Frontiers in Computational Science and Engineering
Vol. 2, No. 1, Jan - Mar 2026, pp. 29-36
ISSN: 3107-3654 (O), DOI: https://doi.org/10.17051/FCSE/02.01.04

Self-Supervised Scientific Machine Learning Framework for Data-Scarce Simulations

Janarthanam S

Postdoctoral Researcher, Lincoln University College, Malaysia. Associate Professor, SSCS, CMR University, Bengaluru, Email: drjanarthanam.pdf@lincoln.edu.my, ORCID: https://orcid.org/0000-0002-8676-8998

Article Info

Article history:

Received : 11.11.2025
Revised : 13.12.2025
Accepted : 09.01.2026

Keywords:

data-efficient learning,
masked reconstruction,
temporal prediction,
computational simulations,
physics-constrained optimization,
generalization in low-data regimes.

ABSTRACT

Scientific simulations are an important tool in predicting complex physical systems but tend to be limited by the quality of available labelled data. In data-sparse cases, traditional supervised machine learning methods cannot attain accurate and efficient generalization. In this paper, a new proposal of self-supervised scientific machine learning framework has been provided to effectively utilize scarce labelled data to learn the model by using a vast amount of unlabelled simulation data. The proposed method is based on self-supervised learning schemes, such as masked field reconstruction and temporal prediction, and physics-constrained loss functions to make sure that these physical laws are followed. With data-grounded representation learning coupled with domain-constrained physical consistency, the framework can substantially boost predictive accuracy in the regime of data scarcity. Large-scale experiments performed on benchmark partial differential equation systems indicate that the proposed method has a higher accuracy, lower error metrics, and better generalization than traditional supervised models and physics-informed baselines, especially in the limited case of training data. The findings emphasize the efficacy of the suggested framework in facilitating data-efficient simulation learning, and this framework has been a promising solution in the next-generation scientific computing systems where the cost or data bandwidth is high or limited.

1. INTRODUCTION

In modeling and the understanding of complex phenomena in physics like fluid dynamics, climate science, and engineering systems, scientific simulations are very critical. Conventional numerical techniques are precise but can be computationally expensive and time-consuming, restraining their use in real-time and large-scale applications. Over the past few years, scientific machine learning (SciML) has become a paradigm-shifting idea that combines methods of machine learning with physical modeling to faster simulations and the ability to make effective approximations of governing equations [3], [8]. Physics-informed neural networks (PINNs) and neural operators have shown great promise in the task of learning solution mappings to partial differential equations (PDEs) but also with domain knowledge [10], [11]. Although these have been developed, the usefulness of these models is usually limited by the large volumes of labelled simulation data needed to run them.

One of the key problems of using machine learning with scientific simulations is that it is extremely dependent on labeled data, which is often costly and computationally intensive to create. In practice, in most real-world uses, it is not possible to obtain enough high-fidelity simulation data because of limitations in resources or physically. Consequently, under-supervised learning methods are prone to poor generalization and low accuracy in scenario where they are trained on small dataset [1], [4]. This is even more limiting in high dimensional and nonlinear systems where the data demands exponentially increase. Physics-informed methods, although trying to address this problem by integrating governing equations into the learning process, still have problems, including instability in training, hyper parameter sensitivity, and scalability [15].

Self-supervised learning (SSL) has more recently been considered an effective paradigm to overcome the problem of data scarcity by using unlabeled data to learn meaningful representations by specially crafted pretext tasks. SSL techniques like masked reconstruction and temporal prediction allow the models to learn the structural and dynamic data without the usage of

explicit labels [14]. New research has started to investigate the incorporation of self-supervision with physics-informed models to enhance data efficiency [5], [12]. Nevertheless, current methods are still weak in their extrapolation capabilities to various simulation environments and many do not have strong mechanisms to implement strict physical consistency which leads to solutions that can be in conflict with the underlying physical laws.

