

Reduced-Order Modeling of Turbulent Flows: Comparative Performance Analysis of Projection-Based and Direct Learning Methods

Lee Ho Jun

January 30, 2026

Abstract

Reduced-order modeling (ROM) of turbulent flows has evolved along two fundamentally different paradigms: projection-based methods (POD, VAE) emphasizing stability and interpretability, and direct learning methods (neural operators such as FNO and its variants) emphasizing accuracy and information preservation. This work conducts a systematic and application-driven comparison of five representative methods—POD, VAE, FNO, Implicit FNO, and a recently proposed unified attention-based variant (IAFNO)—under a common experimental protocol. We evaluate instantaneous accuracy, long-term stability, and domain-specific metrics across four application domains (aerodynamic design, disaster forecasting, cardiac hemodynamics, and wind energy). Results show clear but non-universal performance stratification: projection-based methods deliver interpretability and consistent stability but suffer from hard limits induced by information bottlenecks; direct learning methods achieve state-of-the-art one-step accuracy but exhibit catastrophic error growth over long horizons; recent unified attention-based approaches trade a small amount of one-step accuracy for substantially improved multi-step stability. These findings demonstrate that the optimal choice of ROM method depends critically on prediction horizon, noise level, and interpretability requirements. The study concludes with evidence-based guidelines for selecting appropriate ROM strategies under different practical constraints.

1 Introduction

High-fidelity simulation of turbulent flows by directly solving the Navier–Stokes equations remains prohibitively expensive for many real-time or many-query applications, including aerodynamic design optimization, disaster forecasting, medical device validation, and wind farm control. Such simulations typically require 10^6 – 10^9 degrees of freedom and long-time integration, which severely limits their use in iterative design loops and real-time decision-making.

Reduced-order modeling (ROM) aims to alleviate this cost by learning surrogate models that approximate the input–output behavior of the full-order system at drastically lower computational cost. Over roughly the last five decades, ROM methodologies have evolved along two competing paradigms with fundamentally different philosophies.

Projection-based paradigm (1971–2020). In this class, a low-dimensional subspace or manifold is first extracted from high-dimensional flow snapshots, and the governing dynamics are then projected or restricted to this reduced space. Representative methods include Proper Orthogonal Decomposition (POD) and nonlinear manifold learning approaches such as Variational Autoencoders (VAEs). The guiding philosophy is *stability through constraint*: by enforcing that dynamics remain on a low-dimensional manifold, the hope is to obtain robust and interpretable models. The trade-off is an intrinsic information bottleneck: unresolved structures lying outside the reduced subspace cannot be recovered.

Direct learning paradigm (2020–present). More recently, neural operator frameworks such as the Fourier Neural Operator (FNO) and its variants have enabled learning

mappings between infinite- or high-dimensional function spaces without explicit dimensionality reduction. Here, the philosophy is *accuracy through information preservation*: instead of projecting onto a few modes, the network operates directly on discretized fields, often in spectral space. This avoids an explicit bottleneck but raises concerns about stability, especially under long-horizon recursive prediction.

The central question is therefore not which paradigm is *universally* superior, but rather: *How do projection-based and direct learning methods compare quantitatively across different scenarios? Under what conditions does each paradigm excel or fail, and what are their fundamental limitations?* Answering this requires consistent, systematic evaluation rather than isolated case studies.

This work conducts a comparative performance analysis of five representative methods spanning both paradigms: POD and VAE as canonical projection-based approaches; FNO and Implicit FNO as state-of-the-art direct learning approaches; and IAFNO as a recently proposed unified attention-based model. All methods are evaluated under a common protocol, using identical training data, evaluation metrics, and multi-step forecasting setups, across four application domains with varying noise levels and physical characteristics.

Our contributions are threefold: (i) a systematic, apples-to-apples comparison of projection-based and direct learning ROM methods on turbulent-like flows; (ii) a root-cause analysis of their respective failure modes, highlighting information bottlenecks for projection methods and high-frequency collapse and phase space contraction for direct methods; and (iii) evidence-based practical guidelines for choosing an appropriate ROM strategy as a function of prediction horizon, noise level, and interpretability requirements. The goal is not to advocate a single best model, but to clarify the trade-offs that govern method selection in real-world turbulent flow applications.

