Foundations of Data Science
Vol. 7, No. 1, March 2025, pp. i-x
doi:10.3934/fods.2024051

SCIENTIFIC MACHINE LEARNING: A SYMBIOSIS

Brendan Keith[1], Thomas O'Leary-Roseberry[2], Benjamin Sanderse[3], Robert Scheichl[4], and Bart van Bloemen Waanders[5]

[1]Brown University, Providence RI, USA

[2]University of Texas at Austin, Austin TX, USA

[3]Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

[4]Heidelberg University, Heidelberg, Germany

[5]Sandia National Laboratories, Albuquerque NM, USA

Abstract. This editorial serves as a preface to the "Scientific Machine Learning" (SciML) special issue of the AIMS Foundations of Data Science journal. In this piece, we contend that SciML exists in a symbiotic relationship with the fields of computational science and engineering (CSE) and machine learning (ML). We highlight the progress (and limitations) of CSE and reflect on the recent successes of ML. While ML creates significant possibilities for advancing simulation techniques, it lacks the mathematical guarantees that are typically found in CSE. We argue that as SciML develops and embraces the remarkable capabilities of ML, it will support, not replace, traditional methods of CSE. We then overview some existing challenges and opportunities in this interdisciplinary field and close by introducing the special issue papers.

1. Introduction.

Mathematical modeling of physical phenomena has been a cornerstone of engineering and the natural sciences for centuries. The field of computational science and engineering (CSE) seeks to operationalize the resulting models through computer simulations based on numerical methods. Pursuing numerical methods has led to fundamental mathematical theories and advances in high-performance computing, creating a reliable paradigm for conducting physics-based computer simulations with rigorous mathematical guarantees. This paradigm has enabled major breakthroughs across disciplines by allowing scientists, engineers, and practitioners to investigate physical phenomena and make decisions about complex systems that would be impossible to achieve through theory and experimentation alone. As a result, CSE has been referred to as the "third pillar" of the scientific enterprise, alongside theory and experimentation [39, 43].

The tenets and impediments of traditional CSE. The key tenets of CSE lie in exploiting two fundamental aspects of mathematical models: *generalizability* and *interpretability* [28]. Newton's law of gravitation is an example of a simple mathematical model that is both generalizable (it generalizes from apples falling on Earth to planets orbiting the Sun) and interpretable (the gravitational force depends on the masses of the objects, their relative distance, and a universal constant). Symplectic integrators [44] operationalize this planetary model (as well as far more complicated mathematical models), delivering accurately simulated orbital motions that preserve key invariants of the true mechanics. As such, symplectic integrators are a prime example of a successful class of numerical methods: they preserve fundamental properties deriving from the original mathematical formulation while providing efficient execution and high-accuracy guarantees.

Over the past century, numerical methods from CSE have reformed prediction, decision-making, and design of complex physical and engineering systems, leading to major technological advances. However, significant challenges can persist even when such powerful tools are available. Indeed, direct simulation with state-of-the-art numerical methods can be infeasible in the following practical scenarios and for the following reasons:

Outer-loop and many-query problems. Tasks involving the design, optimization, uncertainty quantification, or control of complex physical systems require varying inputs and, thus, repeatedly evaluating model outputs. In these problems, the overall computational cost multiplies with the number of samples or iterations.

Scales and physical complexity. While mathematically describable, physical systems that present either many scales or many relevant physical phenomena can lead to systems of equations that are too costly to solve. For example, direct numerical simulations of the Navier–Stokes equations are usually not feasible at high Reynolds numbers without case-by-case modeling assumptions that are less well-understood and not universally agreed upon.

Unknown models. Macroscopic balance laws can typically be derived from known physical principles. However, constitutive laws are often based on empirical observations that may come from imperfect experiments or noisy and incomplete datasets. Moreover, complex systems, such as those governing weather and climate, depend on many interacting factors that are usually not fully observable. In such scenarios, the main challenges, next to designing an appropriate numerical method, are selecting or calibrating the appropriate governing equations followed by quantifying the uncertainty in their predictions.