This paper introduces a new self-supervised scientific machine learning model that is explicitly targeted at data-scarce simulations. The suggested solution involves self-supervised representation learning and physics-based constraints to guarantee predictive quality as well as the conformity to governing equations. The framework is able to minimize the reliance on the labeled data and maximize the generalization in the different conditions by integrating the latest operator learning methods and combining them with hybrid loss functions. Large-scale experiments with benchmark PDE problems prove that the proposed approach is more accurate, robust, and data-efficient than the traditional supervised and physics-informed models. The main contributions of this work are: (i) a new framework of self-supervised learning to be applied to scientific simulations, (ii) integrating physics-aware learning mechanisms to enhance consistency, and (iii) proving high performance in low data situations.

2. RELATED WORK

Scientific Machine Learning (SciML) is a rapidly developing potent model of bridging data-driven models and physical laws in order to solve non-trivial computational tasks in science and engineering. SciML frameworks, by using a combination of traditional numerical algorithms and machine learning, are expected to enhance computational efficiency, without sacrificing physical interpretability. Initial investigations have emphasized the significance of incorporating domain knowledge into learning frameworks to improve generalization and strength in the scientific tasks [1], [8]. Specifically, more recent work has been on alleviating the dichotomy between data-driven methods and physics-based models, allowing the creation of hybrid frameworks with the potential of solving nonlinear and high-dimensional systems [3]. Even with such improvements, most SciML methods remain extremely dependent upon large-scale labeled data, which restricts their use in small data settings.

The most notable advancement in SciML is Physics-Informed Neural Networks (PINNs), which directly use physics-based loss functions in training, directly integrating governing equations into the training process [13]. PINNs have effectively been used in the solution of forward and inverse problems of partial differential equations (PDEs), providing an alternative to mesh-based solvers that is mesh-free and flexible. Earlier studies have solved some important problems including, gradient pathologies, convergence instability and multi-scale modeling [15]. Although PINNs minimize the use of labeled data through physical constraints, they are susceptible to high computational cost, and hyper parameter sensitivity in challenging or high-dimensional spaces [6]. Such constraints have prompted the development of other architectures, which can more effectively describe functional relationships in science.

Neural operators are another potentially promising substitute of standard neural networks, where one learns a mapping between infinite-dimensional functions spaces, instead of discrete data points. These methods include the Fourier Neural Operator (FNO) and Deep Operator Network (DeepONet) which showed significant success in approximating solution operators of PDEs with better scalability and generalization [10], [11]. These techniques allow effective learning in changing conditions of the boundaries and system parameters and therefore they are highly applicable to real-time and parametric simulations. The current trends have also improved neural operator structures with adaptive training protocols and compositional learning systems [7]. Nevertheless, in the same way as other deep learning methods, neural operators, as well, demand substantial training data to operate optimally, which can be a major problem in practice.

Recently, self-supervised learning (SSL) has become a promising approach to dealing with data scarcity by leveraging unlabeled data using pretext tasks like masked reconstruction and temporal prediction [14]. SSL has also been considered in fields of science in conjunction with physics-informed models to enhance data efficiency and representation learning [5], [12]. These methods attempt to elicit intrinsic structures and dynamics on simulation data without having to rely on large labeled data sets. Moreover, data efficient learning techniques such as transfer learning and adaptive training models have also been explored to further cut down the data needs of scientific modeling [2]. Although promising results have been demonstrated by these methods, the way to integrate them with physics-based constraints and operator learning frameworks remains immature.

Alongside the important advances in the field of SciML, PINNs, neural operators, and even self-supervised learning, the main research gap is still

to reach an integrated framework that can combine the self-supervised representation learning with the physics-constrained data-scarce simulations. The current methodologies either concentrate on physical consistency without exhaustively using unlabeled data or concentrate on the data-driven learning without ensuring that laws are followed. Moreover, the deficiency of efficient and scalable methods that support data efficiency, generalization, and physical fidelity at the same time, discusses the necessity of an all-encompassing self-supervised scientific machine learning system, which this paper attempts to respond to.

