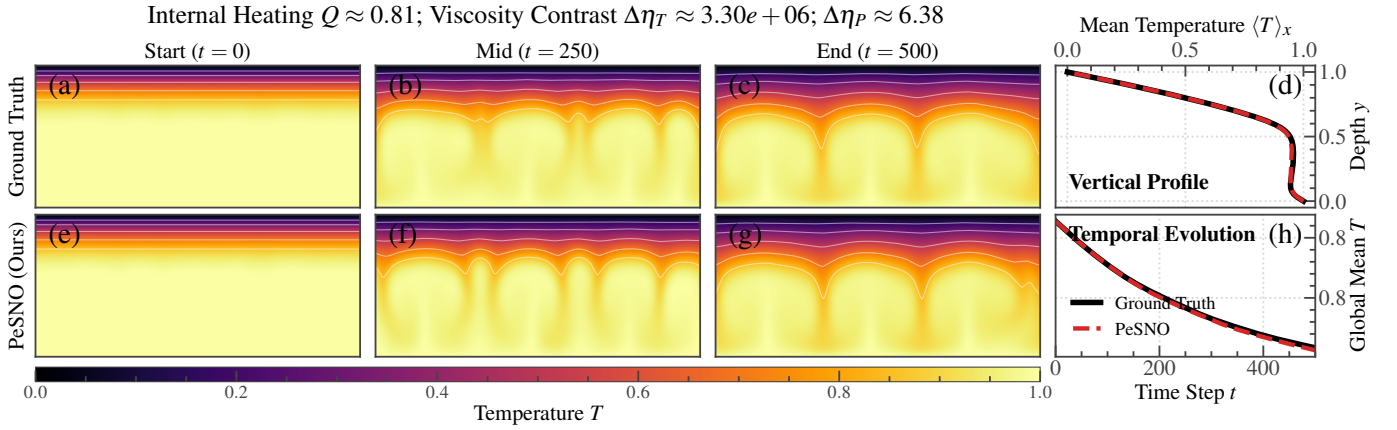


Figure A2: Long-term spatiotemporal evolution in the low-energy sluggish convection regime ($Q \approx 0.81$, $\Delta\eta_T \approx 3.30 \times 10^6$). (a–c, e–g) Temperature snapshots at $t = 0, 250, 500$ (Top: Ground Truth; Bottom: PeSNO). White lines denote isotherms. (d) Horizontally averaged temperature profile $\langle T \rangle_x$ at the final step. (h) Global mean temperature time series. [\[video\]](#)

- [5] P. L. Roe, Characteristic-based schemes for the euler equations, Annual review of fluid mechanics 18 (1) (1986) 337–365.
- [6] J. C. Butcher, Numerical methods for ordinary differential equations, John Wiley & Sons, 2016.
- [7] V. Solomatov, Scaling of temperature-and stress-dependent viscosity convection, Physics of Fluids 7 (2) (1995) 266–274.
- [8] M. Tancik, P. Srinivasan, B. Mildenhall, et al., Fourier features let networks learn high frequency functions in low dimensional domains, Advances in neural information processing systems 33 (2020) 7537–7547.
- [9] N. Rahaman, A. Baratin, D. Arpit, et al., On the spectral bias of neural networks, in: International conference on machine learning, 2019, pp. 5301–5310.
- [10] E. Perez, F. Strub, H. De Vries, et al., Film: Visual reasoning with a general conditioning layer, in: Proceedings of the AAAI conference on artificial intelligence, Vol. 1, 2018.
- [11] Z. Li, N. B. Kovachki, K. Azizzadenesheli, et al., Fourier neural operator for parametric partial differential equations, in: International Conference on Learning Representations, 2021.
- [12] C. Canuto, M. Y. Hussaini, A. Quarteroni, et al., Spectral methods: fundamentals in single domains, Springer, 2006.
- [13] J. P. Boyd, Chebyshev and Fourier spectral methods, Courier Corporation, 2001.
- [14] G. Liu, F. A. Reda, K. J. Shih, et al., Image inpainting for irregular holes using partial convolutions, in: Proceedings of the European conference on computer vision (ECCV), 2018, pp. 85–100.
- [15] I. E. Lagaris, A. Likas, D. I. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations, IEEE transactions on neural networks 9 (5) (1998) 987–1000.
- [16] D. Goldberg, What every computer scientist should know about floating-point arithmetic, ACM computing surveys (CSUR) 23 (1) (1991) 5–48.

