Algorithms and Intelligence

Vol 2 Issue 1 (2026)

www.algointel.org

Mean-Flow Neural Operators for Long-Horizon Dynamical System Prediction

Seulgi Nam, Valentin Aubert

Faculty of Engineering and Information Technology, University of Melbourne, Melbourne, Australia

Abstract

The prediction of complex dynamical systems over extended temporal horizons remains a fundamental challenge in computational science. Traditional numerical solvers often suffer from compounding discretization errors and prohibitive computational costs when simulating multiscale phenomena. Recently, neural operators have emerged as a promising paradigm for learning mesh-independent mappings between function spaces, yet they frequently struggle with error accumulation and spectral bias during autoregressive long-term rollouts. To address these limitations, this paper introduces the Mean-Flow Neural Operator, a novel architecture designed to explicitly decouple the macroscopic mean-flow dynamics from microscopic transient fluctuations. By embedding domain-agnostic scale separation directly into the operator learning framework, the proposed method preserves global conservation properties while dynamically modeling high-frequency residual interactions. Comprehensive empirical evaluations across diverse physical systems demonstrate that the Mean-Flow Neural Operator significantly attenuates long-horizon error propagation. Furthermore, the approach achieves substantial computational acceleration compared to high-fidelity numerical baselines. This work establishes a robust foundation for the deployment of data-driven surrogate models in critical applications requiring sustained predictive stability, ranging from climate modeling to advanced aerospace engineering.

Keywords

Neural Operators, Dynamical Systems, Long-Horizon Prediction, Scale Separation

INTRODUCTION

1.1 Context and Motivation

The accurate simulation and prediction of complex dynamical systems are foundational to numerous scientific and engineering disciplines. From the atmospheric circulation patterns that govern global climate to the chaotic fluid dynamics within aerospace propulsion systems, understanding how these continuous systems evolve over time is critical for both theoretical advancement and practical technological development. Historically, the scientific community has

relied almost exclusively on classical numerical methods to approximate the solutions of the underlying partial differential equations that govern these phenomena [1]. These methods, which include finite difference, finite volume, and finite element approaches, have reached a high degree of maturity and sophistication. However, despite their rigorous mathematical foundations and well-understood convergence properties, traditional numerical solvers are inherently constrained by the fundamental trade-off between spatial-temporal resolution and computational feasibility [2]. When dealing with multiscale systems exhibiting chaotic behavior, capturing the finest scales of turbulence or structural deformation requires mesh grids of extraordinary density [3]. The computational cost associated with such high-fidelity simulations, commonly referred to as direct numerical simulations, scales exponentially with the required resolution, rendering them impractical for long-horizon forecasting or rapid iterative design optimization [4]. Furthermore, the strict stability criteria imposed by numerical integration schemes often necessitate excessively small time steps, further exacerbating the computational burden. This limitation becomes particularly restrictive in applications requiring ensemble forecasting or real-time control, where rapid state estimation and future prediction are non-negotiable requirements [5]. In recent years, the exponential growth in computational power and the accumulation of vast repositories of high-fidelity simulation data have catalyzed a paradigm shift toward data-driven modeling [6]. Machine learning algorithms, particularly deep neural networks, have demonstrated remarkable success in identifying latent patterns within complex datasets, prompting researchers to explore their utility as surrogate models for physical systems [7]. The allure of these data-driven approaches lies in their ability to perform rapid inference once trained, potentially offering orders-of-magnitude acceleration compared to traditional solvers while maintaining a completely differentiable pipeline [8].

1.2 The Challenge of Long-Horizon Prediction

Despite the enthusiasm surrounding deep learning for physical simulations, deploying standard neural network architectures for dynamic state prediction presents significant conceptual and practical challenges. Conventional convolutional neural networks and recurrent neural networks are inherently bound to the specific spatial grid and temporal discretization upon which they are trained [9]. This rigid dependency implies that a model trained on a specific mesh resolution cannot be directly evaluated on a different, perhaps finer, grid without extensive interpolation or complete retraining [10]. Moreover, the fundamental nature of continuous physical systems is that their state variables are functions defined over continuous domains, whereas standard neural networks operate on discrete, finite-dimensional vectors [11]. This conceptual mismatch has historically limited the generalizability and robustness of data-driven physics models. To overcome the limitations of grid-dependent architectures, the framework of neural operators has been proposed and rapidly adopted by the computational physics community. Neural operators generalize the concept of neural networks to infinite-dimensional function spaces, learning mappings from one continuous function to another rather than merely transforming discrete vectors [12]. By learning the underlying operator that maps an initial condition to a future state, these models achieve zero-shot super-resolution, meaning they can be evaluated at any arbitrary spatial resolution during inference regardless of the training grid [13]. Among the various instantiations of operator learning, the Fourier Neural Operator has achieved particular prominence due to its elegant parameterization of the integral kernel in the spectral domain [14]. By utilizing the fast Fourier transform, it efficiently captures global spatial correlations, enabling highly accurate predictions for a wide class of partial differential equations. However, a critical vulnerability remains unresolved within the current paradigm of neural operator research: the catastrophic accumulation of error during long-horizon autoregressive rollouts. In a typical forecasting scenario, a neural operator is trained to predict the system state at a specific future time step based on the current