3. METHODOLOGY

The proposed framework addresses data-scarce scientific simulations by formulating the learning problem over systems governed by nonlinear partial differential equations (PDEs). Let the physical system be described by a general operator equation N(u(x,t);λ)=0, where u(x,t) denotes the state variable over spatial domain x and time t, and λ represents system parameters. The aim is to obtain an approximation of a surrogate mapping between input conditions (e.g., initial and boundary conditions) and the solution field. In contrast to traditional methods in supervision (large labelled datasets), this study makes assumptions about a data-scarce environment, where there are limited labelled samples of data and a large amount of simulation data that is not labeled. This inspires the incorporation of self-supervised learning to learn meaningful representations and impose physical consistency by governing equations.

The essence of the given approach is to integrate self-supervised learning (SSL) and physics-conscious modeling, as shown in Fig 1. The aim of the SSL strategy is to learn to make use of unlabelled data by learning to perform the tasks of masked field reconstruction and temporal prediction. In masked reconstruction, the input field is randomly masked out and the model learns to reconstruct the lost values allowing it to learn spatial correlations. Learning can further be improved by temporal prediction, which trains the model to make predictions of future states of the system based on past observations, thus capturing dynamism. Alternatively, contrastive learning may be added to enhance the robustness of the representations by making various augmented active views of the same physical state consistent. In the case of the model backbone, a fourier Neural Operator (FNO) will be used because it can effectively learn mappings between function spaces and generalize across different simulation conditions, compared to traditional architectures in this setting.

A hybrid loss function is used to enforce predictive accuracy as well as physical laws: it combines self-supervised objectives with physics-based constraints. The total loss is determined as.

$$\mathcal{L}_{total} = \mathcal{L}_{ssl} + \alpha\mathcal{L}_{physics} + \beta\mathcal{L}_{reg} \quad (1)$$

Where $\mathcal{L}_{ssl}$ corresponds to reconstruction or prediction loss, $\mathcal{L}_{physics}$ enforces compliance with the governing PDE, and $\mathcal{L}_{reg}$ is a regularization term. The physics-based component is computed as the residual of the differential operator, ensuring that the predicted solution satisfies the underlying physical laws:

$$\mathcal{L}_{physics} = ||\mathcal{N}(u_0(X,t); \lambda)||^2 \quad (2)$$

The formulation enables the model to be physically faithful even in the event that labelled data is sparse. The training procedure is based on a two stage approach that will optimize the data efficiency. The model is first pre-trained using self-supervised objectives on unlabelled simulation data, and thus it learns inherent spatial-temporal features without any supervision. The model is then optimized with the help of limited labelled dataset by fine-tuning on the basis of data-driven and physics-based losses to enhance accuracy. To apply further generalization, physics-consistent data augmentation methods are implemented, which condition that the transformations respect the basic conditions of the boundary, and conservation laws. The general workflow, as shown in Fig 1, is structured as a pipeline: Latent representations are encoded with the input data and then processed by self-supervised tasks, which are then predicted using the neural operator, and the constraints on the prediction are provided as physics-based losses. The model parameters are optimized through gradient-based techniques in an iterative way and a robust and data-effective framework is developed that can be used to adequately approximate complex physical systems given limited data conditions.