Table B1: Performance comparison of linear solver initialization. Init_Div denotes the L_∞ norm of velocity divergence before projection. Steps indicates iterations required to reach a residual $< 10^{-6}$.

Case	Model	Init Div (L_∞)	Final Res (L_2)	Steps	Total Time (s)
High-Energy ($Q \approx 4.21$)	Teacher	3.37×10^4	9.94×10^{-7}	1106	1.14
	PeSNO	6.13×10^{-2}	9.56×10^{-7}	652	0.44
Low-Energy ($Q \approx 0.81$)	Teacher	7.41×10^2	9.59×10^{-7}	1022	1.18
	PeSNO	1.77×10^{-3}	9.95×10^{-7}	538	0.42

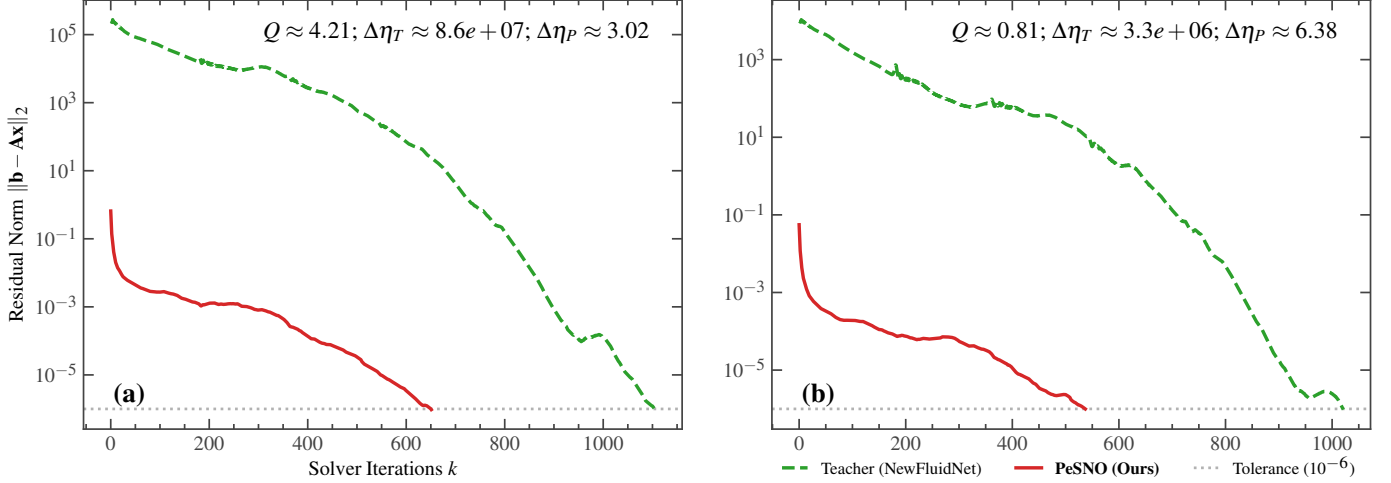


Figure B1: Residual convergence history of the pressure Poisson solver across different heating regimes. Representative instantaneous iterations ($t = 300$) for (a) high-energy ($Q \approx 4.21$) and (b) low-energy ($Q \approx 0.81$) regimes, comparing initial guesses from PeSNO (solid lines) and the Teacher model (dashed lines). The horizontal dotted line marks the convergence tolerance ($\epsilon = 10^{-6}$).

- [17] W. M. Czarnecki, S. Osindero, M. Jaderberg, et al., Sobolev training for neural networks, in: Advances in neural information processing systems, Vol. 30, 2017.
- [18] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving non-linear partial differential equations, Journal of Computational physics 378 (2019) 686–707.
- [19] D. Anderson, J. C. Tannehill, R. H. Pletcher, et al., Computational fluid mechanics and heat transfer, CRC press, 2020.
- [20] M. Ren, W. Zeng, B. Yang, et al., Learning to reweight examples for robust deep learning, in: International conference on machine learning, 2018, pp. 4334–4343.
- [21] Y. Bengio, J. Louradour, R. Collobert, et al., Curriculum learning, in: Proceedings of the 26th annual international conference on machine learning, 2009, pp. 41–48.
- [22] A. Krishnapriyan, A. Gholami, S. Zhe, et al., Characterizing possible failure modes in physics-informed neural networks, Advances in neural information processing systems 34 (2021) 26548–26560.
- [23] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, 2015, pp. 234–241.
- [24] K. He, X. Zhang, S. Ren, et al., Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
- [25] Z. Zhang, Q. Liu, Y. Wang, Road extraction by deep residual u-net, IEEE Geoscience and Remote Sensing Letters 15 (5) (2018) 749–753.
- [26] R. A. Adams, J. J. Fournier, Sobolev spaces, Vol. 140, Elsevier, 2003.
- [27] G. K. Batchelor, An introduction to fluid dynamics, Cambridge university press, 2000.
- [28] A. Quarteroni, A. Valli, Numerical approximation of partial differential equations, Springer, 1994.
- [29] L. N. Trefethen, Spectral methods in MATLAB, SIAM, 2000.
- [30] R. J. LeVeque, Finite difference methods for differential equations, Draft version for use in AMath 585 (6) (1998) 112.
- [31] L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D: nonlinear phenomena 60 (1-4) (1992) 259–268.