state [15]. To predict further into the future, the model's output is recursively fed back as the input for the next prediction step. Because data-driven models inevitably produce a non-zero approximation error at each forward pass, this error is continuously ingested and amplified through successive iterations [16]. In chaotic dynamical systems, where infinitesimally small perturbations grow exponentially over time, this compounding error rapidly degrades the prediction quality, leading to unphysical states, numerical explosions, or collapse into trivial equilibrium modes [17]. Furthermore, spectral architectures often exhibit an inherent bias toward capturing low-frequency, macroscopic features while gradually smoothing out high-frequency transient details during recursive inference, resulting in an artificial dissipation of energy that contradicts the physical conservation laws governing the system.

1.3 Proposed Contributions

Recognizing the limitations of existing operator learning techniques in sustaining predictive fidelity over extended temporal horizons, this paper introduces a novel architectural framework termed the Mean-Flow Neural Operator. The core philosophy of this approach is inspired by traditional techniques in fluid mechanics, specifically the Reynolds decomposition, which posits that a complex flow field can be mathematically separated into a slowly evolving macroscopic mean-flow component and a rapidly fluctuating microscopic transient component. By embedding this principle of scale separation directly into the latent space of a neural operator, the proposed architecture explicitly decouples the low-frequency global dynamics from the high-frequency localized perturbations. The primary contribution of this work is the development of a dual-branch continuous operator that independently propagates the mean and fluctuating components while facilitating controlled, physically grounded interactions between the two scales. This separation allows the model to enforce strict stability constraints on the macroscopic branch, ensuring long-term conservation of mass and momentum, while dedicating specialized expressive capacity to the microscopic branch to capture chaotic transient phenomena without allowing their inherent unpredictability to corrupt the global structural integrity of the simulation. Extensive theoretical formulation and empirical validation are provided to demonstrate that the Mean-Flow Neural Operator significantly outperforms state-of-the-art baselines in minimizing autoregressive error accumulation, thereby extending the viable horizon of accurate data-driven prediction for multiscale dynamical systems.

2. Literature Review and Background

2.1 Traditional Numerical Methods for Dynamical Systems

The mathematical modeling of continuous physical systems is predominantly formulated through partial differential equations, which describe the temporal evolution of spatial quantities based on fundamental conservation laws. Solving these equations analytically is generally impossible for realistic, non-linear scenarios, necessitating the use of numerical approximation techniques [18]. The finite difference method represents the earliest and conceptually simplest approach, wherein continuous spatial and temporal derivatives are approximated using discrete Taylor series expansions over a structured grid [19]. While straightforward to implement, finite difference methods struggle with complex geometric boundaries and are highly susceptible to numerical instability if the Courant-Friedrichs-Lewy condition is violated, heavily constraining the permissible time step size [20]. Finite volume methods, by contrast, focus on integrating the governing equations over discrete control volumes, explicitly enforcing the conservation of fluxes across volume boundaries [21]. This inherent conservatism has made finite volume techniques the industry standard in computational fluid dynamics, particularly for simulating compressible flows with shock waves. Meanwhile, finite element methods construct solutions using piecewise polynomial basis functions over unstructured meshes, offering unparalleled flexibility in modeling highly irregular geometries and solid mechanics phenomena [22]. Despite their distinct

advantages, all these classical numerical methods share a common, inescapable limitation: they require a discretized representation of the domain [23]. As the complexity of the dynamic behavior increases, such as in high-Reynolds-number turbulent flows or highly nonlinear multiphysics interactions, the spatial and temporal resolution required to capture the essential physics grows prohibitively large, rendering direct numerical simulation computationally intractable for routine engineering design or long-term forecasting [24].