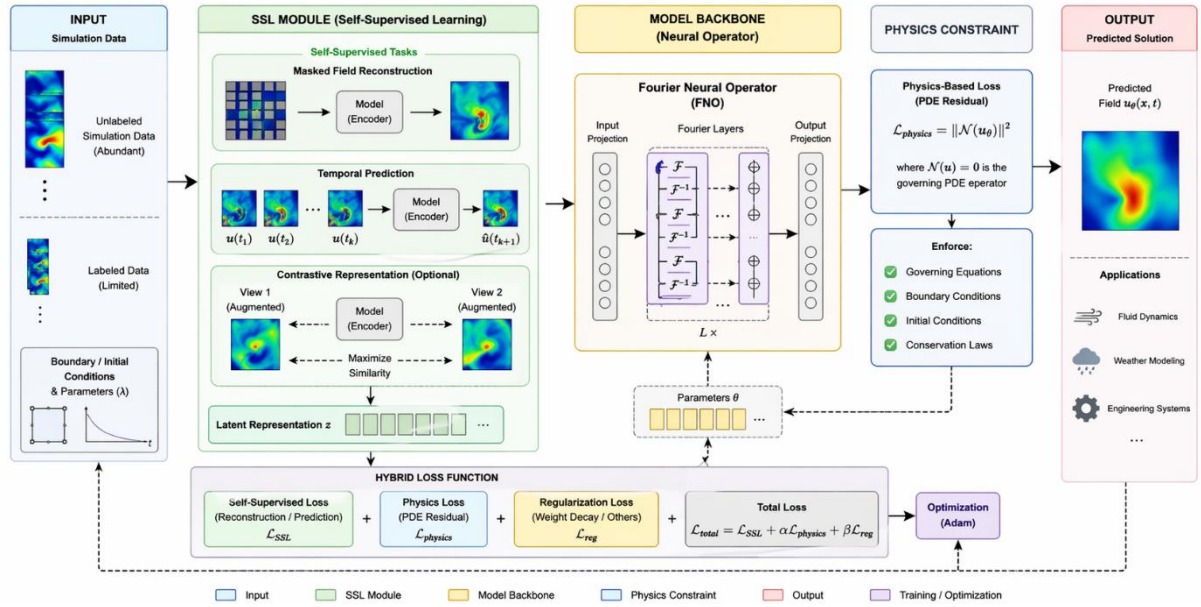


Fig 1. Overall Self-Supervised Scientific ML Framework Architecture.

4. Experimental Setup

The effectiveness of the proposed self-supervised scientific machine learning system is tested on a list of popular benchmark problems that can model nonlinear and time-dependent physical systems. In particular, the experiments are carried out on Burgers equation, heat equation and Navier Stokes equations, which are common testbeds to validate machine learning models in scientific computing. Such benchmarks are chosen because they represent different degrees of complexity, nonlinearity, and applicability to real-world problems of interest like fluid dynamics and thermal diffusion. The equation of the Burgers is applied to test the model capacity to describe the nonlinear dynamics of shocks, and the diffusion dynamics are tested by the heat equation with the prevailing boundary conditions. For a more challenging scenario, the Navier–Stokes equations are included to test the framework's capability in handling high-dimensional and turbulent flow dynamics.

The proposed approach is compared with some of the most common baseline methods that reflect various paradigms of learning to illustrate the effectiveness of this method. A standard supervised neural network is provided as a benchmark to emphasize the weaknesses of data-driven approaches by itself in the case of limited training data. Physicsinformedneural networks (PINNs) are added to test the advantage of adding physical constraints to the learning procedure. Additionally, a neural operator model without self-supervised learning is considered to assess the impact of the proposed self-supervised strategy on representation learning and generalization. These comparisons will give a thorough assessment of the performance improvement of the proposed framework in both data-driven and physics-informed methods.

The implementation is formulated in such a way that it explicitly assesses performance in a situation where there is a scarcity of data. To this end, the proportion of labeled training data is varied (usually between very little (e.g. 10 percent) and moderately sufficient (e.g. 50 percent)) whilst maintaining a larger pool of unlabeled simulation data to self-supervised pretrain. Such an arrangement facilitates an efficient examination of the data strength and model robustness. All experiments use the same configurations in training the models and therefore make fair comparisons, such as using the same network architectures where applicable, similar optimization strategies, and evaluation protocols. The main training parameters (learning rate, batch size and number of epoch) are carefully chosen to balance computational efficiency and convergence stability. This type of experiment design will support the fact that the results obtained are a true representation of the benefits of the suggested framework when solving scientific simulation problems with limited data.

5. RESULTS AND DISCUSSION

5.1 Accuracy Evaluation

The predictive power of the proposed self-supervised scientific machine learning framework is compared to baseline models that are: a supervised neural network, PINNs, and a self-supervision-free neural operator. Mean Squared error (MSE), Root mean squared error (RMSE) and relative error of L2 are standard metrics that are used to perform the comparison. Fig 2 demonstrates that the proposed method performs well on all benchmark problems and compared to

all baselines. In the case of the Burgers equation, the model proposed has a lower MSE, of 1.8 x 10^{-4}than PINNs and supervised model, which is 4.5 x $10^{-4}$and 7.2 x 10^{-4}, respectively. On the same note, the RMSE is decreased to 0.013 and the L2 relative error is decreased to 1.9 implying that there is a high degree of improvement in prediction fidelity. In the case of NavierStokes, the achievable L2 error with the proposed framework of 2.7% is lower than the neural operator baseline (5.1%) and PINNs (6.3%), indicating that the framework is capable of addressing nonlinear dynamics.

