

FORECASTING VARIABLE-DENSITY 3D TURBULENT FLOW

Xingyu Su

Department of Computer Science
Zhejiang University
Hangzhou 310058, P.R. China
su_xingyu@zju.edu.cn

Robin Walters

Khoury College of Computer Sciences
Northeastern University
Boston, MA 02115, USA
rwalters@northeastern.edu

Denis Aslangil

Computational Physics and Methods (CCS-2)
Los Alamos National Lab
Los Alamos, NM 87545, USA
denis.aslangil@lanl.gov

Rose Yu

Department of Computer Science and Engineering
University of California, San Diego
La Jolla, CA 92093, USA
roseyu@eng.ucsd.edu

ABSTRACT

Deep learning has shown the potential to significantly accelerate numerical simulation of fluids without sacrificing accuracy, but prior works are limited to stationary flows with uniform density. In real-world engineering applications, turbulent flows are mostly three-dimensional, non-stationary and have variable-density. Here we propose Taylor-Net, a hybrid model that combines deep neural networks with numerical Taylor series method for 3D turbulent flow prediction. Across flows with different density-ratio, our method over 3 orders of magnitude faster than high-fidelity numerical simulations. It also achieves higher accuracy than several strong physics-informed deep learning baselines. Most importantly, the predictions of our Taylor-Net pertain consistent physical characteristics including mass conservation, and turbulent energy spectrum.

1 INTRODUCTION

Modeling the evolution of fluid flow is critical to climate science, aerospace industry, combustion and bio-fluid applications, and far more. In most cases, fluid flows are observed in the turbulence phase in which the flow is composed of a wide range of spatial and temporal scales. Unfortunately, performing high-fidelity numerical simulations of turbulence, which requires extremely large resolutions, is computationally infeasible (Pope, 2000; Sagaut et al., 2013). As a result, turbulence-closure models are used to reduce the computational complexity of the simulation. These classical techniques, however, require hand-engineering, are specific to particular flows, and still often do not produce as accurate results as desired.

Recent approaches to accelerate fluid simulation use physics-informed deep learning: Wiewel et al. (2019); Xie et al. (2018); Wang et al. (2020a); Kochkov et al. (2021); Li et al. (2020); Mohan et al. (2020); Beucler et al. (2019); Wu et al. (2020); Mohan et al. (2019); Kim & Lee (2020); Fang et al. (2018); Font et al. (2021). However, current methods still rely on simplifying assumptions in Navier-Stokes equations to make the problem more tractable, such as working on two-dimensional data, studying stationary flows, using low-resolution data, or asserting uniform density. These assumptions dramatically limit the applicability of these methods to real-world turbulent flows.

In contrast, we tackle the general case of predicting three-dimensional, non-stationary variable-density (VD) turbulent flow used in engineering applications. In such flows, the statistical features of the turbulence are changing over time. Specifically, higher and lower-density fluids within the flow mix, and the turbulent kinetic energy of the flow decays. From a machine learning perspective, this causes covariate shift when splitting train-test data across time, which leads to poor generalization. In addition, in VD flows, the density field is considered as a variable rather than a constant and its behaviour needs to be captured separately than the velocity field.

We consider two type of homogeneous variable-density turbulence (HVDT) Aslangil et al. (2020c) flow with different density-ratio in response to buoyant forces. The input and output of our model is a 4-dimensional vector field consisting of 3 velocity components $\mathbf{v} = (u, v, w)$ and a density component ρ which vary in time and space. Since the flow evolution is highly coupled with the density and velocity fields the momentum $\mathbf{p} = \rho\mathbf{v}$ plays a larger role in the governing equations than velocity. Hence, we find that predicting momentum and density and then computing velocity $\mathbf{v} = \mathbf{p}/\rho$ is more stable than the predicting velocity directly.

Our model is a hybrid model which uses deep learning to learn the remainder of the Taylor series. We predict the future time steps of the flow with the n^{th} -order Taylor series approximation fit to the past data. The deep learning model takes as input past time steps of flow and the output is added to the prediction from Taylor approximation. We use a U-net encoder-decoder architecture shown to be effective for dynamics prediction (de Bezenac et al., 2018; Wang et al., 2020a;b). We find that learning the Taylor remainder can dramatically reduce the error of the prediction, compared to learning the future momentum and density directly.

In summary, our contributions include,

- Taylor-Net: first deep learning method to predict variable-density three-dimensional turbulent flows, by combining Taylor series approximation and U-net.
- Significantly improved accuracy and physical consistency over competitive baselines on more general (e.g. non-stationary and anisotropic VD turbulence) fluid dynamics,
- Theoretical analysis of the interplay between forecasting horizon, step size, and the order of Taylor approximation.

2 METHODS

2.1 BACKGROUND IN VARIABLE DENSITY TURBULENCE

Simulating 3D variable-density (VD) turbulent fluids has a broad impact in many important engineering applications such as in high-speed combustion, and high-energy-density processes like inertial confinement fusion (Aslangil et al., 2019; 2020b;c; Saenz et al., 2021). The readers are referred to the comprehensive review Livescu (2020) for more details.