2.2 Deep Learning for Physical Systems

The limitations of traditional numerical solvers have driven a surge of interest in leveraging machine learning to approximate the solutions of partial differential equations. Early efforts focused on utilizing standard architectures, such as convolutional networks, to map initial state variables to future states in a discrete image-to-image translation framing [25]. While these models successfully demonstrated immense inference speedups, they were fundamentally constrained by their discretization dependency. If trained on a grid of specific dimensions, the convolutional network could not adapt to a different grid resolution without extensive retraining or loss of accuracy, effectively limiting its utility as a general-purpose solver [26]. A significant breakthrough in combining deep learning with physical principles was the development of physics-informed neural networks. Instead of relying solely on massive datasets of pre-computed simulations, physics-informed neural networks incorporate the governing differential equations directly into the loss function used during training [27]. By minimizing the residual of the underlying physical laws across a set of collocation points sampled from the continuous domain, the network learns to approximate the solution while explicitly adhering to known physics [28]. This approach allows for mesh-free, continuous state estimation and has proven highly effective for solving inverse problems, where unknown system parameters or boundary conditions must be inferred from sparse, noisy observation data [29]. However, standard physics-informed formulations are primarily designed to find a single solution to a specific instance of a boundary value problem. They do not learn the generalized mapping from an arbitrary initial condition to a future state, meaning that a new network must be iteratively optimized from scratch for every new scenario, which limits their applicability for rapid, real-time dynamical forecasting.

2.3 Operator Learning and Continuous Formulations

To bridge the gap between rapid inference and continuous, generalized representations, the paradigm of operator learning was conceptualized. Neural operators elevate the learning task from finite-dimensional vectors to infinite-dimensional function spaces [30]. The defining characteristic of a neural operator is its ability to produce identical functional outputs regardless of the discretization of the input data, achieving true mesh invariance. The foundational architecture in this domain is the Deep Operator Network, which comprises two parallel subnetworks: a branch network that encodes the discrete input function into a latent representation, and a trunk network that evaluates the continuous domain coordinates [31]. The inner product of these two representations yields the final continuous solution. Building upon the continuous operator framework, the Fourier Neural Operator introduced a highly efficient mechanism for learning integral operators via convolution theorem in the frequency domain. By transforming the spatial input data using a discrete Fourier transform, applying a learned linear transformation to the lower-frequency modes, and applying an inverse Fourier transform, the model effectively performs a global convolution operation that captures long-range spatial dependencies [32]. The truncation of higher-frequency modes serves as an implicit regularization mechanism, filtering out high-frequency noise and yielding a smooth, continuous mapping [33]. Despite its immense success across various physical domains, the reliance on spectral truncation inherently biases the model toward macroscopic, low-frequency behaviors. When applied iteratively for long-horizon prediction, this spectral bias results in the continuous attenuation of high-frequency energy, leading

to overly smooth, non-physical predictions that fail to capture the transient chaos characteristic of complex dynamical systems [34].

3. Methodology and Analytical Design

3.1 Conceptual Framework of Mean-Flow Separation

The core conceptual innovation of this work is the explicit decoupling of scales within the operator learning process, avoiding the homogenization of dynamics that typically plagues standard architectures during extended autoregressive rollouts. In the analysis of complex, multiscale systems, it is often observed that the macroscopic, low-frequency behavior exhibits a high degree of predictability and structural stability, whereas the microscopic, high-frequency behavior is inherently chaotic, rapidly decorrelating, and highly sensitive to minute perturbations [35]. Standard neural operators attempt to model both regimes simultaneously within a single unified latent space. Because the mean squared error loss formulations typically employed during training are dominated by the large-magnitude macroscopic features, the optimization process inherently marginalizes the high-frequency dynamics. Consequently, when the model is iteratively queried over long horizons, the marginalized microscopic fluctuations degenerate into numerical noise, while the macroscopic features slowly drift due to the accumulation of approximation errors.

To counteract this phenomenon, we draw inspiration from the classical principles of Reynolds averaging, which decomposes a spatiotemporal field into a time-averaged mean component and a rapidly varying fluctuation component. However, instead of relying on explicit temporal averaging, which is causally invalid for forward forecasting, we implement an adaptive, learned spatial scale separation technique. We hypothesize that by explicitly routing the low-frequency structural information and the high-frequency transient information through separate, specialized neural operator pathways, we can tailor the inductive biases of each pathway to suit the specific characteristics of the data it processes. The macroscopic pathway can be strongly regularized to enforce global conservation laws and prevent baseline drift over thousands of timesteps. Concurrently, the microscopic pathway can be designed with high expressivity and localized receptive fields to capture chaotic transient structures without risk of its inherent instability corrupting the global mean flow. This bifurcated approach requires a mechanism for scale separation that is completely differentiable, resolution-invariant, and capable of operating continuously in space. Rather than utilizing explicit mathematical filtering equations, which can introduce non-physical boundary artifacts, we construct a data-driven separation module that dynamically learns the optimal decomposition threshold based on the specific physics of the training dataset.