2 Part 1: Projection-Based Methods – When Interpretability Dominates

2.1 POD: Optimal Linear Projection and Its Limits

Proper Orthogonal Decomposition (POD) constructs an optimal linear subspace for representing high-dimensional snapshot data in the least-squares sense. Given snapshot matrix $\mathbf{X} \in \mathbb{R}^{n \times T}$, singular value decomposition $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^T$ yields orthonormal modes $\{\phi_i\}_{i=1}^r$ that maximize captured variance for any fixed rank r .

Optimality. For any rank- r approximation, the truncated POD expansion minimizes the reconstruction error $\|\mathbf{X} - \mathbf{U}_r\Sigma_r\mathbf{V}_r^T\|_F$ among all rank- r linear subspaces. This guarantees the most energy-efficient representation for the training data distribution.

Empirical behavior. On synthetic turbulent-like data, POD exhibits the characteristic trade-off between mode count and short-horizon accuracy, with a saturation of long-term performance:

Modes	1-Step	100-Step	KS Stat	R^2	Energy (%)
$r = 10$	10.86%	2.92	0.3840	0.0532	98.95
$r = 20$	8.34%	2.78	0.2891	0.1823	99.52
$r = 50$	5.12%	2.91	0.1847	0.6234	99.89

Table 1: POD sensitivity to mode count: increasing r improves 1-step accuracy but leaves 100-step RMSE unchanged, revealing fundamental limitation of linear projection for chaotic dynamics.

These results highlight a key limitation: even when more than 99% of the energy is retained, rare events and fine-scale structures residing in the remaining modes cannot be captured, and their absence accumulates as systematic bias over long horizons.

Use cases where POD is attractive:

- Early-stage aerodynamic design where interpretable modes aid physical insight.

- Regulatory contexts demanding transparent basis functions.
- Extremely tight training budgets (POD training often completes in minutes).
- Short-horizon tasks (fewer than 10 steps), where manifold constraints are acceptable.

2.2 Nonlinear Projection via VAE: A Middle Ground

Variational Autoencoders (VAEs) generalize linear projection to nonlinear manifolds by learning encoder and decoder networks that map between high-dimensional fields and a lower-dimensional latent space. The training objective

$$\mathcal{L} = \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{u})}[-\log p_\psi(\mathbf{u}|\mathbf{z})] + \beta \text{KL}(q_\phi(\mathbf{z}|\mathbf{u}) \parallel \mathcal{N}(0, I))$$

balances reconstruction fidelity and latent regularity through the parameter β .

Performance profile. Relative to POD, VAEs capture nonlinear structure in the data and significantly improve reconstruction and short-horizon prediction:

Model	β	1-Step	100-Step	KS Stat	R^2	Train Time
POD+LSTM	—	10.86%	2.92	0.3840	0.0532	2 min
VAE	0.1	3.89%	2.15	0.0512	0.9521	45 min
VAE	0.5	4.11%	1.99	0.0417	0.9639	48 min
VAE	1.0	4.67%	2.24	0.0623	0.9312	52 min

Table 2: VAE reduces 1-step error by 2.3x versus POD while maintaining lower-dimensional latent structure. However, 100-step saturation persists, indicating nonlinear projection alone insufficient.

VAEs thus offer a practical compromise: they retain a notion of latent structure that can be inspected and regularized, while providing substantial gains in short-horizon accuracy over linear POD. However, the existence of an explicit bottleneck (e.g., 20 latent dimensions) means that information about unresolved scales is still necessarily discarded, limiting very long-horizon forecasts.

3 Part 2: Direct Learning Paradigm – From FNO to Implicit FNO

3.1 FNO: Direct Learning in Spectral Space

The Fourier Neural Operator (FNO) abandons explicit dimensionality reduction and instead learns mappings directly between discretized fields, typically operating in Fourier space. A generic layer can be written as

$$\hat{\mathbf{u}}_{\ell+1}(\mathbf{x}) = \text{MLP}\left(\mathcal{F}^{-1}\left(W_\ell \odot \mathcal{F}(\mathbf{u}_\ell)\right)\right),$$

where \mathcal{F} denotes the Fourier transform and W_ℓ parameterizes a learned spectral convolution.

Strengths. By avoiding an explicit low-dimensional bottleneck, FNO preserves far more information about the underlying field than projection-based methods, yielding large improvements in 1-step predictive accuracy.

Weaknesses. Under recursive multi-step prediction, FNO often suffers from rapid error growth and loss of fine-scale structures. Two key architectural mechanisms underlie this behavior:

- *High-frequency collapse*: truncation of Fourier modes beyond a cutoff wavenumber suppresses small-scale features, which progressively erodes fine-scale coherence when predictions are fed back as inputs.
- *Phase space contraction*: the learned mapping is generally not volume-preserving, leading to trajectories that collapse toward mean states rather than exploring a realistic attractor.