The VD turbulence emerges from the mixing of two fluids with different densities. Using the Einstein summation convention, the governing equations of this physical system are:

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u_j}{\partial x_j} = 0, \quad \frac{\partial \rho u_i}{\partial t} + \frac{\partial \rho u_i u_j}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j} + \frac{\rho g_i}{Fr^2} \quad (1)$$

where ρ is the whole density field, u_i is the whole velocity field in direction i , p is the pressure, g_i is the gravity in direction i and the stress tensor is assumed Newtonian, $\tau_{ij} = (\rho/Re_0)(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\frac{\partial u_k}{\partial x_k}\delta_{ij})$. Unlike uniform density flows whose velocity is divergence-free, the velocity divergence in VD turbulence is none-zero and can be written as:

$$\frac{\partial u_i}{\partial x_i} = -\frac{1}{Re_0 Sc} \frac{\partial \rho}{\partial^2 x_i}. \quad (2)$$

The non-dimensional parameters are the computational Reynolds number, Re_0 , Schmidt number, Sc , and Froude number, Fr , defined in (Aslangil et al., 2020c). In addition, in VD mixing problems, the density contrast between the two mixing fluids is defined by an Atwood number $A = (\rho_h - \rho_l)/(\rho_h + \rho_l)$ where ρ_h and ρ_l are the heavy and light fluids density values, respectively. We generate two datasets using HVDT Direct Numerical Simulation (DNS) in a triply periodic domain $[(2\pi)^3]$ with a 64^3 resolution, with $A = 0.05$ and $A = 0.75$, whose Re_0 values are 100 and Sc and Fr are both unity. The prediction task is to forecast the future density and velocity field given past data sequences. The homogeneity property of HVDT means the flow is statistically space-translation invariant. It is also notable that HVDT flow is non-stationary. That is, in contrast to space translations, time translations do alter the flow. In HVDT, the spatial homogeneity allows us to isolate the VD mixing problem from layer edges and inhomogeneities. Aslangil et al.

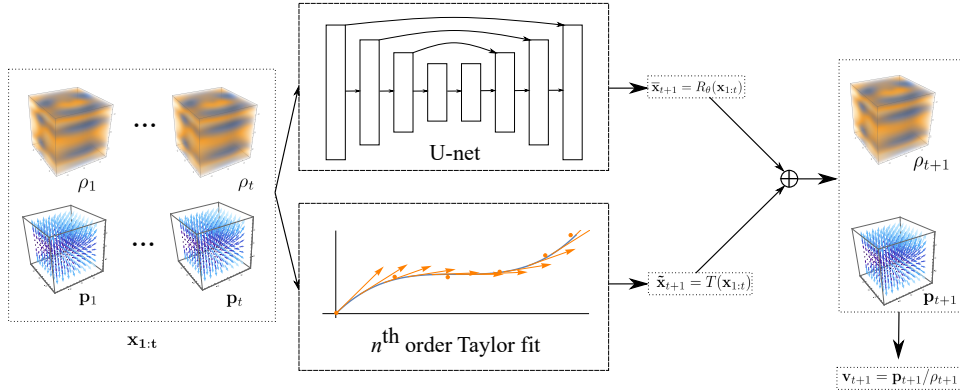


Figure 1: Diagram of the Taylor-Net- n method. First, the model computes the n^{th} -order polynomial fit, i.e., Taylor approximation with it uses to approximate $\tilde{\mathbf{x}}_{t+1} = T(\mathbf{x}_{1:t})$. The U-net predicts the Taylor remainder $\bar{\mathbf{x}}_{t+1} = R_{\theta}(\mathbf{x}_{1:t})$. For the output we take the sum $(\rho_{t+1}, \mathbf{p}_{t+1}) = \mathbf{x}_{t+1} = \bar{\mathbf{x}}_{t+1} + \tilde{\mathbf{x}}_{t+1}$ from which we compute the velocity $\mathbf{v}_{t+1} = \mathbf{p}_{t+1}/\rho_{t+1}$.

(2019; 2020c) has connected the idealized homogeneous VD flow to the real applications such as to the buoyancy-driven Rayleigh-Taylor Instability (Rayleigh, 1884; Taylor, 1950) under constant and variable-acceleration histories (Aslangil et al., 2016; 2020a), shock-driven Richtmyer–Meshkov Instability (Richtmyer, 1960; Meshkov, 1969), and VD mixing layers (Baltzer & Livescu, 2020).

2.2 TAYLOR-NET: LEARNING TAYLOR SERIES REMAINDER WITH DEEP NETS

We employ a hybrid method which combines Taylor approximation and a U-net to learn the Taylor remainder. Note that our method does not explicitly use any part of the Navier-Stokes equations. The high-level structure of our model is illustrated in Figure 1.

Taylor Approximation Given input data x_{-t}, \dots, x_0 , we can interpolate the n -th order polynomial fit to the data where $n \leq t$. Let $p_n(t) = a_0 + a_1 t + \dots + a_n t^n$. The coefficients of p_n can be determined by the inverse of the Vandermonde matrix

$$\begin{pmatrix} x_{-n} \\ \vdots \\ x_0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & n & \dots & n^n \end{pmatrix}^{-1} = \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix}$$

Then the n -th order Taylor approximation is

$$\hat{x}_1 = T_n(x_{-n}, \dots, x_0) = p_n(1) = \sum_{i=-n}^0 x_i \binom{n}{-i} (-1)^{i+1} \quad (3)$$

For $n = 2$, this yields $x_{-2} - 3x_{-1} + 3x_0$ and for $n = 1$ this gives $-x_{-1} + 2x_0$.