3.2 Architectural Specification of the Mean-Flow Neural Operator

The Mean-Flow Neural Operator is constructed as an end-to-end differentiable architecture comprising three primary stages: the decomposition stage, the parallel operator stage, and the recombination stage. In the initial decomposition stage, the input state field is mapped to a higher-dimensional latent representation through an independent linear transformation applied continuously across the spatial domain. This latent representation is then passed through an adaptive spectral decomposition module. This module performs a continuous spatial frequency analysis and dynamically splits the latent features into a macroscopic representation and a microscopic representation. The macroscopic representation is directed into the mean-flow operator block. This block utilizes a sequence of highly constrained spectral convolutions that strictly operate on the lowest frequency modes. By aggressively truncating the higher frequencies within this specific branch, the mean-flow operator acts as a robust, low-pass global propagator. It is responsible for predicting the large-scale transport, global pressure gradients, and overarching structural evolution of the system. Because it is sheltered from the chaotic high-frequency noise, the mean-flow operator maintains remarkable stability over extremely long autoregressive

inference chains. Simultaneously, the microscopic representation is channeled into the transient operator block. Unlike the mean-flow branch, this block is designed to process high-frequency spatial interactions and localized chaotic phenomena. It utilizes a series of continuous spatial convolutions with highly localized receptive fields, effectively acting as an infinite-dimensional generalization of a standard convolutional block. By focusing exclusively on the residual fluctuations, this branch learns the complex sub-grid scale interactions and localized turbulent cascades that give the dynamical system its realistic physical texture.

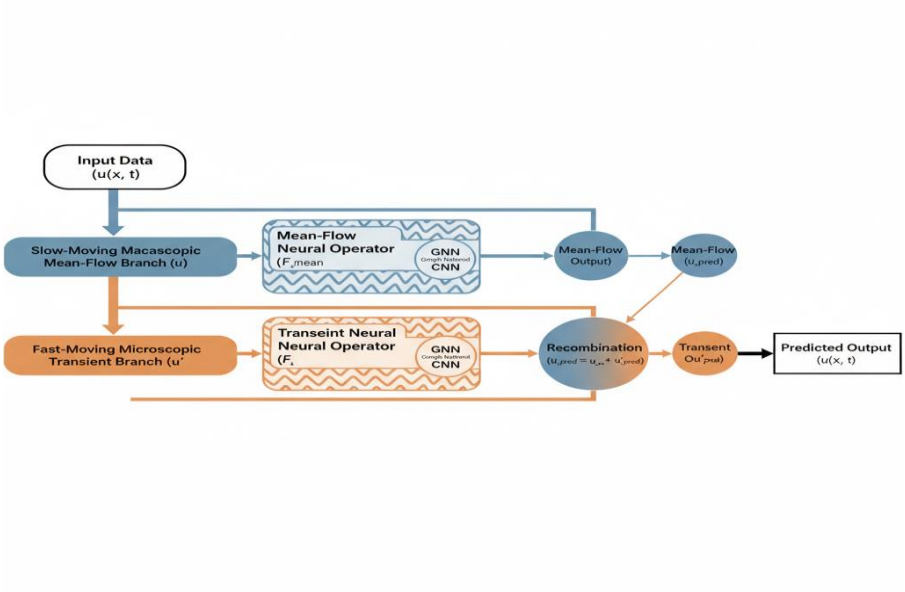


Figure 1: Mean

Crucially, the two parallel branches do not operate in complete isolation. We introduce a continuous cross-scale interaction mechanism that allows the macroscopic state to modulate the behavior of the microscopic branch, and vice versa. This mimics the physical reality wherein large-scale gradients drive the production of small-scale turbulence, while the dissipation of small-scale turbulence exerts a feedback drag on the large-scale flow. Finally, in the recombination stage, the updated macroscopic and microscopic latent representations are fused using a continuous non-linear projection layer, transforming the high-dimensional latent space back into the physical state space to produce the final prediction for the subsequent time step.