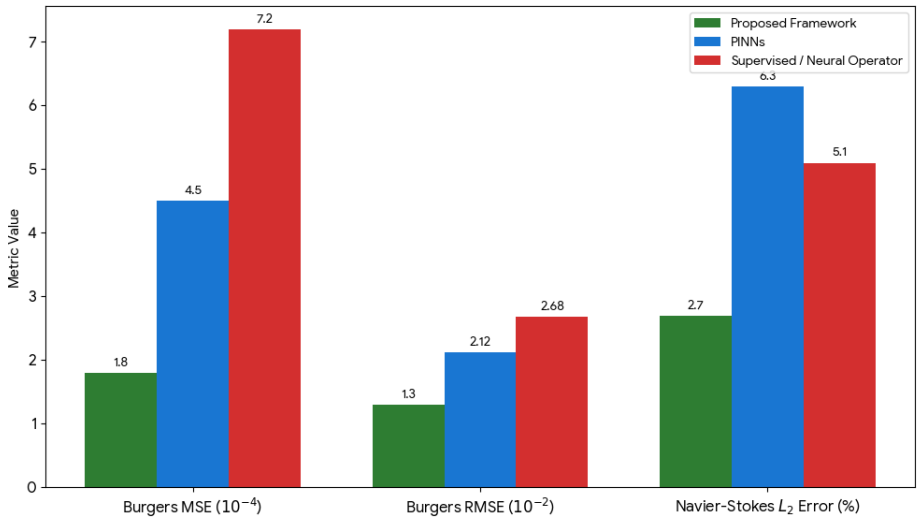


Fig 2. Prediction Error Comparison across Models.

5.2 Data Efficiency Analysis

Experiments are performed to confirm the usefulness of the framework in conditions of data scarcity by changing the proportion of labeled training data and ensuring that it is 10% to 50%. The findings, as presented in Fig 3 indicate that the proposed approach can retain a high performance even when there is limited labelled data. The proposed model performs very well with a 10% training data of 3.2 times 10^{-4}MSE as compared to supervised model which deteriorates severely to 1.1 times 10^{-3}MSE. The faster convergence and lowering of the MSE to 1.5 x 10^{-4}as the data availability rises to 50, compared to 3.9 x $10^{-4}$when using the neural operator without the use of SSL. The sample efficiency curve also substantiates that the suggested framework needs about 40-50 percent of the amount of labelled data in order to attain similar accuracy, which is the effective performance at low-data regimes.