U-net The U-net encoder-decoder architecture has been shown to be effective for dynamics prediction (de Bezenac et al., 2018; Wang et al., 2020a;b). We use a 9 layer-U-net with 3 down-sampling convolutional layers with stride 2, followed by 2 convolutional layers which preserve the dimension, followed by 3 up-sampling transposed convolution layers which repeat the sizes of the down-sampling layers in repeat order. Layers of the same size are directly connected by skip connections. There is one final convolutional output layer which shrinks the channel dimension to the size of the output.

Density and Momentum Features Our goal is to predict both the density field ρ and velocity field \mathbf{v} . We find our model performs better when we use density and momentum \mathbf{p} as input and output features instead and compute $\mathbf{v} = \mathbf{p}/\rho$. We hypothesize this is because the momentum plays a larger roll in the defining equations (1). To solve these equations using DNS for velocity, it would be necessary to repeatedly multiply the density and velocity and then divide by density. These products and quotients are somewhat difficult for the neural network to approximate, and thus learning is improved by using momentum features.

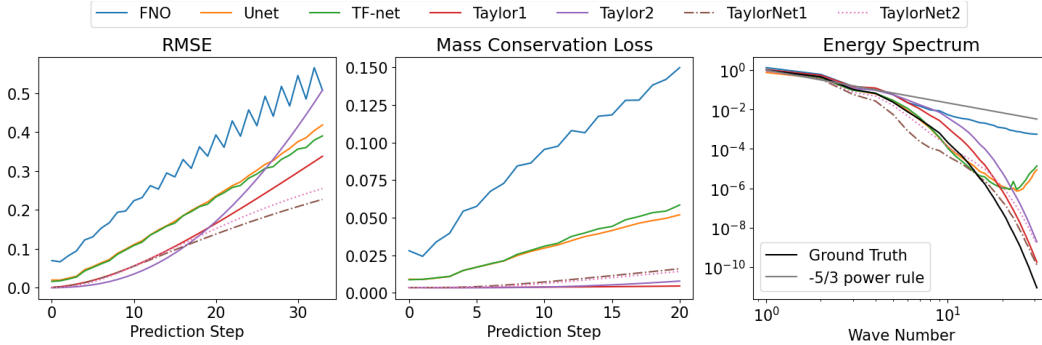


Figure 2: Comparison of our model versus several baselines across 3 metrics for \mathbf{p} prediction: RMSE, Mass Conservation, and Energy Spectrum. Our errors grow more slowly than others across the forecasting horizon. The flow predicted by our model has an energy spectrum closer to the truth.

Periodic Convolution Since the input and output spatial domain is periodic, we implement U-net using periodic 3D convolutions. Consider input tensor x of size 64^3 and kernel ϕ of size $(2c + 1)^3$. We index the kernel symmetrically about $(0, 0, 0)$. We define periodic convolution as

$$y_{m,n,p} = \sum_{i=-c,j=-c,k=-c}^{c,c,c} \phi_{i,j,k} x_{\overline{m-i}, \overline{n-j}, \overline{p-k}}$$

where we take the indices of x as $\overline{i} = i \bmod 64$. This is implemented using the “circular” padding mode of the PyTorch `conv3D` function.

Theoretical Analysis on Taylor Order Although higher order Taylor approximations achieve better fit over the short-term, they also diverge faster in the long term, a fact we see reflected in our experiments. Thus, the optimal order of n represents a trade-off. The proof is deferred to Supp.

Proposition 2.1. *Assume the true time series $|x_t| \leq C$ is bounded for all t . For almost all values x_{-n}, \dots, x_0 , the iterated n -th order Taylor progression $x_{i+1} = T_n(x_{i-n}, \dots, x_i)$ diverges and the rate of divergence is $\Theta(i^n)$.*

3 EXPERIMENTS

Our turbulent datasets contain flows with two different Atwood numbers. We benchmark the performance of different methods w.r.t three metrics: Root Mean Square Error (RMSE), Energy Spectrum Error (ESE) and Mass Conservation. We compare with SoTA baselines for turbulent flow prediction including U-Net, TF-Net (Wang et al., 2020a), and Fourier Neural Operator (FNO) (Li et al., 2020).

Table 1 shows the parameter counts and run times of different models on Dataset 1. Previous state-of-arts deep learning models consistently fail to outperform simple numerical approximations. Our combined method of numerical and deep learning has the best performance and shows a great improvement in comparison with numerical approximations.