Appendix A. Generalization Across Rheological Conditions

We evaluate PeSNO generalization across a broad parameter space by examining the FiLM mechanism. We compare two regimes. These include high-intensity vigorous convection and sluggish convection.

Figure A1 illustrates flow evolution under high internal heating ($Q \approx 4.21$) and strong viscosity variation ($\Delta\eta_T \approx 10^7$). PeSNO accurately reconstructs the initial flow structure. Morphological deviations from the NewFluidNet baseline are expected. The baseline uses a Finite Element formulation (Citcom). In contrast, our ground truth generator employs a Staggered Grid Finite Difference solver. Discretization errors differ between these schemes. In chaotic regimes, these microscopic differences amplify over time. Additionally, our data pipeline applies inertial blending and divergence-free projection. This enforces strict mass conservation but slightly modifies instantaneous flow structures. PeSNO learns this specific physics-consistent target rather than the raw Citcom output. Consequently, the temperature profile in Fig. A1(d) remains consistent with the stagnant lid physics despite spatial offsets. However, by $t = 500$, the predicted spatial morphology deviates from the numerical ground truth. This trajectory separation stems from reduced internal viscosity, which enhances system nonlinearity and renders the flow highly sensitive to initial numerical perturbations. Despite this phase shift, the horizontally averaged temperature profile in Figure A1(d) remains consistent with the ground truth. This confirms the model correctly captures the thermal boundary layer and geothermal gradient governed by the high-viscosity stagnant lid.¹⁵ The global mean temperature in Fig. A1(h) exhibits a minor positive drift. This suggests thermal accumulation in unsteady flows. However, the evolution remains stable.

Figure A2 displays sluggish convection under low internal heating ($Q \approx 0.81$). The weaker driving force results in regular flow evolution. PeSNO maintains phase consistency over 500 time steps without morphological differences. Predicted curves in Fig. A2(d) and Fig. A2(h) coincide with the ground truth. This confirms that the Sobolev loss constrains the differential structure and energy conservation under weak nonlinear conditions.

These results demonstrate that PeSNO captures underlying physical operators. It avoids memorizing trajectories. The model achieves precise spatiotemporal reconstruction in low-energy regimes. It maintains macroscopic statistical conservation in chaotic flows.

Appendix B. Robustness of Solver Acceleration

We quantitatively evaluate the acceleration performance of the hybrid solver across different dynamic regimes. We analyze

single-step efficiency using instantaneous snapshots at $t = 300$. We compare high-intensity and low-intensity convection.

Table B1 summarizes performance metrics. In the high-intensity regime ($Q \approx 4.21$), the viscosity contrast reaches 10^7 . This induces sharp velocity gradients. The Teacher model yields an initial divergence of 3.37×10^4 . This violates the incompressibility constraint. In contrast, PeSNO reduces the initial divergence to 6.13×10^{-2} . This represents a reduction of six orders of magnitude. The method provides a physically consistent initial guess. This efficiency reduces iterations from 1106 to 652. The low-intensity regime ($Q \approx 0.81$) shows similar gains. PeSNO efficacy remains independent of driving intensity.

Figure B1 illustrates residual convergence trajectories. In the high-intensity field, the Teacher model exhibits a stagnation plateau. This implies the initial error contains high-frequency components difficult for the solver. Conversely, PeSNO shows a monotonic descent in both states. It avoids bottlenecks. Initial values align with the physical pressure gradient. This prevents stagnation and ensures robustness.

¹⁵[Zehua 15]: Response to "Contours look off / Difference from NewFluidNet": 1. Explained that deviations stem from the solver mismatch: Citcom (FEM) vs. our Staggered Grid solver (FDM). 2. Clarified that our projection strategy ensures mass conservation, creating a different target from the raw Citcom output.