Neural Manifold Operator for Geophysical Fluid Dynamics Prediction

Wei Xiong[1], Kun Wang[3], Yuxuan Liang[2], Hao Wu[3] ✉, Xiaomeng Huang[1] ✉,

[1]Tsinghua University [2]Hong Kong University of Science and Technology (Guangzhou)
[3]University of Science and Technology of China
xiongw21@mails.tsinghua.edu.cn; wk520529@mail.ustc.edu.cn; yuxliang@outlook.com;
wuhao2022@mail.ustc.edu.cn; hxm@tsinghua.edu.cn; ✉ represents the corresponding authors.

Abstract

Physics preserving, efficiency, performance in extreme value, and generalization capacity are the key challenges for predicting geophysical fluid dynamics. Although several recent works have proposed various models to address these issues, satisfying all four challenges mentioned above with a single model is still not solved. To address these problems, we propose Neural Manifold Operator (NMO), a novel operator learning paradigm for calculating the intrinsic dimension of the evolution operator, to learn the intrinsic dynamics of the geophysical fluid. By optimal dimensional representation, our model can accurately learn the underlying evolution operators by a compact parameter space, which not only significantly reduces the computational cost but also enhances generalization capability. Several benchmarks of geophysical fluid dynamics, including shallow water equations, extreme weather prediction, and ocean currents prediction, are used for evaluating our model. Our model has demonstrated the best performance in both statistical and physical metrics. Besides, we set ablation experiments and demonstrate that the intrinsic dimension calculated by our paradigm is the optimal dimension.

Introduction

Geophysical fluid, which mainly refers to the atmosphere and ocean on Earth, is one of the most significant parts of the Earth. Most weather and climate phenomena such as typhoons, precipitation, and ocean currents can be described as geophysical fluid dynamics. Modeling geophysical fluid dynamics is a prerequisite for weather prediction and disaster warning. Therefore, improving the accuracy and efficiency of modeling for the evolution of geophysical fluid variables has a huge socioeconomic and scientific influence. However, there are several challenges in predicting geophysical fluid variables. Geophysical fluid dynamics are complex physical systems with continuous spectrum in planetary rotation coordinates. Lacking the analytical solution of the governing equation of the geophysical fluid dynamics, accurately calculating the evolution of variables requires high-resolution grids by numerical algorithm, which causes huge computing costs.

With the rapid development of deep learning, a new paradigm for modeling and predicting physical dynamics is widely discussed. Deep learning models can learn underlying physical relationships from data and predict the future state at a lower cost. To date, deep learning methods designed for modeling and predicting physical dynamics have achieved many achievements in several studies(De Bézenac, Pajot, and Gallinari 2019). Massive geophysical fluid data, such as observation, simulation, and reanalysis data, provide conditions for deep learning methods. Therefore, modeling geophysical fluid dynamics by using learning methods shows great potential. Different from time series prediction, geophysical fluid modeling has some special requirements ($\mathcal{R}$) for deep learning methods outlined below:

- $\mathcal{R}1$: **Accuracy in statistical and physical metrics.** For various complex geophysical variables with different distributions and magnitudes or even disconnected, deep learning models require good enough performance with physical preserving.
- $\mathcal{R}2$: **Efficiency in cost and speed.** Some weather forecasting scenarios (such as nowcasting) require timeliness, which requires the model to be lightweight and efficient enough.
- $\mathcal{R}3$: **Robust in extreme value.** Disaster weather prediction scenarios generally require models to predict extreme values. Such as typhoon prediction with anomaly high wind speed and marine heat wave with anomaly high sea temperature.
- $\mathcal{R}4$: **Generalization capacity in Banach space.** Operational prediction requires deep learning methods to have a strong generalization capacity with various initial conditions and coefficients.

Despite all the recent progress, four requirements are so hard to be satisfied by one deep learning model. Several studies regard physics variables as computer vision tasks and get good performance in statistical metrics, these methods usually get poor performance in physics consistency and are hard to generalize into similar scenarios in the same physics system (e.g. different initial conditions or configure parameters). Several weather prediction models with a large number of parameters, trained by huge volumes of reanalysis data, get good performance in statistical metrics such as Root Mean Square Error (RMSE), but they still lack physical preserving analysis and cause high computation costs in the training process.

To address these problems, we propose Neural Manifold Operator (NMO), an operator learning paradigm for learning the intrinsic dimension representation of the underlying operator, to model the geophysical fluid dynamics and meet the above four requirements. By calculating the minimum dimensional submanifold representation of the geophysical variables in the latent space, NMO can adaptively determine the intrinsic dimension of the system. By projecting the latent space into a compact space of the intrinsic dimension, NMO enables accurate ($\mathcal{R}1$) and efficient ($\mathcal{R}2$) learning of the underlying operators and preserves the physics consistency of the system. We introduce four geophysical benchmarks, including shallow water equations, weather prediction, extreme weather events and ocean stream prediction, aiming to evaluate the capacity of our model for extreme value ($\mathcal{R}3$), generalization ($\mathcal{R}4$), and preserving physical properties. These benchmarks include satellite images and physical variables. Compared to several baseline models, NMO achieves state-of-the-art performance in statistical and physical metrics. Besides, we experimentally demonstrate that the intrinsic dimension calculated by our paradigm is the optimal dimension of the latent space in efficiency and accuracy.