Table 1: Parameters number and run time comparison

Models	DNS	Unet	TF-Net	FNO	TaylorNet
# params ($1e7$)	—	1.587	2.528	0.083	1.587
Runtime (s)	≈ 100	0.0407	0.0986	0.0478	0.0404

Figure 2 shows the performance comparison for (a) RMSE (b) Mass-Conservation and (c) energy spectrum at the 20th prediction step. As noted earlier, HVDT flow is a non-stationary flow even at later times, so the eddy turnover time is not a constant during the flow evolution. For comparison, the eddy turnover time at the beginning of the training data is ≈ 6 prediction steps, but at the end of the test data set, eddy turnover time increases to ≈ 30 prediction steps. Taylor-Net- n denotes n -th order Taylor-Net. Taylor-Net has lower RMSE compared to all other methods including classic Taylor expansions. In terms of Mass Conservation loss, Taylor-Net outperforms all ML models, but is not as good as Taylor expansion. For ESE based on total momentum, Taylor-Net’s largest advantage is in small-scale motion prediction (at large wave numbers) which is a great challenge for many ML turbulence models. Figure 3 visualizes the Taylor-Net1 predictions on Dataset 2 for 30

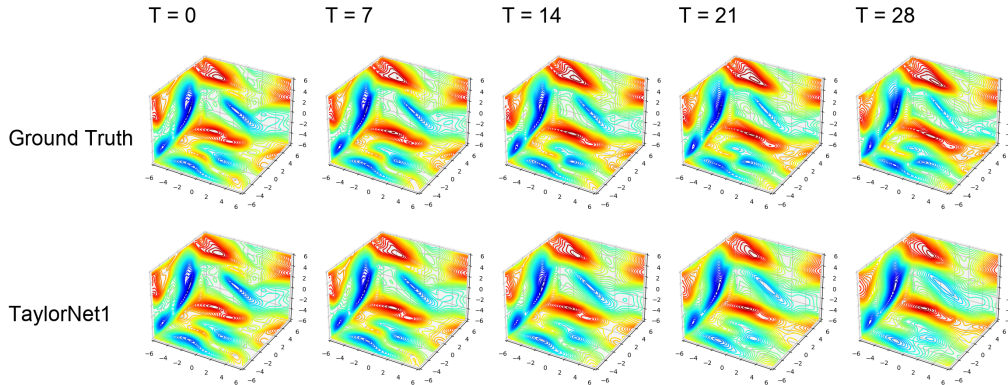


Figure 3: Prediction visualization for momentum field over 30 time steps ($\approx .2$ eddy-turnover times). Comparison between Taylor-Net1 and ground truth shows our hybrid method can capture the small scale dynamics.

steps ahead forecasting. For each frame, we pick only the momentum field along x -axis and cast 3 surfaces of the data cube to xy , yz , zx planes.

4 CONCLUSION

We apply deep learning for the first time to variable-density and non-stationary turbulent flow. This creates unique challenges for learning, which we address with a hybrid model which predicts the Taylor remainder of future time steps. We achieve higher accuracy than several strong baselines. A limitation of our model is that it does not show significant improvement for highly mixed or stationary flows. Future work includes larger resolution data and different Reynolds number flows.

5 ACKNOWLEDGEMENT

We thank Google for GCP Research Credits to help with this work. We thank Rui Wang, Arvind Mohan, Karthik Kashinath, and Mustafa Mustafa for helpful conversations. We thank Arvind Mohan for also sharing his data and code. We thank Zongyi Li for answering our questions regarding the Fourier Neural Operator baseline. Robin Walters was supported by the Roux Institute under the Experiential AI Fellowship. This work was supported in part by Google Faculty Research Award, NSF Grant #2037745, and the U. S. Army Research Office under Grant W911NF-20-1-0334. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

REFERENCES

- Denis Aslangil, Arindam Banerjee, and Andrew G. W. Lawrie. Numerical investigation of initial condition effects on Rayleigh-Taylor instability with acceleration reversals. *Physical Review E*, 94:053114, 2016.
- Denis Aslangil, Daniel Livescu, and Arindam Banerjee. Flow regimes in buoyancy-driven homogeneous variable-density turbulence. In Ramis Örlü, Alessandro Talamelli, Joachim Peinke, and Martin Oberlack (eds.), *Progress in Turbulence VIII*, pp. 235–240, Cham, 2019. Springer International Publishing.
- Denis Aslangil, Zachary Farley, Andrew G. W. Lawrie, and Arindam Banerjee. Rayleigh–Taylor Instability With Varying Periods of Zero Acceleration. *Journal of Fluids Engineering*, 142(12), 10 2020a. ISSN 0098-2202. doi: 10.1115/1.4048348. URL <https://doi.org/10.1115/1.4048348>. 121103.
- Denis Aslangil, Daniel Livescu, and Arindam Banerjee. Variable-density buoyancy-driven turbulence with asymmetric initial density distribution. *Physica D: Nonlinear Phenomena*, 406:132444,