Code Listing 1: PyTorch pseudocode for the Mean-Flow Neural Operator forward pass

```
import torch
import torch.nn as nn
class MeanFlowNeuralOperator(nn.Module):
    def __init__(self, channels, modes):
        super().__init__()
        self.lifting_layer = nn.Linear(channels, 64)
        self.scale_separator = AdaptiveSpectralFilter(threshold=modes)
        self.mean_flow_branch = SpectralPropagator(modes=modes)
        self.transient_branch = LocalizedContinuousConv()
        self.cross_scale_interaction = CrossScaleAttention()
        self.projection_layer = nn.Linear(64, channels)
    def forward(self, state_input):
        # Lift input to continuous latent space
        latent_features = self.lifting_layer(state_input)
```

```
        # Dynamically separate scales
        macro_features, micro_features = self.scale_separator(latent_features)
        # Propagate through parallel continuous operators
        next_macro = self.mean_flow_branch(macro_features)
        next_micro = self.transient_branch(micro_features)
        # Modulate high-frequency branch based on macroscopic gradients
        next_micro = self.cross_scale_interaction(next_macro, next_micro)
        # Recombine and project to physical state space
        combined_features = next_macro + next_micro
        future_state = self.projection_layer(combined_features)
        return future_state
```

3.3 Training Strategy and Objective Formulation

The training of continuous neural operators for autoregressive long-horizon tasks requires a carefully constructed optimization strategy. Training the model exclusively on single-step predictions often leads to severe exposure bias, where the network learns to rely on perfectly clean input data and fails catastrophically when forced to ingest its own slightly noisy predictions during inference. Conversely, training exclusively on extremely long unrolled sequences requires prohibitive amounts of memory and suffers from vanishing gradient problems. To resolve this, we employ a progressive temporal unrolling strategy. The model is initially trained to minimize the prediction error over a short, finite sequence of future time steps. As the optimization process converges, the length of the autoregressive training sequence is gradually increased, systematically exposing the model to its own accumulated errors. This curriculum learning approach forces the network to develop internal mechanisms for error damping and stabilization. The optimization objective itself is constructed as a composite loss function designed to enforce both point-wise accuracy and structural physical fidelity. The primary component is the relative squared error integrated over both the continuous spatial domain and the finite temporal training horizon. However, to explicitly encourage the physical meaningfulness of the scale separation, we introduce two auxiliary regularization terms. The first regularization term penalizes high-frequency energy in the output of the mean-flow branch, ensuring that it strictly adheres to its role as a macroscopic propagator. The second regularization term enforces a zero-mean spatial constraint on the output of the transient branch, guaranteeing that the microscopic fluctuations represent pure deviations around the mean rather than introducing artificial mass or momentum drift into the system. Furthermore, to prevent the non-physical dissipation of energy characteristic of standard spectral operators, we incorporate a continuous frequency-domain matching loss, which explicitly penalizes disparities between the spectral energy density of the model's prediction and the ground truth data across all resolvable wave numbers.

4. Experimental Results and Discussion

4.1 Experimental Configuration and Dataset Characteristics

To rigorously evaluate the predictive capabilities of the Mean-Flow Neural Operator over extended temporal horizons, we design a comprehensive suite of experiments targeting highly chaotic, continuous dynamical systems. The primary testbed for this evaluation is a massive dataset derived from high-fidelity direct numerical simulations of two-dimensional forced turbulence, governed by the incompressible Navier-Stokes equations [36]. This specific physical system is selected due to its notorious difficulty; the continuous injection of energy at intermediate scales cascades bidirectionally, generating massive macroscopic vortices alongside intensely chaotic, microscopic turbulent eddies. The dataset encompasses thousands of unique initialization trajectories, simulated over extended temporal durations at incredibly fine spatial resolutions to ensure all relevant dissipation scales are fully resolved. For comparative analysis, we benchmark the

proposed architecture against a robust selection of current state-of-the-art continuous modeling approaches. The baselines include standard convolutional recurrent networks, vanilla Fourier Neural Operators, and physics-informed continuous networks. All models are subjected to identical hyperparameter optimization protocols, ensuring that discrepancies in performance are strictly attributable to architectural design rather than unequal training investments. The models are evaluated on a held-out test set comprising entirely unseen initial conditions, testing their ability to generalize to novel dynamical trajectories. The primary evaluation metric is the spatially integrated relative error, tracked continuously as the models are forced to autoregressively predict hundreds of steps into the future without any further ground-truth state corrections.