Related Work

Spatio-temporal Prediction Spatio-temporal Spatio-temporal prediction, crucial in fields like meteorology, environmental monitoring (Nguyen, Hu, and Spanos 2018; Singh et al. 2010), urban planning (Li et al. 2020a; Falahatkar and Rezaei 2020), traffic management, economics, sociology, and public health (Sun et al. 2020; Wang, Cao, and Philip 2020; Jiang et al. 2021; Jiang, Ji, and Li 2019; Wang et al. 2022a), involves analyzing temporal and spatial data patterns using machine learning, statistical, and deep learning methods (Oh et al. 2015; Mathieu, Couprie, and LeCun 2015; Tulyakov et al. 2018). These methods often merge multiple models for enhanced accuracy and robustness, especially important in real-time data stream analysis with IoT advancements (Kim, Ahn, and Bengio 2019; Wang et al. 2022c; Weissenborn, Täckström, and Uszkoreit 2019; Kumar et al. 2019). Although these models achieve good performance in statistical metrics, preserving physics ($\mathcal{R}1$) and generalizing in similar scenarios ($\mathcal{R}4$) is still challenging when using these methods to model the evolution of geophysical fluid dynamics.

Geophysical Fluid Prediction In the early stage, due to similar tensor shapes, several models designed by computer vision are used for predicting geophysical fluid dynamics. There are several models based on CNN(Kareem, Hamad, and Askar 2021). Recently, more works are focused on building the foundation model of weather prediction(Pathak et al. 2022; Bi et al. 2023; Lam et al. 2023; Chen et al. 2023a,b; Zhang et al. 2023). Massive reanalysis datasets are used to train these model to achieve excellent performance in extreme weather forecasting ($\mathcal{R}3$) such as tropical cyclone, and demonstrating good generalization capability ($\mathcal{R}4$). For several statistical metrics of RMSE and ACC, these AI-based models seem to surpass numerical weather prediction systems(Bi et al. 2023). However, high computational cost(Kurth et al. 2023) of training and inference ($\mathcal{R}2$) limits their application in several scenarios.

Operator Learning Operator learning is designed for operators mapping between infinite-dimension function spaces. The well-trained model can achieve generalization to Banach spaces. DeepONet(Lu et al. 2021) and its further works(Wang, Wang, and Perdikaris 2021; Wang and Perdikaris 2023; Goswami et al. 2023) proved that the model can converge to the target operator if the sampling in the function space is sufficient. However, these methods are hard to model complex real-world dynamics. Inspired by the Green Function method, several neural operators(Kovachki et al. 2023; Li et al. 2020b,b,c) are designed for solving partial differential equations. To represent the evolution of the nonlinear dynamics, some neural operator methods(Xiong et al. 2023a,b; Wang et al. 2022b) based on Koopman theory use linear representation to model the evolution of the observation state of the variables. However, these methods do not further discuss the finite-dimensional representation of the infinite-dimensional operator.

Methodology

Problem Definition

The governing equations of geophysical fluid are derived from the Navier-Stokes equation with Coriolis force and conservation law. A simplified barotropic shallow water equation describing the velocity $\mathbf{u} = \{u, v\}$ and potential height h of the geophysical fluid can be expressed as:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + \frac{\partial \varphi}{\partial x} - fv = 0, \tag{1}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + \frac{\partial \varphi}{\partial y} + fu = 0, \tag{2}$$

$$\frac{\partial \varphi}{\partial t} + \frac{\partial u\varphi}{\partial x} + \frac{\partial v\varphi}{\partial y} = 0. \tag{3}$$

The general form of the evolution equation for the geophysical fluid variables can be expressed as:

$$\frac{\partial F}{\partial t} + \mathcal{L}F = \mathcal{G}, \tag{4}$$

where F denotes the geophysical fluid variables, $\mathcal{L}$ denotes the operator describing the evolution of the variables and $\mathcal{G}$ denotes the computable function. The general form of Eq. (1-3) can be defined as an operator

$$\mathcal{L} = \begin{pmatrix} u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} & -f & \frac{\partial}{\partial x} \\ f & u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} & \frac{\partial}{\partial y} \\ 0 & 0 & \frac{\partial}{\partial x}u + \frac{\partial}{\partial y}v \end{pmatrix} \tag{5}$$

to describe the evolution of geophysical variables $F = (u, v, \phi)$. The discrete form of the general evolution equation can be expressed as

$$F_{t+dt} = LF_t \tag{6}$$

where L is discrete operator of $\mathcal{L}$. Assuming that the evolution of geophysical variables F is time-invariant, different