- 2020b. ISSN 0167-2789. doi: <https://doi.org/10.1016/j.physd.2020.132444>. URL <http://www.sciencedirect.com/science/article/pii/S016727891930483X>.
- Denis Aslangil, Daniel Livescu, and Arindam Banerjee. Effects of atwood and reynolds numbers on the evolution of buoyancy-driven homogeneous variable-density turbulence. Journal of Fluid Mechanics, 895:A12, 2020c. doi: 10.1017/jfm.2020.268.
- Jon R. Baltzer and Daniel Livescu. Variable-density effects in incompressible non-buoyant shear-driven turbulent mixing layers. Journal of Fluid Mechanics, 900:A16, 2020. doi: 10.1017/jfm.2020.466.
- Tom Beucler, Michael Pritchard, Stephan Rasp, Jordan Ott, Pierre Baldi, and Pierre Gentine. Enforcing analytic constraints in neural-networks emulating physical systems. arXiv preprint arXiv:1909.00912, 2019.
- Emmanuel de Bezenac, Arthur Pajot, and Patrick Gallinari. Deep learning for physical processes: Incorporating prior scientific knowledge. In International Conference on Learning Representations, 2018. URL <https://openreview.net/forum?id=By4HsfWAZ>.
- Rui Fang, David Sondak, Pavlos Protopapas, and Sauro Succi. Deep learning for turbulent channel flow. arXiv preprint arXiv:1812.02241, 2018.
- Bernat Font, Gabriel D. Weymouth, Vinh-Tan Nguyen, and Owen R. Tutty. Deep learning of the spanwise-averaged navier–stokes equations. Journal of Computational Physics, 434:110199, 2021. ISSN 0021-9991. doi: <https://doi.org/10.1016/j.jcp.2021.110199>. URL <https://www.sciencedirect.com/science/article/pii/S0021999121000942>.
- Junhyuk Kim and Changhoon Lee. Deep unsupervised learning of turbulence for inflow generation at various reynolds numbers. Journal of Computational Physics, 406:109216, 2020.
- Dmitrii Kochkov, Jamie A Smith, Ayya Alieva, Qing Wang, Michael P Brenner, and Stephan Hoyer. Machine learning accelerated computational fluid dynamics. arXiv preprint arXiv:2102.01010, 2021.
- Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. arXiv preprint arXiv:2010.08895, 2020.
- Daniel Livescu. Turbulence with large thermal and compositional density variations. Annu. Rev. Fluid Mech., 52:309–341, 2020.
- Daniel Livescu and J. Raymond Ristorcelli. Buoyancy-driven variable-density turbulence. J. Fluid Mech., 591:43–71, 11 2007. doi: 10.1017/S0022112007008270.
- E. E. Meshkov. Instability of the interface of two gases accelerated by a shock wave. Fluid Dynamics, 4(5):101–104, Sep 1969. ISSN 1573-8507. doi: 10.1007/BF01015969. URL <https://doi.org/10.1007/BF01015969>.
- Arvind Mohan, Don Daniel, Michael Chertkov, and Daniel Livescu. Compressed convolutional lstm: An efficient deep learning framework to model high fidelity 3d turbulence. arXiv preprint arXiv:1903.00033, 2019.
- Arvind T Mohan, Nicholas Lubbers, Daniel Livescu, and Michael Chertkov. Embedding hard physical constraints in convolutional neural networks for 3d turbulence. In ICLR 2020 Workshop on Integration of Deep Neural Models and Differential Equations, 2020.
- S. B. Pope. Turbulent Flows. Cambridge University Press, 2000. ISBN 9780521598866. URL <https://books.google.com/books?id=HZsTw9SMx-0C>.
- Rayleigh. Investigation of the character of the equilibrium of an incompressible heavy fluid of variable density. Proceedings of the London Mathematical Society, s1-14(1):170–177, 1884. doi: 10.1112/plms/s1-14.1.170. URL <https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/plms/s1-14.1.170>.

- R. D. Richtmyer. Taylor instability in shock acceleration of compressible fluids. Communications on Pure and Applied Mathematics, 13(2):297–319, 1960. doi: 10.1002/cpa.3160130207. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.3160130207>.
- Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pp. 234–241. Springer, 2015.
- Juan A. Saenz, Denis Aslangil, and Daniel Livescu. Filtering, averaging, and scale dependency in homogeneous variable density turbulence. Physics of Fluids, 33(2):025115, 2021. doi: 10.1063/5.0040337. URL <https://doi.org/10.1063/5.0040337>.
- Pierre Sagaut, Sébastien Deck, and Marc Terracol. Multiscale and Multiresolution Approaches in Turbulence. IMPERIAL COLLEGE PRESS, 2nd edition, 2013. doi: 10.1142/p878. URL <https://www.worldscientific.com/doi/abs/10.1142/p878>.
- Taylor. The instability of liquid surfaces when accelerated in a direction perpendicular to their planes. i. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 201(1065):192–196, 1950. ISSN 0080-4630. doi: 10.1098/rspa.1950.0052. URL <http://rspa.royalsocietypublishing.org/content/201/1065/192>.
- Rui Wang, Karthik Kashinath, Mustafa Mustafa, Adrian Albert, and Rose Yu. Towards physics-informed deep learning for turbulent flow prediction. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 1457–1466, 2020a.
- Rui Wang, Robin Walters, and Rose Yu. Incorporating symmetry into deep dynamics models for improved generalization. arXiv preprint arXiv:2002.03061, 2020b.
- Steffen Wiewel, Moritz Becher, and Nils Thuerey. Latent space physics: Towards learning the temporal evolution of fluid flow. In Computer graphics forum, volume 38, pp. 71–82. Wiley Online Library, 2019.
- Jin-Long Wu, Karthik Kashinath, Adrian Albert, Dragos Chirila, Heng Xiao, et al. Enforcing statistical constraints in generative adversarial networks for modeling chaotic dynamical systems. Journal of Computational Physics, 406:109209, 2020.
- You Xie, Erik Franz, Mengyu Chu, and Nils Thuerey. tempogan: A temporally coherent, volumetric gan for super-resolution fluid flow. ACM Transactions on Graphics (TOG), 37(4):1–15, 2018.