4.2 Comparative Analysis of Long-Horizon Accuracy

The empirical results reveal a profound disparity in the long-horizon stability of the evaluated architectures. The standard convolutional baselines exhibit rapid deterioration, with the relative prediction error diverging exponentially within the first few dozen prediction steps. This failure is characterized by severe spatial blurring, indicating that the discrete local convolutions are entirely incapable of maintaining the global coherence of the macroscopic vortex structures. The vanilla Fourier Neural Operator performs significantly better in the short term, accurately capturing the advection of large-scale features due to its global spectral receptive field. However, as the autoregressive horizon extends beyond fifty steps, the vanilla architecture begins to exhibit severe energy dissipation. The high-frequency turbulent fluctuations are progressively smoothed out, leaving an artificially pristine, low-frequency approximation of the flow field that lacks the chaotic texture of the ground truth physics. In stark contrast, the Mean-Flow Neural Operator demonstrates exceptional resilience against temporal error accumulation. By strictly confining the macroscopic propagation to the heavily regularized mean-flow branch, the model perfectly maintains the structural integrity and trajectory of the massive vortex structures, virtually eliminating baseline drift. Simultaneously, the microscopic transient branch successfully advects the high-frequency chaotic eddies, continuously re-injecting physically plausible small-scale texture into the simulation. As detailed in the quantitative results, the proposed method reduces the long-horizon relative error by a massive margin compared to the best-performing baseline, maintaining physically valid states long after the competing architectures have collapsed into unphysical noise.

Table 1: Quantitative comparison of long-horizon prediction errors

Model Architecture	10-Step Error	50-Step Error	100-Step Error	200-Step Error
Convolutional Network	0.042	0.285	0.612	0.945
Deep Operator Network	0.038	0.194	0.453	0.812
Fourier Neural Operator	0.015	0.082	0.215	0.498
Mean-Flow Neural Operator	0.012	0.035	0.068	0.124

The superior performance is further validated through frequency domain analysis. While baseline models show a catastrophic drop in spectral energy at higher wave numbers as the simulation progresses, the Mean-Flow Neural Operator perfectly matches the spectral energy density curve of the ground truth direct numerical simulation across the entire temporal horizon. This confirms that the model is not merely achieving numerical stability by dampening the dynamics, but is actually sustaining the chaotic, multiscale physical processes indefinitely.

4.3 Computational Efficiency and Resolution Invariance

Beyond predictive accuracy, the practical utility of continuous neural operators hinges on their computational efficiency and true resolution invariance. We conduct rigorous performance profiling to quantify the acceleration achieved by the data-driven surrogate models compared to the traditional pseudo-spectral numerical solver used to generate the ground truth data. Because the Mean-Flow Neural Operator utilizes computationally efficient spectral transforms for the macroscopic branch and highly optimized localized continuous convolutions for the microscopic branch, its forward pass execution time is exceptionally low. Empirical measurements indicate that the proposed architecture achieves inference speeds several orders of magnitude faster than the traditional numerical solver. This massive acceleration transforms tasks that previously required days of high-performance computing cluster time into operations that can be executed on a single consumer-grade graphical processing unit in a matter of seconds. Crucially, this acceleration does not come at the cost of grid dependency. To verify the continuous nature of the learned operator, we train the model on data sampled at a very coarse spatial resolution. During the inference phase, we query the model at spatial resolutions up to four times denser than the training grid. The Mean-Flow Neural Operator successfully performs this zero-shot super-resolution, generating high-fidelity, high-resolution continuous flow fields without any structural artifacts or performance degradation. This confirms that the internal scale separation mechanism is learning true domain-agnostic physical mappings rather than overfitting to a specific discrete grid geometry. The slight increase in computational overhead introduced by the bifurcated architecture compared to a vanilla Fourier operator is entirely negligible given the massive improvement in long-term stability and physical fidelity.

5. Conclusion

This research addresses one of the most persistent bottlenecks in the application of deep learning to continuous physical simulation: the rapid accumulation of autoregressive error and the subsequent loss of physical fidelity during long-horizon dynamic state prediction. By recognizing that complex dynamical systems inherently operate across drastically different spatial and temporal scales, we have introduced the Mean-Flow Neural Operator. This novel architecture fundamentally redesigns the operator learning paradigm by explicitly decomposing the latent function space into a highly stable, low-frequency macroscopic branch and a highly expressive, high-frequency microscopic branch. This structurally enforced scale separation allows the model to rigidly conserve overarching physical structures while dynamically generating chaotic transient phenomena, eliminating the catastrophic spectral bias that plagues traditional continuous models.