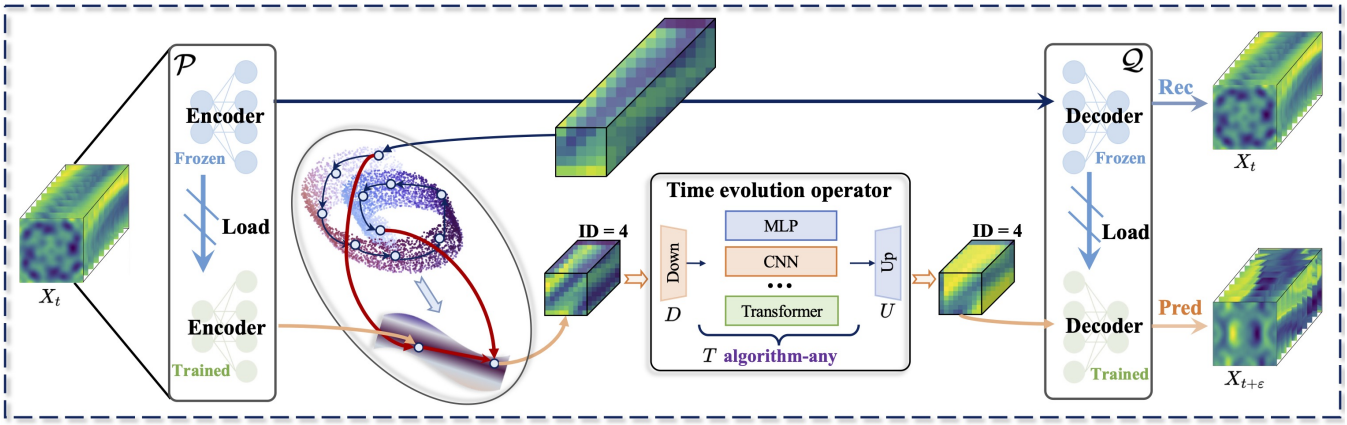

Figure 1: Overview of NMO. $\mathcal{P}$ and $\mathcal{Q}$ are the encoder and decoder, which are trained by construction constraints. Having trained the encoder and decoder, the manifold algorithm is used to calculate the intrinsic dimension of the underlying operator. Projecting into the intrinsic dimension, the time evolution module is trained with a frozen encoder and decoder.

initial states and future states of F can be regarded as separable Banach spaces $\mathcal{X}$ and $\mathcal{Y}$. Although the governing equation of some geophysical fluid variables differs from the shallow water equations or is even unknown, the general form of dynamic evolution can be expressed similarly. Therefore, our paradigm is designed for learning the discrete evolution operator L mapping from the initial states $\mathcal{X}$ to its future states $\mathcal{Y}$. Define Ω be a bounded open set in $\mathbb{R}^d$, and let X and Y be separable Banach spaces defined on Ω with dimensions d^x and d^y, respectively. Suppose that we have observation pairs $\{x_i, y_i\}_{i=1}^n$ from X and Y, we can build an approximation of L by constructing a parametric map G_θ,

$$G_\theta : \mathcal{X} \to \mathcal{Y}, \theta \in \Theta, \tag{7}$$

where θ is the parameter of the parametric map in finite-dimensional parameter space Θ. Define a cost function C to seek a minimizer of the problem

$$\min_{\theta \in \Theta} \mathbb{E}\left[C\left(G(x,\theta), L(x)\right)\right], \tag{8}$$

where x is the observation of the initial state F from the Banach space $\mathcal{X}$, to calculate the optimal parameter $\theta^\dagger \in \Theta$ and converge to

$$G\left(\cdot, \theta^\dagger\right) = G_{\theta^\dagger} \approx L. \tag{9}$$

Overall Structure. There are several main structures in our architecture: an encoder $\mathcal{P} : \mathcal{X} \to \hat{\mathcal{X}}$ for projecting from physics space to latent space, a decoder $\mathcal{Q} : \hat{\mathcal{Y}} \to \mathcal{Y}$ for projecting from latent space to physics space, where $\hat{\mathcal{X}}$ and $\hat{\mathcal{Y}}$ is the latent Banach space on Ω with a higher-dimension d^l, a linear projection $\mathcal{W} : \hat{\mathcal{X}} \to \mathcal{T}_{\mathcal{X}}$ to project from latent space with d^l to the Banach space with d^m, where m is the estimated intrinsic dimension of the physics variables calculated by maximum likelihood estimation algorithm, and its pseudo-inverse operator $\mathcal{W}^+ : \mathcal{T}_{\mathcal{Y}} \to \hat{\mathcal{Y}}$, and time evolution operator $\mathcal{T} : \mathcal{T}_{\mathcal{X}} \to \mathcal{T}_{\mathcal{Y}}$ to learn the evolution of physics systems in the Banach space with estimated intrinsic dimension d^m. Therefore, the parametric map G_θ can be defined as

$$G_\theta := \mathcal{Q} \circ \mathcal{W}^+ \circ \mathcal{T} \circ \mathcal{W} \circ \mathcal{P}. \tag{10}$$

where $\circ$ denotes the operator composition.

Encoder

The encoder $\mathcal{P}$ consists of multiple convolutional layers, each processing input geophysical fluid variables. Given input $x \in \mathbb{R}^{T\times C\times H\times W}$, where C is the input dimension, the encoder's convolution layer is defined as:

$$\hat{\mathcal{X}} = \text{LeakyReLU}(\text{GroupNorm}(\text{Conv2d}(\mathcal{X}))), \tag{11}$$

This structure transforms the physical space into a higher latent space. The transformed tensor $\hat{\mathcal{X}} \in \mathbb{R}^{T\times \hat{C}\times \hat{H}\times \hat{W}}$ represents a high-dimensional latent variable of the geophysical fluid variables.