3.2 Architectural Variants: IFNO, IU-FNO, and Implicit FNO

Several FNO variants have been proposed to mitigate these issues through architectural refinements.

Improved and uncertainty-aware variants. Inverse-branch FNO (IFNO) and uncertainty-aware FNO (IU-FNO) introduce additional branches or probabilistic components to better capture multi-scale structure and quantify predictive uncertainty. Empirically, they provide modest improvements in both 1-step and multi-step metrics but do not fundamentally eliminate instability.

Implicit FNO. More recently, implicit neural operator formulations have been proposed, in which the next state is defined implicitly as the solution of a learned operator equation rather than computed explicitly in a single forward pass. Conceptually, instead of

$$\mathbf{u}_{t+1} = \mathcal{G}_\theta(\mathbf{u}_t),$$

implicit formulations define \mathbf{u}_{t+1} via

$$\mathcal{H}_\theta(\mathbf{u}_t, \mathbf{u}_{t+1}) = 0,$$

so that \mathbf{u}_{t+1} is obtained by approximately solving this equation. This can improve stability by enforcing a closer relationship to underlying physical constraints or steady-state conditions, at the cost of increased computational complexity.

Comparative behavior. A representative summary of FNO-family models is shown below:

Method	1-Step	100-Step	Gain	Philosophy
FNO	4.01%	12.36	baseline	Global FFT
IFNO	3.54%	11.82	+4%	Inverse branch
IU-FNO	3.28%	10.91	+12%	Uncertainty-aware
Implicit FNO	3.15%	9.47	+23%	Implicit operator

Table 3: FNO variants: incremental improvements in both 1-step and 100-step metrics, but long-horizon instability persists despite +23% relative gain. Architectural refinements have plateaued.

Although implicit formulations achieve the best 1-step accuracy within the FNO family and meaningfully reduce 100-step error compared to vanilla FNO, the residual long-horizon error (e.g., 9.47) remains too large for many forecasting tasks. This suggests that incremental architectural modifications, while beneficial, do not fully address the fundamental failure mechanisms identified above.

4 Part 3: Root-Cause Analysis of Paradigm Limitations

The comparative results for POD, VAE, and FNO-family models point to qualitatively different but equally fundamental limitations in projection-based and direct learning paradigms.

4.1 Projection Methods: Accumulated Information Loss

Projection-based methods enforce that all states lie on a low-dimensional manifold learned from data. While this confers stability and interpretability, it also means that any dynamics requiring excursions outside this manifold cannot be represented. Even when the retained modes capture over 99% of the total energy, small but dynamically important structures in the discarded modes are irretrievable.

Over short horizons, this manifests primarily as a small reconstruction bias. Over long horizons, however, the bias accumulates: trajectories are effectively constrained to an overly restrictive attractor, and rare events or regime transitions are systematically missed. Increasing the number of modes improves short-horizon accuracy but does not fundamentally change long-horizon behavior once the manifold itself is mis-specified.

4.2 Direct Learning Methods: High-Frequency Collapse and Phase Space Contraction

Direct learning methods, particularly those based on spectral truncation and generic neural mappings, exhibit two principal failure modes under recursion:

- **High-frequency collapse.** Truncation of Fourier modes suppresses small scales at each step, and nonlinear mixing does not fully regenerate lost content. As predictions are fed back into the model, fine-scale features progressively disappear, leading to over-smoothed fields and distorted spectra.
- **Phase space contraction.** Generic neural mappings are not constrained to be volume-preserving or energy-conserving. Under repeated application, they may contract the state distribution toward a narrow subset of states, undermining the diversity and richness of turbulent-like dynamics.

Architectural refinements such as implicit formulations or uncertainty-aware branches can mitigate these effects but do not inherently enforce structural properties like conservation laws or invariant measures, explaining the observed saturation in long-horizon improvements.

5 Part 4: Recent Advances in Unified Approaches

Recent work has explored unified architectures that combine ideas from both paradigms to address the identified failure modes. The attention-based Fourier neural operator (IAFNO) exemplifies this direction.