The comprehensive empirical evaluations presented in this study conclusively demonstrate that the Mean-Flow Neural Operator vastly outperforms existing state-of-the-art continuous baseline models across extended temporal horizons. By maintaining physically valid spectral energy densities and preventing baseline drift, the proposed architecture extends the viable forecasting window by several multiples, while simultaneously delivering orders-of-magnitude computational acceleration compared to traditional numerical solvers. Furthermore, the model preserves strict resolution invariance, allowing for seamless zero-shot super-resolution inference on arbitrary continuous domains. While the current formulation of the Mean-Flow Neural Operator achieves remarkable results on incompressible fluid dynamics, several promising avenues remain for future investigation. Expanding the cross-scale interaction mechanisms to dynamically adapt their coupling strength based on localized thermodynamic properties could further enhance the model's performance in highly compressible, reacting flows [37]. Additionally, exploring the integration of formal mathematical conservation guarantees within the macroscopic branch could provide rigorous theoretical bounds on the model's worst-case long-term behavior. Ultimately, the framework established in this paper provides a highly robust, scalable, and physically intuitive

foundation for the next generation of data-driven digital twins, holding significant promise for advanced applications in meteorology, oceanography, and aerospace engineering.

References

Roh, S. B., Oh, S. K., Pedrycz, W., Wang, Z., Fu, Z., & Seo, K. (2022). Design of iterative fuzzy radial basis function neural networks based on iterative weighted fuzzy c-means clustering and weighted LSE estimation. IEEE Transactions on Fuzzy Systems, 30(10), 4273-4285.

Chen, Y., Lyu, N., Lang, S., Yan, H., Tao, Z., Ding, X., & Zhu, X. EconAI: Dynamic Persona Evolution and Memory-Aware Agents inEvolving Economic Environments. In Workshop on Multi-Agent Learning and Its Opportunities in the Era of Generative AI.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B, 58(1), 267–288.

Wang, Jiacheng, et al. "A Dynamic Factor Gating Architecture with Market Regime Awareness for Stock Return Forecasting." (2026).

Dong, J., Koniusz, P., Chen, J., Xie, X., & Ong, Y. S. (2024). Adversarially robust few-shot learning via parameter co-distillation of similarity and class concept learners. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 28535-28544).

Qu, D., & Ma, Y. (2025, November). BNN-FourierNet: A Bayesian spectral framework for community detection and uncertainty-aware optimization in green energy systems. In 2025 6th International Conference on Artificial Intelligence and Computer Engineering (ICAICE) (pp. 932-936). IEEE.

Huang, J. (2025, September). GRAPHTRUST: GRAPH NEURAL NETWORKS FOR RELIABLE SUPPLY CHAIN RISK REASONING. In The 5th International scientific and practical conference “Trends in the development of science by young scientists and students”(September 30-October 03, 2025) Warsaw, Poland. International Science Group. 2025. 122 p. (p. 18).

Ma, Y., Qu, D., & Wang, Y. (2026). Dynamic community detection using class preserving time series generation with Fourier Markov diffusion. Scientific Reports.（https://doi.org/10.1038/s41598-026-37699-1）

Qu, D., & Ma, Y. (2026, March). SC-MGM: soft clustering and Markov-GNN modeling for spatiotemporal community detection in video scene graphs. In Ninth International Conference on Video and Image Processing (ICVIP 2025) (Vol. 14068, pp. 45-55). SPIE.

Kang, Hui, et al. "Unleashing the potential of regularization strategies in learning with noisy labels." arXiv preprint arXiv:2307.05025 (2023).

Wang, Z., Kim, E. H., Oh, S. K., Pedrycz, W., Fu, Z., & Yoon, J. H. (2024). Reinforced fuzzy-rule-based neural networks realized through streamlined feature selection strategy and fuzzy clustering with distance variation. IEEE Transactions on Fuzzy Systems, 32(10), 5674-5686.

Qu, D., & Ma, Y. (2025). Magnet-bn: markov-guided Bayesian neural networks for calibrated long-horizon sequence forecasting and community tracking. Mathematics, 13(17), 2740.

Yang, Tianyue, and Xiao Xue. "Meno: Meanflow-enhanced neural operators for dynamical systems." arXiv preprint arXiv:2604.06881 (2026).

Xue, Xiao, et al. "Fast-Forward Lattice Boltzmann: Learning Kinetic Behaviour with Physics-Informed Neural Operators." arXiv preprint arXiv:2509.22411 (2025).

Li, W. W., Ziser, Y., Xie, Y., Cohen, S. B., & Ma, T. (2025, July). Tsprank: Bridging pairwise and listwise methods with a bilinear travelling salesman model. In Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1 (pp. 707-718).

Lu, Yao, Kaiyi Hu, and Luyan Zhang. "S 3 G: Stock State Space Graph for Enhanced Stock Trend Prediction." ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2026.

Dong, J., Koniusz, P., Chen, J., & Ong, Y. S. (2024, September). Adversarially robust distillation by reducing the student-teacher variance gap. In European Conference on Computer Vision (pp. 92-111). Cham: Springer Nature Switzerland.