A APPENDIX

A.1 THEOREM PROOFS

Proof. Define \hat{x}_i for $i \geq -n$ inductively as follows. Set $\hat{x}_i = x_i$ for $i < 0$. For $i \leq 0$, let p_i be the n -th order polynomial such that $p_i(i - k) = \hat{x}_{i-k}$ for $0 \leq k \leq n$. Then set $\hat{x}_{i+1} = p_i(i + 1) = T_n(\hat{x}_{i-n}, \dots, \hat{x}_i)$. Note that p_i is uniquely determined by any $n + 1$ points, and thus $p_i = p_{i+1}$. Thus denoting $p = p_i$ for all i , we have $\hat{x}_i = p(i)$. The lead term of $p(i)$ is i^n times a linear combination of x_{-n}, \dots, x_0 which is non-zero for almost all values of x_{-n}, \dots, x_0 . Since we assume the true values of the time series are bounded the error of polynomial expansion will grow asymptotically as i^n . \square

A.2 EXPERIMENT SET UP

Datasets The 3 datasets for our experiments are buoyancy-driven homogeneous variable-density turbulence (HVDT) simulated at different Atwood numbers. In each dataset, each frame has 4 channels, which are density and velocity along 3 axes. The DNS data is generated by using petascale variable-density version of the CFDNS code (Livescu & Ristorcelli, 2007; Aslangil et al., 2020c). The code uses the variable time step third order Adams-Bashforth-Moulton scheme, coupled with a fractional time method for time advancement, and the spatial derivatives are evaluated using Fourier transforms.

For each dataset, we use a sliding window approach to generate samples which contain sequences of data frames. The detailed information of 3 datasets is listed below:

- Dataset 1: Atwood number = 0.75;
- Dataset 1b: Atwood number = 0.75, different random initial condition;
- Dataset 2: Atwood number = 0.05.

Baselines We compare our model with several state-of-the-art baselines which have shown great performance on 2D turbulent flow prediction.

- U-Net (Ronneberger et al., 2015): A fully convolutional net work developed for image segmentation, also used for video prediction.
- TF-Net (Wang et al., 2020a): A hybrid deep learning framework based on multi-level spectral decomposition of turbulent flow.
- Fourier Neural Operator (FNO): (Li et al., 2020): A novel neural operator based on Fourier transformation for learning partial differential equations. In Li et al. (2020), FNO is applied to predict a single-channel vorticity field. For 3D HVDT, we require all four channels, which we concatenate before inputting to the model.

Evaluation Metrics We use Root Mean Square Error (RMSE) to measure the pixel-wise prediction performance. We also use the Energy Spectrum Error (ESE) to measure the turbulence kinetic energy of the predictions. (See Wang et al. (2020a) for details.) To handle variable-density, we use momentum instead of velocity to calculate the spectrum as it is suggested in Aslangil et al. (2020c). Another metric is Mass Conservation. For simulated variable-density datasets, the mass-conservation equation holds: $\frac{\partial \rho u_i}{\partial x_i} = -\frac{\partial \rho}{\partial t}$. We calculate the divergence of momentum using the model predicted momentum \mathbf{p} and compare to the time derivative of the ground truth density. The time derivative is approximated by the finite difference over one time step of our data, which is approximately 2 orders of magnitude larger than the DNS time step. Hence, we incur some error which is visible even for our ground truth data.

Trade-off between Interval and Precision To understand why Taylor approximation improved performance so significantly relative to strong baselines such as Fourier Neural Operators (Li et al., 2020), we varied the input step size and found that higher order Taylor approximations are primarily helpful when the step size in numerical differentiation between inputs is small. As the step size between inputs increases, the additional benefit of higher order Taylor approximation decreases. We also find that Taylor approximation is more helpful for the truly turbulent and non-stationary chaotic data we consider relative to the more structured, cyclic, and stationary data such as Rayleigh-Bénard Convection or isotropic stationary homogeneous turbulence considered in earlier works.

Table 3 shows the trade-off between the magnitude of time intervals and the precision of prediction. To control the variables, we fix the input length to 15 frames and predict 20 frames for $\Delta t = 1$, 10 frames for $\Delta t = 2$ and 5 frames for $\Delta t = 4$. This experiment is done on Dataset 1. All the RMSE losses are based on the average of the predicted frames.