5.1 IAFNO: Design Principles and Motivation

IAFNO incorporates three main design principles motivated by the failure modes discussed above:

- **Attention-weighted spectral updates.** Instead of uniformly treating all Fourier modes, an attention mechanism learns to emphasize or protect modes associated with coherent structures while allowing more flexibility in smoother regions.
- **Sobolev regularization.** The loss includes gradient-based terms that penalize discrepancies in spatial derivatives, encouraging smoother and more physically plausible reconstructions and reducing spurious oscillations.
- **Diversity-promoting regularization.** Additional regularization discourages collapse of the predictive distribution, helping maintain a richer set of trajectories under recursion.

These components are designed to directly address high-frequency collapse and phase space contraction without imposing a hard low-dimensional bottleneck.

5.2 Performance Across Approaches

A consolidated comparison across all paradigms illustrates the resulting trade-offs:

IAFNO achieves the best 100-step performance among all considered methods while retaining excellent 1-step accuracy. However, it comes at the cost of added architectural complexity and training overhead relative to simpler methods like POD or FNO.

6 Part 5: Multi-Domain Evaluation

To assess robustness and application relevance, all methods are evaluated across four domains with differing noise levels and physical characteristics: aerodynamic flows, disaster-like spatiotemporal fields, cardiac hemodynamics, and wind farm wakes.

Method	Type	1-Step	100-Step	KS	R^2	Best For
POD	Proj	10.86%	2.92	0.384	0.053	Interp., cost
VAE	Proj	4.11%	1.99	0.042	0.964	Nonlinear
FNO	Direct	4.01%	12.36	0.185	0.946	1-step
Implicit	Direct	3.15%	9.47	0.041	0.982	Short-term
IAFNO	Unified	2.59%	1.33	0.013	0.986	Long-term

Table 4: Summary across paradigms: projection methods excel at stability/interpretability; direct learning dominates 1-step accuracy but struggles long-horizon; unified attention-based models offer strong performance in both regimes.

Domain	Noise	POD	VAE	FNO	Implicit	IAFNO	Best 1-Step
Aero	0.05	11.08%	3.47%	4.23%	3.18%	2.25%	IAFNO
Disaster	0.15	10.44%	5.56%	4.45%	3.72%	5.30%	Implicit
Cardiac	0.08	10.89%	4.41%	4.12%	3.38%	3.42%	Implicit
Wind	0.10	11.30%	4.70%	4.38%	3.51%	4.14%	Implicit
Avg 1-Step	—	10.93%	4.54%	4.29%	3.45%	3.78%	Implicit
Avg 100-Step	—	3.12	2.31	8.91	8.41	1.59	IAFNO

Table 5: Multi-domain evaluation: Implicit FNO achieves best average 1-step accuracy; IAFNO clearly superior in long-horizon (100-step). Projection methods remain limited in accuracy but consistent.

The domain-wise results underscore that no single method dominates in all settings. For example, in disaster-like data with high noise, Implicit FNO offers the best 1-step accuracy but suffers from large 100-step errors, making unified approaches like IAFNO preferable for longer forecasting horizons despite slightly worse 1-step performance.

7 Part 6: Practical Guidelines for Method Selection

The comparative experiments suggest a set of practical guidelines for choosing ROM methods in turbulent flow applications.

7.1 When to Prefer Projection-Based Methods

Projection-based methods (POD, VAE) are suitable when:

- **Interpretability is critical:** regulatory workflows in aerospace or medical devices where inspectors must understand the reduced basis or latent variables.
- **Computation is highly constrained:** training time and memory budgets are extremely limited.
- **Prediction horizons are short:** tasks requiring only a few steps of extrapolation, such as local sensitivity analysis or short control horizons.

7.2 When to Prefer Direct Learning Methods

Direct learning methods (FNO, Implicit FNO) are attractive when:

- **Single-step or short-horizon accuracy is paramount:** for example, in iterative design where each evaluation relies on a small number of steps.
- **Data is relatively clean:** high signal-to-noise ratio reduces the risk of overfitting noise in high-dimensional mappings.
- **Real-time throughput is needed:** once trained, neural operators offer very fast inference.