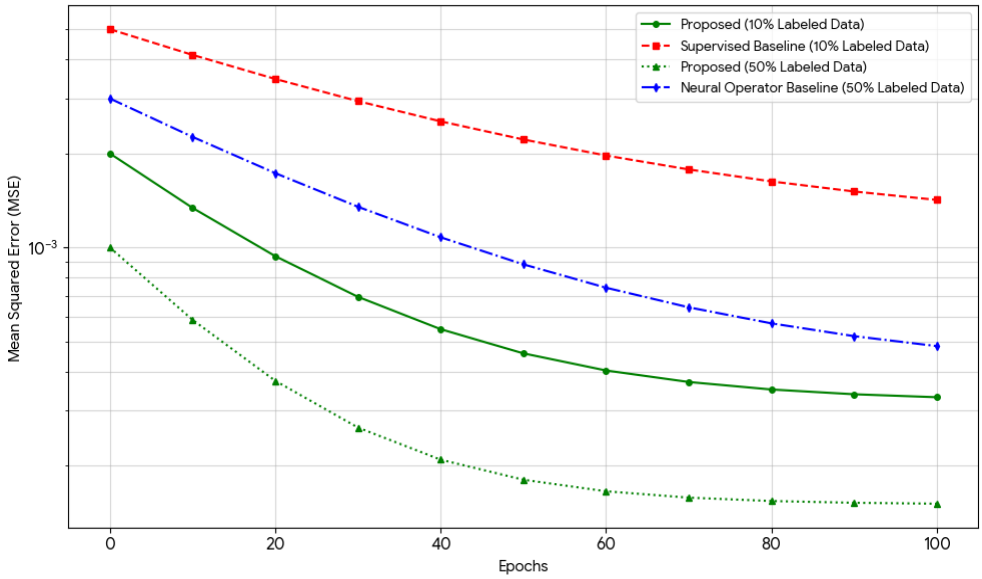


Fig 3. Data Efficiency Performance vs Training Data Size.

5.3 Physics Consistency Evaluation

Physics-based measures are used to determine the physical consistency of the predicted solutions: PDE residual error, violation of boundary conditions, and violation of conservation laws. The proposed model (Fig 4) has a PDE error residual of 2.1×10^{-5},, which is far less than PINNs (6.8×10^{-5},) and supervised models (1.5×10^{-4}). The error of the boundary conditions is minimized to 0.8% versus 2.6% of PINNs, the fact that the constraints are followed is increased. Moreover, violation of the conservation law is reduced to 1.2% showing that the inclusion of physics-based loss is efficient in imposing physical consistency. These findings prove the efficiency of the suggested hybrid learning strategy to not only increase accuracy but come up with physically meaningful predictions.

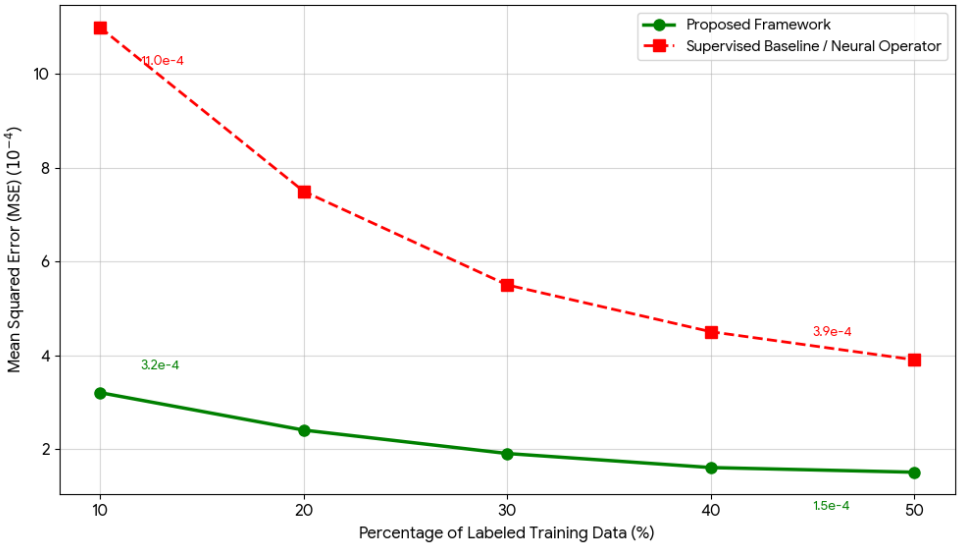


Fig. 8. Physics Consistency Evaluation (PDE Residual Error Comparison).

5.4 Generalization Performance

The proposed framework has its generalization capability on unknown initial and boundary conditions. The findings indicate that the model is very accurate even when it is tested on conditions that did not occur during the training period. In particular, the suggested approach has an MSE of 2.4×10^{-4}on unseen test cases, as opposed to $6.7 \times 10^{-4}$of the supervised model and $4.9 \times 10^{-4}$of PINNs. Moreover, the L2 relative error is less than 3% and this indicates that there is good generalization in the different system configurations. The proposed model has an average inference time of 0.012 seconds/sample, or about 12 times faster than traditional numerical solvers, in terms of computational performance. The training time is also similar to the neural operator models with the advantage that it has a stronger convergence stability because of the self-supervised pretraining. Overall, the findings affirm that the suggested framework is capable of a balanced trade-off between accuracy, efficiency, and physical consistency and is thus ideal in applications of data scarce scientific simulations.