Intrinsic Dimension Calculation

Post-encoding, the latent tensor $\hat{\mathcal{X}} = \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n\}$ are considered as a set of vectors in latent space $\mathbb{R}^{d^l}$. Based on the theory of manifold learning, there exists a minimal submanifold $\mathcal{M}$ in $\mathbb{R}^{d^l}$, with $m < d^l$, for lower-dimensional representation of the high-dimensional tensor $\hat{\mathcal{X}}$. The dimension m of the minimal submanifold $\mathcal{M}$ is regarded as the intrinsic dimension of $\hat{\mathcal{X}}$. We assume that m is the intrinsic dimension of the underlying operators for the evolution of the latent space.

Observation points on $\mathcal{M}$, representing latent tensors, are used to estimate intrinsic dimension. The estimation is based on the density functional of these points (Costa, Girotra, and Hero 2005):

$$\log \int_{B(\boldsymbol{v}_0, r)} g(f(\boldsymbol{v}))\mu(\mathrm{d}\boldsymbol{v}), \tag{12}$$

where g is a metric on $\mathcal{M}$ and $B(\boldsymbol{v}_0, r)$ is a ball of radius r. The density functional approximates the count of points in $B(\boldsymbol{v}_0, r)$. The distance $T_k(\boldsymbol{v}_0)$ to the k-nearest neighbor

determines r (Pettis et al. 1979). The count in the k-nearest ball can be modeled as a Poisson process for maximum likelihood estimation of local intrinsic dimension (Levina and Bickel 2004; Costa, Girotra, and Hero 2005):

$$\hat{m_0} = \frac{1}{k-1} \sum_{j=1}^{k-1} \log \frac{T_k(\boldsymbol{v}_0)}{T_j(\boldsymbol{v}_0)}. \tag{13}$$

The overall intrinsic dimension m is the average across all points (Levina and Bickel 2004; Chen et al. 2022):

$$\hat{m} = \frac{1}{n} \sum_{i=1}^{n} \hat{m}_k, \quad m = E(\hat{m}(\boldsymbol{v})). \tag{14}$$

Time Evolution Operator

Having calculated the intrinsic dimension m. Linear projection $\mathcal{W}$ transforms the tensor from latent space to the Banach space with m. Time Evolution Operator to learn the map from $\mathcal{T}_{\mathcal{X}} \to \mathcal{T}_{\mathcal{Y}}$ in the Banach space with m, which is divided into three layers to implement, downsampling layers, Evolution layers, and upsampling layers.

- **Downsampling.** A series of convolutional layers reduce the data dimensions, extracting more abstract features, represented as $\mathcal{T}_{\mathcal{X}}^{\text{down}} = \text{Downsample}(\mathcal{T}_{\mathcal{X}})$.
- **Evolution layer.** Complex nonlinear transformations are applied on the reduced dimensions to simulate temporal evolution, represented as $\mathcal{T}_{\mathcal{Y}}^{\text{down}} = \text{EvoLayers}(\mathcal{T}_{\mathcal{X}}^{\text{down}})$.
- **Upsampling.** The data dimensions are restored to their initial state, outputting the predicted future state $\mathcal{T}_{\mathcal{Y}} = \text{Upsample}(\mathcal{T}_{\mathcal{Y}}^{\text{down}})$.

It is worth noting that the evolution layer can be implemented by MLP, CNN or Transformer modules, which means our paradigm applies to various structures.

Decoder

After the Time Evolution Operator $\mathcal{T}$ and the inverse of linear projection $\mathcal{W}$, the Decoder reconstructs the tensor to the original physical space $\mathbb{R}^{T \times C \times H \times W}$. The formal mathematical definition is as follows:

$$\mathcal{Y} = \text{LeakyReLU}(\text{GroupNorm}(\text{ConvTranspose2d}(\hat{\mathcal{Z}}))) \tag{15}$$

Experiments

Experiment setting

Baseline Models. Several advanced and representative models are used for evaluating our model.

$\mathcal{B}1$: **Generic deep learning models:** U-Net (Ronneberger, Fischer, and Brox 2015), Residual Networks (ResNet) (He et al. 2016) and Swin-Transformer (Swin) (Liu et al. 2021) are representative and mainstream computer vision backbone model, which is often used for various tasks.

$\mathcal{B}2$: **Time series prediction models:** SimVP-v2 (Tan et al. 2022), PredRNN-V2 (Wang et al. 2022c) are representative general models for time series prediction. EarthFormer (Gao et al. 2022) is designed for the time series of the Earth system.

$\mathcal{B}3$: **Operator-learning models:** Fourier Neural Operator (FNO) (Li et al. 2020b) is one of the most representative neural operator models designed for learning mapping between Banach space.

$\mathcal{B}4$: **Physics-guided deep learning models:** Turbulence-Flow Net (TF-Net) (Wang et al. 2020) and Latent Spectral Models (LSM) (Wu et al. 2023) are advanced physics-guided models incorporating physics knowledge into inductive bias.

Evaluation Metrics. Our model is trained by the loss function of mean square error (MSE). Root-mean-square deviation (RMSE) is used for evaluation. Additionally, we use mass conservation as the physical metric in shallow water equations scenario to test physics preserving, expressed as

$$\frac{d}{dt} \int_D h d\sigma = 0 \tag{16}$$

where η is potential height, D is the computational domain of the equations in 2-dimensional Euclidean space $\mathbb{R}^2$.