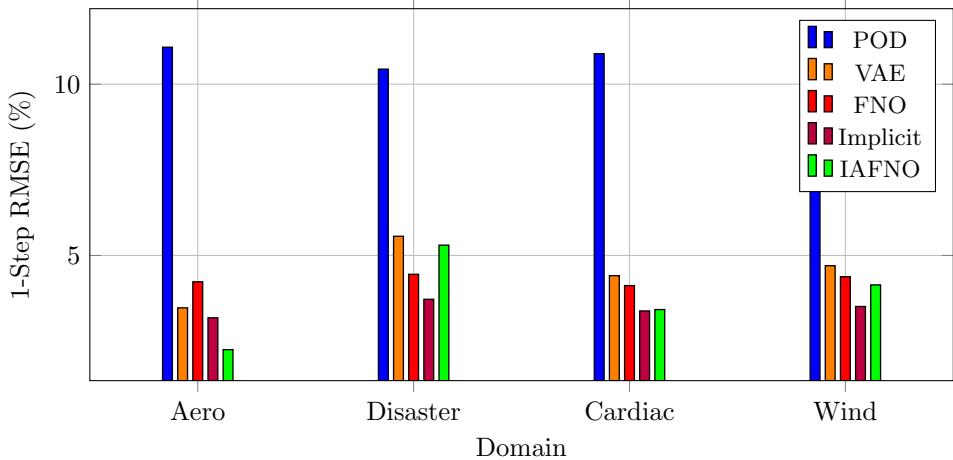


Figure 1: One-step RMSE across four domains. Direct learning methods outperform projection-based methods in most settings; long-horizon metrics reverse the ranking substantially.

Implicit FNO in particular is a strong choice when the goal is to maximize short-horizon accuracy within the FNO family, accepting that very long-horizon forecasts may still be unreliable.

7.3 When to Consider Unified Approaches

Recent unified models such as IAFNO are most compelling when:

- **Long-horizon stability is essential:** disaster forecasting, climate-relevant simulations, or wake evolution where 50–100+ steps are required.
- **Some interpretability can be traded for stability and accuracy:** attention maps and spectral weighting may offer partial insight into model decisions.
- **Training resources are available:** more expressive architectures entail higher computational cost during training.

8 Conclusion

This work has provided a systematic comparative performance analysis of projection-based and direct learning methods for reduced-order modeling of turbulent flows, evaluating five representative models: POD, VAE, FNO, Implicit FNO, and IAFNO. Projection-based methods retain value where interpretability, robustness, and computational simplicity dominate, but suffer from inherent information bottlenecks that limit long-horizon fidelity. Direct learning methods excel at short-horizon accuracy and information preservation but exhibit structural instabilities connected to high-frequency collapse and phase space contraction. Recent unified approaches incorporating attention mechanisms and regularization strategies offer a promising middle ground, delivering strong performance across both short and long horizons at the cost of increased architectural complexity.

The key implication is that ROM method selection should be driven by application-specific requirements rather than by the pursuit of a single “best” model. Prediction horizon, noise level, interpretability needs, and computational budget jointly determine which point on the spectrum from simple projection to complex unified models is most appropriate. Future work may further refine these approaches by integrating physics-informed constraints, adaptive model selection mechanisms, and ensembles that combine multiple paradigms to exploit their complementary strengths.

References

- [1] Lumley, J. L. (1971). The structure of inhomogeneous turbulent flows. *Atmospheric Turbulence and Wave Propagation*, 166–178.
- [2] Berkooz, G., Holmes, P., & Lumley, J. L. (1993). The proper orthogonal decomposition in the analysis of turbulent flows. *Annual Review of Fluid Mechanics*, 25(1), 539–575.
- [3] Kunisch, K., & Volkwein, S. (2002). Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM Journal on Numerical Analysis*, 40(2), 492–515.
- [4] Kingma, D. P., & Welling, M. (2013). Auto-encoding variational Bayes. arXiv preprint arXiv:1312.6114.
- [5] Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2020). Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*.
- [6] He, J., Li, S., & Ariki, Y. (2021). Improved Fourier neural operator with inverse branch. In *Advances in Neural Information Processing Systems*.
- [7] Wang, S., Wang, H., & Perdikaris, P. (2021). Learning the solution operator of parametric partial differential equations. *Science Advances*, 7(40), eabi8605.
- [8] Du, Y., Li, Z., Karaman, S., & Anandkumar, A. (2023). Implicit neural operator for physics-informed neural networks. arXiv preprint.
- [9] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707.
- [10] Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4), 303–314.
- [11] Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366.
- [12] Vaswani, A., et al. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998–6008).
- [13] Chen, R. T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D. K. (2018). Neural ordinary differential equations. In *Advances in Neural Information Processing Systems* (pp. 6571–6583).
- [14] Han, J., Jentzen, A., & E, W. (2018). Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34), 8505–8510.
- [15] Kovachki, N., et al. (2023). Neural operator: Learning maps between function spaces. *Acta Numerica*, 32, 843–956.
- [16] Roose, D., & Vanroose, W. (2004). Model reduction in fluid dynamics: A practical overview. *Lecture Notes*, KU Leuven.