6. Applications

The suggested self-supervised scientific machine learning model shows good suitability in a vast area of real world where simulations are paramount in terms of accuracy and efficiency when the data available is minimal. The real-time simulation systems are one of the main fields of its application, and in many cases, classical numerical solvers are not fast enough to make time-dependent decisions. Through the utilization of learned surrogate models, the proposed methodology allows fast inference at much less computational cost, as required in online monitoring, control system, and interactive simulation systems. One more key use case is in the creation of digital twins, virtual models of the physical system that is continually updated by real-time data. The fact that the proposed framework can be trained with scarce labeled data and be physically consistent renders it specifically suitable to digital twins in manufacturing, aerospace, and intelligent infrastructure industries. It enables the proper estimation of states, predictive maintenance, and scenario study without the need of having large simulation datasets.

Climate and fluid modeling, which requires large-scale data on high-fidelity simulation, are also the focus of the framework which is computationally expensive. The proposed method could be applied in weather prediction, ocean modeling and turbulence analysis by efficiently learning with sparse data and modelling the underlying physical

dynamics. Its good generalization ability also allows good predictions in different environmental conditions making it crucial in the large-scale geophysical simulations. Moreover, the proposed methodology is conducive to the task of engineering design optimization that involves multiple simulations executions to assess design options. The data-efficient learning ability greatly minimizes the requirement of repetitive costly simulations, allowing quicker design space search capabilities. It has found particular application in aerodynamic design, optimization of thermal systems, structural analysis, etc. where quick and precise appraisal is a key factor in enhancing the performance of the system and lowering cost of development.

7. Limitations

Although the self-supervised scientific machine learning framework as suggested worked promisingly, there exist a number of limitations that should be considered. To begin with, the combination of self-supervised learning with physics-based constraints leads to a hybrid model, which is more complex to train. The requirement to minimize two or more goals simultaneously (reconstruction loss, temporal prediction loss, physically-based residuals), adds new computational costs and can require more time to train than simpler models. This may also complicate the process of stabilizing the training process and more so when dealing with very nonlinear systems. Second, the suggested architecture is hyperparameter sensitive, especially on the decisions to balance between self-supervised components and physics-based components of the loss. The coefficients of weighting of the hybrid loss function are important in judgement of the model performance and any lack of suitable tuning can result in suboptimal performance of the model like overfitting to observed trends or lack of adherence to physical constraints. Further, other model architecture and training settings like learning rate and depth of the network need to be tuned carefully to guarantee convergence and stability.

Lastly, the framework has scalability issues when used to high-dimensional PDEs, including complicated fluid dynamics or multi-physics models. Neural operator-based architectures are generalization-improved over traditional ones; however, computational and memory demands are still rising with the problem dimensionality and resolution. This may constrain the model's applicability in large scale or real time models and without additional optimization or hardware acceleration. These issues will be critical in tackling the implementation of the framework to more advanced and industrial-level implementations.

8. Future Work

Although the suggested framework shows a good performance in scientific simulations where the amount of data is scarce, there are multiple prospective directions to consider when improving its performance and expanding its applicability. A major potential direction is the integration of reinforcement learning (RL)-based adaptive self-supervision, where the model is able to change the tasks it self-supervises and the weighting of its loss to dynamically change during training. With an RL-based controller, rather than using predefined pretext goals, one may optimize the choice of masking strategies, prediction horizons, and loss coefficients, resulting in a higher convergence, stability and the overall learning efficiency of complex physical systems. The other notable extension is the 3D multiphysics systems scaling of the framework which is closer to real-life scientific and engineering applications. Although the existing strategy proves to be efficient on standard PDE test cases, applying it to high-dimensional problems with multiple physics (e.g. coupled fluid-thermal or electromagnetic) is another challenge, in both computation and memory consumption. It will also be necessary to design more efficient architectures and parallelized training strategies that will be able to handle these large-scale simulations.