Implementation Details. We implement our model using the PyTorch framework and leverage the four A100-PCIE-40GB as computing support.

Main Results

In this section, We analyze the excellence of the model in several scenarios, as described in the subsequent sections.

Shallow Water Equations

Scenario Description. Shallow water equations, derived from the barotropic approximation of the Navier-Stokes equation, are widely used in the atmosphere and ocean modeling to simulate the surface layer, which is appropriate for testing the solving algorithm about physics preserving. We consider the 2-D general form of shallow water equations under a rotation coordinate system expressed as

$$\frac{du}{dt} - fv = -g\frac{d(\eta)}{dx} + \frac{\tau_x}{(\rho_0 H)} - \kappa u, \tag{17}$$

$$\frac{dv}{dt} + fu = -g\frac{d(\eta)}{dy} + \frac{\tau_y}{(\rho_0 H)} - \kappa v, \tag{18}$$

$$\frac{d\eta}{dt} + \frac{d((\eta + H) * u)}{dx} + \frac{d((\eta + H) * v)}{dy} = 0, \tag{19}$$

where u, v, η denotes u-component velocity, v-component velocity, and potential height respectively, f denotes Coriolis's force, and H, g, τ_x and τ_y is the the thickness of the fluid layer, gravitational acceleration, and wind stress in the x-direction and y-direction, respectively. In this scenario, our model learns the operator mapping the variables $F = (u, v, \eta)$ up to time 10 to the variables up to a future time $T > 10$. We conduct experiments with and without the

Table 1: Performance comparison with 9 baseline models in extreme events and west boundary stream scenarios. RMSE is used for the evaluation of these models, with a smaller RMSE value indicating greater accuracy. Since FNO is designed for single variables prediction, we only evaluate these models in single variables scenarios to ensure optimal performance of the baseline models. The underline indicates the most accurate result in baseline models. The bold font indicates the most accurate of all models. The asterisk (*) denotes GPU memory overflow (exceeding 40GB). The forward slash (/) indicates that the original model is only designed for single variable prediction.

MODEL	SEVIR	HIMAWARI-8	WEST BOUNDARY STREAM
U-NET	2.0280	0.0546	<u>0.0591</u>
RESNET	2.0787	0.1246	0.0709
PREDRNN-V2	1.9741	0.0234	0.0651
SWIN-TRANSFORMER	2.0067	0.0273	0.1682
SIMVP-V2	0.7943	0.0193	0.0658
EARTHFORMER	<u>0.2877</u>	0.0671	0.1612
TF-NET	2.1946	<u>0.0172</u>	0.1033
FNO	1.0099	/	/
LSM	1.2569	/	/
NMO	**0.1698**	**0.0404**	**0.0161**
PROMOTION	41.01%	31.64%	6.40%

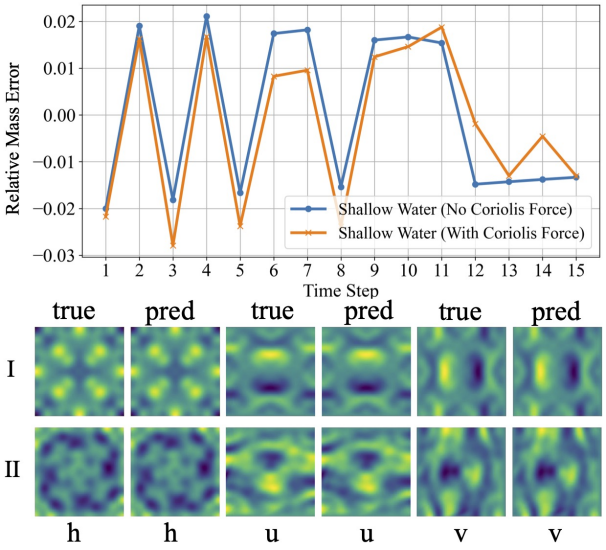

Figure 2: Relative mass error at each time step and visualization of prediction results of each model on the Shallow-Water equations scenario.

Coriolis force f to assess the performance of our model. The result is shown in the Figure. 2. According to the result, our model learns mass conservation and preserves the physical property of the system.

Extreme Weather Events

Extreme weather events have important impacts on humanity and society, which is significant for weather prediction. In our paper, we mainly focus on the prediction of tropical cyclones, thunderstorms, and intense precipitation. In operational extreme weather events prediction, atmospheric variables and satellite images are mainly used to predict, which means the model must be available for both physical variables and satellite images. We use Storm EVent ImageRy (SEVIR) as our benchmark in this scenario. SEVIR dataset (Veillette, Samsi, and Mattioli 2020) including satellite and radar weather data, are used to evaluate our model's accuracy in forecasting short-term severe weather events like thunderstorms and intense precipitation.

We also use the Japanese Himawari-8 Geostationary Satellite Data (Himawari-8) as our benchmark for tropical cyclone prediction by satellite image. The data is a three-layer water vapor channel dataset covering the East and Southeast Asian Pacific coastal regions. We use it to test the model's ability to predict tropical cyclones' structure and water vapor distribution in the next 36 hours. The result is shown in Table. 1 and our model gets state-of-the-art performance and over 30% promotion.