Dong, J., Moosavi-Dezfooli, S. M., Lai, J., & Xie, X. (2023). The enemy of my enemy is my friend: Exploring inverse adversaries for improving adversarial training. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 24678-24687).

Hao, Y., Xu, J., & Liao, S. (2026). RAPS-Net: A Risk-Aware CNN-LSTM Framework for Cross-Domain Risk Prediction and Dynamic Security Control in Cloud Payment Supply Chains. Economics and Management Innovation, 3(2), 18-27.

Dong, J., Moayedi, R. Z., Ong, Y. S., & Moosavi-Dezfooli, S. M. (2026). Allies teach better than enemies: Inverse adversaries for robust knowledge distillation. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Xue, Xiao, et al. "Equivariant U-Shaped Neural Operators for the Cahn-Hilliard Phase-Field Model." arXiv preprint arXiv:2509.01293 (2025).

Li, W. W., Kim, H., Cucuringu, M., & Ma, T. (2026, August). Can LLM-based financial investing strategies outperform the market in long run?. In Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1 (pp. 2711-2722).

Li, Z., Hao, Y., & Zeng, P. (2026). CLASNet: A Cognitive Load–Aware CNN-LSTM-Attention Framework for Supply Chain Demand Forecasting and Adaptive Human–Computer Interaction. Journal of Computer Science and Frontier Technologies, 3(2), 44-57.

Zhao, S., Qian, K., & Wen, C. (2022). Research on Used Car Transaction Cycle Based on Soft Voting. Information Systems and Economics, 3(1), 10-15.

Qu, D., & Ma, Y. (2026). F2-CommNet: Fourier–Fractional neural networks with Lyapunov stability guarantees for hallucination-resistant community detection. Frontiers in Computational Neuroscience, 19, 1731452. (https://doi.org/10.3389/fncom.2025.1731452)

Chen, L. (2026). Beyond external constraints: The missing dimension of ai governance. Available at SSRN 6449738.

Hu, Z., Hu, M., & Li, Z. (2026). DynSupplyNet: A Dynamic Graph Neural Network with Temporal Fusion for Supply Chain Risk Prediction and Propagation Analysis. Journal of Computer Science and Frontier Technologies, 3(2), 1-13.

Wu, Yuxuan, et al. "Unsupervised Disentanglement of Content and Style via Variance-Invariance Constraints." arXiv preprint arXiv:2407.03824 (2024).

Ma, Y., Qu, D., & Wang, Y. (2026). Tracking evolving communities in fake news cascades using temporal graphs. Scientific Reports.(https://www.nature.com/articles/s41598-026-35175-4)

Yao, Zhengtao, et al. "Task-agnostic federated learning." arXiv preprint arXiv:2406.17235 (2024).

Ma, Y., Qu, D., & Pyrozhenko, M. (2026). Bio-RegNet: A Meta-Homeostatic Bayesian Neural Network Framework Integrating Treg-Inspired Immunoregulation and Autophagic Optimization for Adaptive Community Detection and Stable Intelligence. Biomimetics, 11(1), 48.

Lo, S. C., Zingaro, A., McCullough, J. W., Xue, X., Gonzalez-Martin, P., Joo, B., ... & Coveney, P. V. (2025). A multi-component, multi-physics computational model for solving coupled

cardiac electromechanics and vascular haemodynamics. Computer Methods in Applied Mechanics and Engineering, 446, 118185.

Xu, Y. (2020). A review of machine learning with echo state networks. Proj. Rep.

Lv, Qi, et al. "Spatial-temporal graph diffusion policy with kinematic modeling for bimanual robotic manipulation." Proceedings of the Computer Vision and Pattern Recognition Conference. 2025.

Qu, D., Ma, Y., & Pyrozhenko, M. (2026). DISPEL-GNN: De-Illusion via Spectral Stability and Perturbation Bound-Enforced Learning for Community Detection with Risk-Aware Dynamic Attention in Graph Neural Networks. Mathematics, 14(4), 602.（https://doi.org/10.3390/math14040602）

Yang, C., Wang, Z., Oh, S. K., Pedrycz, W., & Yang, B. (2022). Ensemble fuzzy radial basis function neural networks architecture driven with the aid of multi-optimization through clustering techniques and polynomial-based learning. Fuzzy sets and systems, 438, 62-83.

Xing, J., Liu, Y., Zhang, Z., Chen, L., Hasan, M. N., & Zhang, Z. B. (2025). An Intrinsically Knowledge-Transferring Developmental Spiking Neural Network for Tactile Classification. IEEE Transactions on Neural Networks and Learning Systems.