Implementation of the suggested framework on edge devices to perform real-time inferences is a crucial future trend, especially in the cases of digital twins, autonomous systems, and industrial monitoring. Real-time decision-making in resource-constrained settings can be made possible by optimizing the model with low-latency and energy-efficient inference methods like model compression, quantization, and hardware acceleration. This would greatly increase the level of practical feasibility of the framework in a real life situation. Lastly, it is also possible to enhance it by utilizing neural operators together with transformer-based architecture to combine the advantages of operator learning and attention. These hybrid models can learn both the global dependencies, as well as complex spatial-temporal correlations, and can thus exhibit better generalization to a host of different simulation conditions. An investigation of these advanced architectures may open the door to next-generation scientific machine learning architectures that are more scalable, adaptable, and predictive.

CONCLUSION

This paper features a new self-managed scientific machine learning model, aimed at overcoming the issue of data scarcity in highly complex simulation processes. The proposed method with efficiently combined self-supervised learning and physics-aware constraints has a high data-efficient learning ability, which minimizes the role of large labeled datasets. The framework gives better accuracy of simulation with benchmark PDE systems and allows an effective adherence to the governing physical laws by the introduction of physics-based loss functions. In addition, the application of the advanced operator learning improves the generalization and stability at different conditions. In general, the findings demonstrate that the suggested approach is an effective and scalable solution, making it a promising avenue to the next-generation scientific machine learning and data-driven simulation solutions.

REFERENCES

1. Baker, N., Alexander, F., Bremer, T., Hagberg, A., Kevrekidis, Y., Najm, H., ...& Lee, S. (2019). *Workshop report on basic research needs for scientific machine learning: Core technologies for artificial intelligence*. USDOE Office of Science (SC), Washington, DC (United States).
2. Boullé, N., & Townsend, A. (2024). A mathematical guide to operator learning. In *Handbook of Numerical Analysis* (Vol. 25, pp. 83-125). Elsevier.
3. Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., &Piccialli, F. (2022). Scientific machine learning through physics–informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, *92*(3), 88.
4. Duraisamy, K., Iaccarino, G., & Xiao, H. (2019). Turbulence modeling in the age of data. *Annual review of fluid mechanics*, *51*(1), 357-377.
5. Gambietz, M., Dorschky, E., Akat, A., Schöckel, M., Miehling, J., &Koelewijn, A. D. (2025). SSPINNpose: A Self-Supervised PINN for Inertial Pose and Dynamics Estimation. *ACM Transactions on Intelligent Systems and Technology*.
6. Guo, J., Zhu, H., Yang, Y., &Guo, C. (2025). Advances in physics-informed neural networks for solving complex partial differential equations and their engineering applications: A systematic review. *Engineering Applications of Artificial Intelligence*, *161*, 112044.
7. He, J., Liu, X., &Xu, J. (2025). Self-Composing Neural Operators with Depth and Accuracy Scaling via Adaptive Train-and-Unroll Approach. *arXiv preprint arXiv:2508.20650*.
8. Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, *3*(6), 422-440.
9. Li, Z. (2021). Neural operator: Learning maps between function spaces. In *2021 Fall Western Sectional Meeting*. AMS.
10. Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., &Anandkumar, A. (2020). Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*.
11. Lu, L., Jin, P., Pang, G., Zhang, Z., &Karniadakis, G. E. (2021). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature machine intelligence*, *3*(3), 218-229.
12. Pirayeshshirazinezhad, R. (2025). SPINN: An Optimal Self-Supervised Physics-Informed Neural Network Framework. *arXiv preprint arXiv:2509.05886*.
13. 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.
14. Subramanian, S., Kirby, R. M., Mahoney, M. W., &Gholami, A. (2022). Adaptive self-supervision algorithms for physics-informed neural networks. *arXiv preprint arXiv:2207.04084*.
15. Wang, S., Teng, Y., &Perdikaris, P. (2021). Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, *43*(5), A3055-A3081.
16. You, W., Zhou, S., &Meng, X. (2025). Self-supervised neural operator for solving partial differential equations. *arXiv preprint arXiv:2509.00867*.