Table 2: Interval vs Precision

	U-net	Taylor 1	Taylor-Net1	Taylor 2	Taylor-Net2
RMSE $\Delta t = 1$ step	0.2713	0.03732	0.03036	0.01752	0.01548
RMSE $\Delta t = 2$ step	0.07167	0.04146	0.02603	0.02176	0.01691
RMSE $\Delta t = 4$ step	0.05096	0.06918	0.0232	0.03305	0.02656

It is shown that the performance of U-net grows while the performance of Taylor approximation declines as the magnitude of time interval grows. The performance of higher order Taylor approximation declines faster, which makes sense because the error of n^{th} order Taylor approximation is proportional to $O(\Delta t^n)$. Lower order Taylor approximation, in contrast, works better when the interval is large.

A.3 ADDITIONAL RESULTS

Zero Padding In Figure 4, we perform an ablative study with periodic padding. Instead of periodic padding, we use zero padding. We find that the mass conservation is much worse.

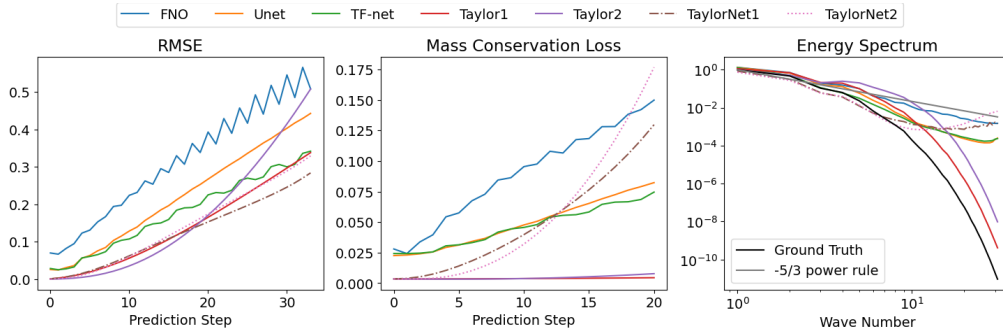


Figure 4: Comparison of our model with zero padding versus several baselines across 3 metrics for p prediction: RMSE, Mass Conservation, and Energy Spectrum.

Train on Dataset 1 and Test on Dataset 1b Instead of separating our train and test set in time, in this experiment, we instead vary the random seed which determines the initial conditions of our simulation. The results, shown in Figure 5, show similar errors to the other test set indicating generalizability of our model.

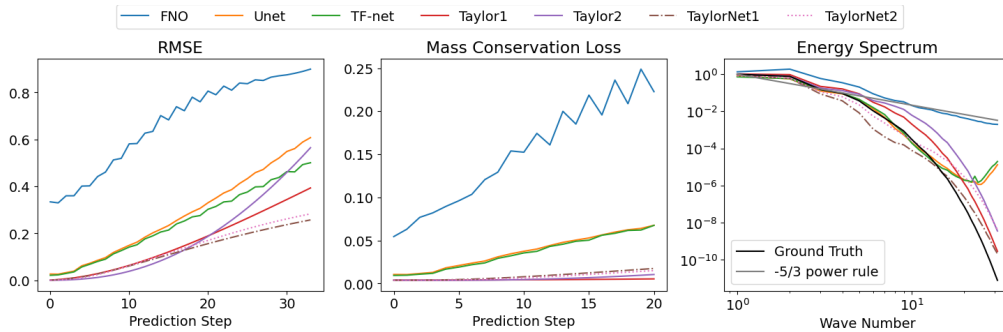


Figure 5: Train on Dataset 1 and test on Dataset 1b

Other Visualizations In figure 6, we include the full version of Figure 3 which displays visualizations of all of our baselines, our models, and the ground truth.

Velocity and Density Visualizations In Figure 7 and Figure 8 we show the predicted velocity and density fields of Taylor-Net1 versus the ground truth.