Ocean Stream Prediction

The ocean, which covers about 71% of the Earth, is an important geophysical fluid. Different from the atmosphere, ocean fluid with complex land and topology boundaries is considered incompressible. Because of the Coriolis force, western intensification, which is a special phenomenon in oceanography, is a key challenge for ocean stream prediction because affects the dynamics on the global scale. In the western boundary of the ocean, the stream velocity of the currents is much higher than in other areas and it is highly related to mesoscale eddies, which means it needs much more spatial resolution to resolve in traditional numeric methods. Therefore, we utilize the sea surface stream velocity data from the Copernicus Marine Environment Monitoring Service (CMEMS) as our benchmark. The Kuroshio region (10-42°N, 123-155°E) and the Gulf Stream region (20-52°N, 33-65°W), the two most representative west boundary currents, are used to analyze. The result is shown in Table. 1 and our model gets 6.4% promotion in two boundary stream

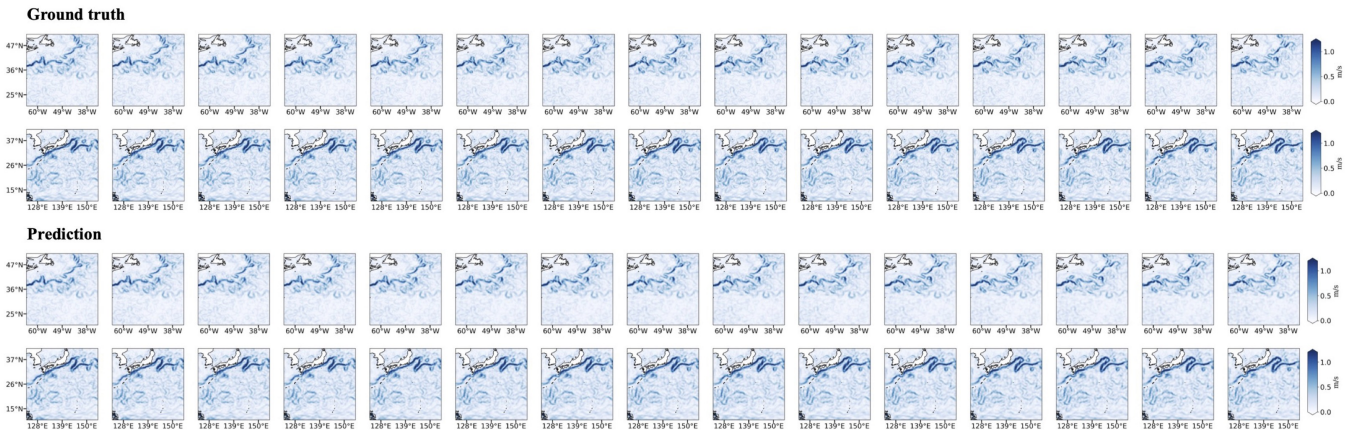

Figure 3: Visualization of Gulf Stream and Kuroshio.

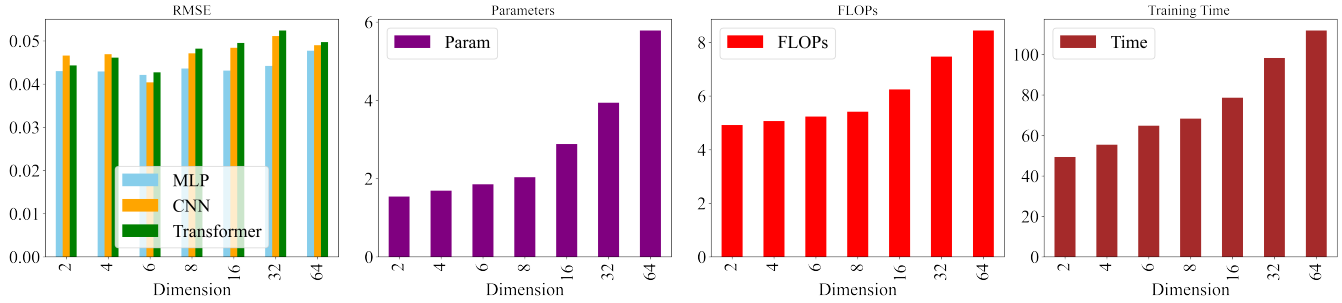

Figure 4: Ablation study evaluates different dimensions of the time evolution operator using three types: MLP, CNN, and Transformer. Results include optimal model parameters, floating point operations (FLOPs), and training time. The most accurate results and intrinsic dimensions are highlighted with underlining and bold font.

scenarios average. The visualization of our model is shown in Figure. 3.

Ablation Study

To demonstrate the intrinsic dimension calculated by our paradigm is the optimal dimension of the latent space, we set several experiments on various dimensions for three types of time evolution operator implementations. As Table 4 shows, when the latent space is projected onto the calculated intrinsic dimension to learn the underlying operators, NMO implementations consistently achieve the best performance across all scenarios.

Conclusions

Neural Manifold Operator accurately predicts geophysical fluid dynamics and addresses the four challenges mentioned above through a single model. Furthermore, we experimentally demonstrated that our paradigm applies to various network structure implementations.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (42125503) and the National Key Research and Development Program of China (2022YFE0195900, 2021YFC3101600, 2020YFA0608000, 2020YFA0607900).

References

Bi, K.; Xie, L.; Zhang, H.; Chen, X.; Gu, X.; and Tian, Q. 2023. Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, 619(7970): 533–538.

Chen, B.; Huang, K.; Raghupathi, S.; Chandratreya, I.; Du, Q.; and Lipson, H. 2022. Automated discovery of fundamental variables hidden in experimental data. *Nature Computational Science*, 2(7): 433–442.

Chen, K.; Han, T.; Gong, J.; Bai, L.; Ling, F.; Luo, J.-J.; Chen, X.; Ma, L.; Zhang, T.; Su, R.; et al. 2023a. FengWu: Pushing the Skillful Global Medium-range Weather Forecast beyond 10 Days Lead. *arXiv preprint arXiv:2304.02948*.

Chen, L.; Zhong, X.; Zhang, F.; Cheng, Y.; Xu, Y.; Qi, Y.; and Li, H. 2023b. FuXi: A cascade machine learning forecasting system for 15-day global weather forecast. *arXiv preprint arXiv:2306.12873*.

Costa, J. A.; Girotra, A.; and Hero, A. 2005. Estimating local intrinsic dimension with k-nearest neighbor graphs. In *IEEE/SP 13th Workshop on Statistical Signal Processing, 2005*, 417–422. IEEE.

De Bézenac, E.; Pajot, A.; and Gallinari, P. 2019. Deep learning for physical processes: Incorporating prior scientific knowledge. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(12): 124009.

Falahatkar, S.; and Rezaei, F. 2020. Towards low carbon

cities: Spatio-temporal dynamics of urban form and carbon dioxide emissions. *Remote Sensing Applications: Society and Environment*, 18: 100317.

Gao, Z.; Shi, X.; Wang, H.; Zhu, Y.; Wang, Y. B.; Li, M.; and Yeung, D.-Y. 2022. Earthformer: Exploring space-time transformers for earth system forecasting. *Advances in Neural Information Processing Systems*, 35: 25390–25403.

Goswami, S.; Bora, A.; Yu, Y.; and Karniadakis, G. E. 2023. Physics-informed deep neural operator networks. In *Machine Learning in Modeling and Simulation: Methods and Applications*, 219–254. Springer.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Jiang, R.; Yin, D.; Wang, Z.; Wang, Y.; Deng, J.; Liu, H.; Cai, Z.; Deng, J.; Song, X.; and Shibasaki, R. 2021. Dl-traff: Survey and benchmark of deep learning models for urban traffic prediction. In *Proceedings of the 30th ACM international conference on information & knowledge management*, 4515–4525.

Jiang, X.; Ji, P.; and Li, S. 2019. CensNet: Convolution with Edge-Node Switching in Graph Neural Networks. In *IJCAI*, 2656–2662.

Kareem, S.; Hamad, Z. J.; and Askar, S. 2021. An evaluation of CNN and ANN in prediction weather forecasting: A review. *Sustainable Engineering and Innovation*, 3(2): 148–159.

Kim, T.; Ahn, S.; and Bengio, Y. 2019. Variational temporal abstraction. *Advances in Neural Information Processing Systems*, 32.

Kovachki, N.; Li, Z.; Liu, B.; Azizzadenesheli, K.; Bhattacharya, K.; Stuart, A.; and Anandkumar, A. 2023. Neural operator: Learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research*, 24(89): 1–97.

Kumar, M.; Babaeizadeh, M.; Erhan, D.; Finn, C.; Levine, S.; Dinh, L.; and Kingma, D. 2019. Videoflow: A flow-based generative model for video. *arXiv preprint arXiv:1903.01434*, 2(5): 3.

Kurth, T.; Subramanian, S.; Harrington, P.; Pathak, J.; Mardani, M.; Hall, D.; Miele, A.; Kashinath, K.; and Anandkumar, A. 2023. Fourcastnet: Accelerating global high-resolution weather forecasting using adaptive fourier neural operators. In *Proceedings of the Platform for Advanced Scientific Computing Conference*, 1–11.

Lam, R.; Sanchez-Gonzalez, A.; Willson, M.; Wirnsberger, P.; Fortunato, M.; Alet, F.; Ravuri, S.; Ewalds, T.; Eaton-Rosen, Z.; Hu, W.; et al. 2023. Learning skillful medium-range global weather forecasting. *Science*, eadi2336.

Levina, E.; and Bickel, P. 2004. Maximum likelihood estimation of intrinsic dimension. *Advances in neural information processing systems*, 17.

Li, R.; He, H.; Wang, R.; Huang, Y.; Liu, J.; Ruan, S.; He, T.; Bao, J.; and Zheng, Y. 2020a. Just: Jd urban spatio-temporal data engine. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, 1558–1569. IEEE.

Li, Z.; Kovachki, N.; Azizzadenesheli, K.; Liu, B.; Bhattacharya, K.; Stuart, A.; and Anandkumar, A. 2020b. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*.

Li, Z.; Kovachki, N.; Azizzadenesheli, K.; Liu, B.; Bhattacharya, K.; Stuart, A.; and Anandkumar, A. 2020c. Neural operator: Graph kernel network for partial differential equations. *arXiv preprint arXiv:2003.03485*.

Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; and Guo, B. 2021. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF international conference on computer vision*, 10012–10022.

Lu, L.; Jin, P.; Pang, G.; Zhang, Z.; and Karniadakis, G. E. 2021. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature machine intelligence*, 3(3): 218–229.

Mathieu, M.; Couprie, C.; and LeCun, Y. 2015. Deep multi-scale video prediction beyond mean square error. *arXiv preprint arXiv:1511.05440*.

Nguyen, L. V.; Hu, G.; and Spanos, C. J. 2018. Efficient sensor deployments for spatio-temporal environmental monitoring. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 50(12): 5306–5316.

Oh, J.; Guo, X.; Lee, H.; Lewis, R. L.; and Singh, S. 2015. Action-conditional video prediction using deep networks in atari games. *Advances in neural information processing systems*, 28.

Pathak, J.; Subramanian, S.; Harrington, P.; Raja, S.; Chattopadhyay, A.; Mardani, M.; Kurth, T.; Hall, D.; Li, Z.; Azizzadenesheli, K.; et al. 2022. Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *arXiv preprint arXiv:2202.11214*.

Pettis, K. W.; Bailey, T. A.; Jain, A. K.; and Dubes, R. C. 1979. An intrinsic dimensionality estimator from near-neighbor information. *IEEE Transactions on pattern analysis and machine intelligence*, (1): 25–37.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, 234–241. Springer.

Singh, A.; Ramos, F.; Whyte, H. D.; and Kaiser, W. J. 2010. Modeling and decision making in spatio-temporal processes for environmental surveillance. In *2010 IEEE International Conference on Robotics and Automation*, 5490–5497. IEEE.

Sun, J.; Zhang, J.; Li, Q.; Yi, X.; Liang, Y.; and Zheng, Y. 2020. Predicting citywide crowd flows in irregular regions using multi-view graph convolutional networks. *IEEE Transactions on Knowledge and Data Engineering*, 34(5): 2348–2359.

Tan, C.; Gao, Z.; Li, S.; and Li, S. Z. 2022. Simvp: Towards simple yet powerful spatiotemporal predictive learning. *arXiv preprint arXiv:2211.12509*.

Tulyakov, S.; Liu, M.-Y.; Yang, X.; and Kautz, J. 2018. Mocogan: Decomposing motion and content for video generation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1526–1535.

Veillette, M.; Samsi, S.; and Mattioli, C. 2020. Sevir: A storm event imagery dataset for deep learning applications in radar and satellite meteorology. *Advances in Neural Information Processing Systems*, 33: 22009–22019.

Wang, K.; Zhou, Z.; Wang, X.; Wang, P.; Fang, Q.; and Wang, Y. 2022a. A2DJP: A two graph-based component fused learning framework for urban anomaly distribution and duration joint-prediction. *IEEE Transactions on Knowledge and Data Engineering*.

Wang, R.; Dong, Y.; Arik, S. O.; and Yu, R. 2022b. Koopman Neural Operator Forecaster for Time-series with Temporal Distributional Shifts. In *The Eleventh International Conference on Learning Representations*.

Wang, R.; Kashinath, K.; Mustafa, M.; Albert, A.; and Yu, R. 2020. Towards physics-informed deep learning for turbulent flow prediction. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1457–1466.

Wang, S.; Cao, J.; and Philip, S. Y. 2020. Deep learning for spatio-temporal data mining: A survey. *IEEE transactions on knowledge and data engineering*, 34(8): 3681–3700.

Wang, S.; and Perdikaris, P. 2023. Long-time integration of parametric evolution equations with physics-informed deeponets. *Journal of Computational Physics*, 475: 111855.

Wang, S.; Wang, H.; and Perdikaris, P. 2021. Learning the solution operator of parametric partial differential equations with physics-informed DeepONets. *Science advances*, 7(40): eabi8605.

Wang, Y.; Wu, H.; Zhang, J.; Gao, Z.; Wang, J.; Philip, S. Y.; and Long, M. 2022c. Predrnn: A recurrent neural network for spatiotemporal predictive learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2): 2208–2225.

Weissenborn, D.; Täckström, O.; and Uszkoreit, J. 2019. Scaling autoregressive video models. *arXiv preprint arXiv:1906.02634*.

Wu, H.; Hu, T.; Luo, H.; Wang, J.; and Long, M. 2023. Solving High-Dimensional PDEs with Latent Spectral Models. In Krause, A.; Brunskill, E.; Cho, K.; Engelhardt, B.; Sabato, S.; and Scarlett, J., eds., *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, 37417–37438. PMLR.

Xiong, W.; Huang, X.; Zhang, Z.; Deng, R.; Sun, P.; and Tian, Y. 2023a. Koopman neural operator as a mesh-free solver of non-linear partial differential equations. *arXiv preprint arXiv:2301.10022*.

Xiong, W.; Ma, M.; Huang, X.; Zhang, Z.; Sun, P.; and Tian, Y. 2023b. Koopmanlab: machine learning for solving complex physics equations. *APL Machine Learning*, 1(3).

Zhang, Y.; Long, M.; Chen, K.; Xing, L.; Jin, R.; Jordan, M. I.; and Wang, J. 2023. Skilful nowcasting of extreme precipitation with NowcastNet. *Nature*, 619(7970): 526–532.