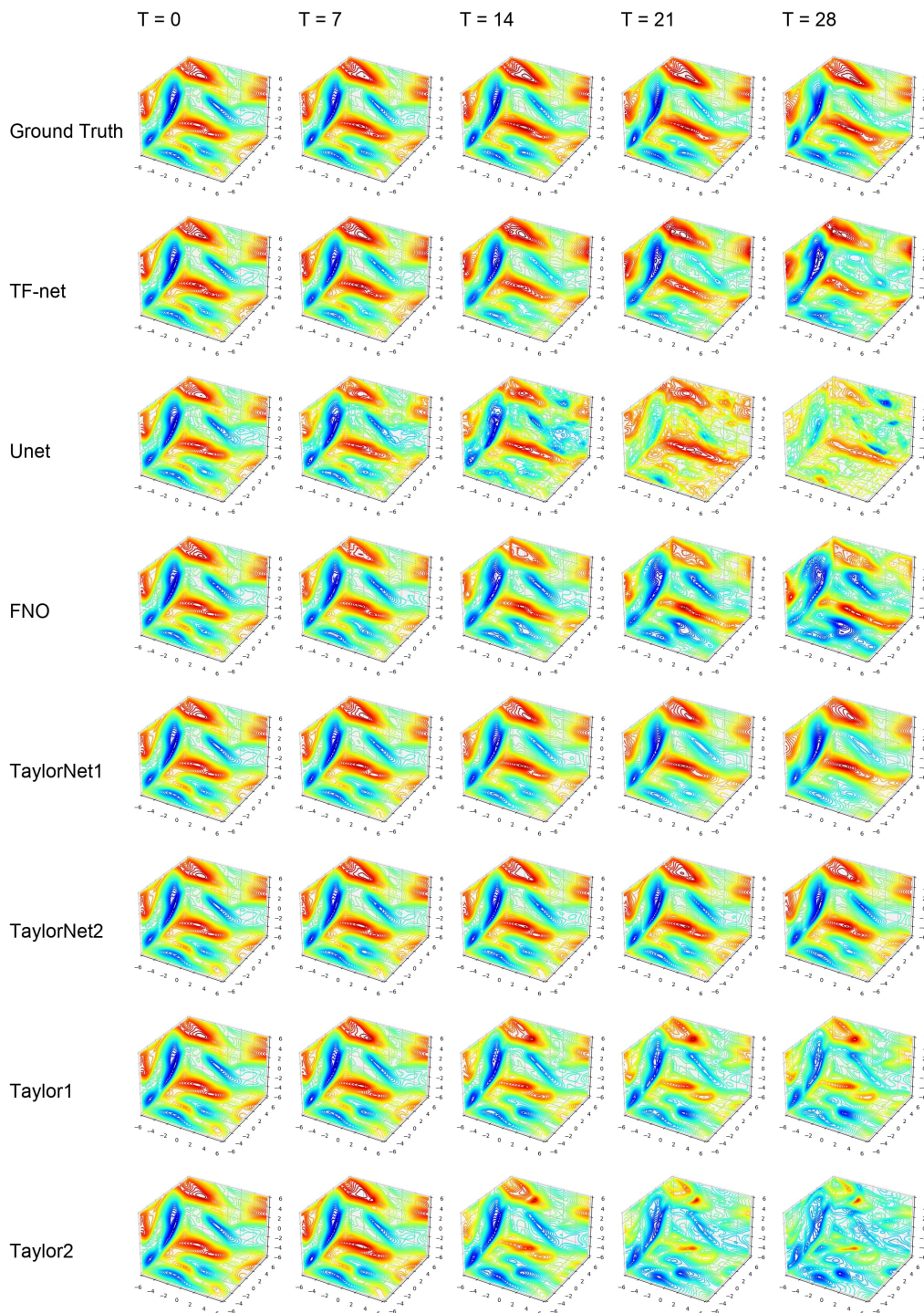


Figure 6: Visualizations of all of our baselines, our models, and the ground truth for the predicted momentum field at different points along the forecasting horizon.

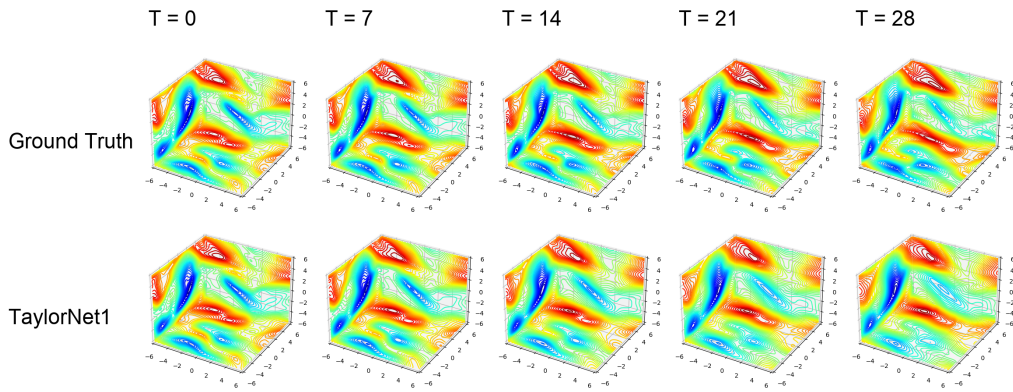


Figure 7: Visualization of predicted velocity field magnitude versus the ground truth.

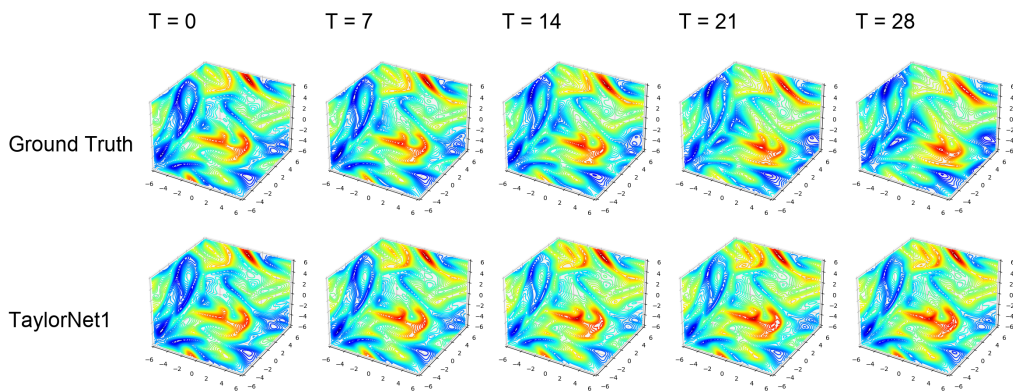


Figure 8: Visualization of predicted density field versus the ground truth.