The advent of machine learning (ML). The parallel developments of explicitly parametrized architectures and data-driven machine learning (ML) techniques have led to revolutionary breakthroughs in the fields of computer vision and natural language processing, as evidenced by self-driving cars [4], foundational large language models such as GPT-4 [2], and bots capable of mastering complex games such as Go [46]. These disruptive technologies have propelled research into applying similar techniques to scientific applications, suggesting a path to overcome the impediments of traditional simulations described above. In particular, ML models that are cheap to evaluate (once trained) could be substitutes or surrogates for expensive physical models [32]. When the model itself is unknown, one can learn effective models from a fusion of experimental and computational data and physical principles [45]. The appeal is undeniable, and a significant ongoing effort has emerged focusing on how to blend ML techniques with traditional CSE frameworks [25].

The symbiosis. Applying the recent advances in ML toward the goals of CSE defines a new and ascendant scientific discipline with remarkable opportunities for innovation and discovery. Borrowing a term from biological sciences, *symbiosis* best describes this interplay, extracting the most beneficial characteristics and capabilities from the two fields. This symbiosis has emerged as a field unto itself: Scientific Machine Learning (SciML).

To realize SciML's full potential, ML's flexible, scalable, nonlinear approximation techniques must be combined with the rich mathematical foundations underlying

CSE. SciML is sure to benefit from advances in ML. Indeed, as better techniques and richer neural network architectures appear in ML, scientists and engineers will benefit from applying them to their problems [7]. On the other hand, SciML should not be viewed as a replacement for traditional CSE. Instead, it can — and should — benefit tremendously from the knowledge and principles underlying classical methods, existing theoretical foundations, and supporting technology (e.g., computing infrastructure and legacy codes). Moreover, traditional numerical methods are often essential or unavoidable in the training of SciML models [15, 49]. Therefore, despite the abundant and warranted interest in SciML, it is essential to maintain a strong focus and investment in the traditional methods of CSE.

BIRS workshop on SciML. In June 2023, a multidisciplinary group of researchers (including the authors) gathered at the Banff International Research Station (BIRS) for a workshop on SciML [26]. The purpose was to exchange ideas, spark collaborations, and secure the foundation of a lasting research community. Attendees of this workshop were invited to submit to this special issue of the AIMS Foundations of Data Science (FoDS) journal. The key themes of the workshop included i) neural network design and approximation theory; ii) modeling, inference, prediction, and data assimilation; and iii) high-performance algorithms and scalability. The workshop participants provided complementary expertise in numerical analysis, approximation theory, functional analysis, probability theory, nonlinear programming, high-performance computing, computer science, statistics, engineering, and industrial applications. The workshop resulted in the following observations, which we elaborate on in the next section:

1. The combination of physics-based and learning-based models provides a unique opportunity for scientific discovery.
2. Techniques from computational science can guide the verification and validation of ML tools, resulting in greater understanding.
3. The enormous power of data-driven models (e.g., large language models) has been made possible through strong advances in software and hardware. SciML requires a similar investment in community software and institutional hardware.
4. Curricula at universities should be designed so that students are well-versed in both mathematics and machine learning topics. Accordingly, these institutions should promote the design of such curricula.

2. Challenges and opportunities.

In this section, we illustrate how a symbiosis of CSE and ML is essential for SciML to tackle persistent and emerging challenges in engineering and the natural sciences. We then discuss new challenges arising from natural and historical differences between CSE and ML, such as their typical computing workflows and preferences for mathematical guarantees. We also highlight the significant opportunities that exist to overcome these challenges, e.g., by promoting reproducible research, joint educational programs, and collaborative efforts.

Grand challenge problems. Our society faces urgent and important challenges, including mitigating the detrimental effects of climate change, achieving affordable and clean energy production, and controlling epidemiological events. For concreteness, consider the challenges of climate change and energy supply. Critical decisions and mitigation strategies will be eventually required to address CO_2 and methane emissions to curtail global warming. Such strategies could be based on an analysis of the underlying dynamics, which is, however, complex and uncertain. To address uncertainties, significant amounts of data are collected on different aspects of the Earth's climate in an attempt to infer model parameters and thereby improve the predictability of the system. Similarly, in the case of energy systems (e.g., solar, wind, fusion), the decision-making process requires model calibration to improve prediction accuracy and to ensure controllability of the system.

In short, these challenge problems lead to a complex combination of physics-driven and data-driven models to be used in outer-loop analyses. The decisions in these applications are often safety-critical or entail dramatic economic consequences, thus requiring reliability and interpretability of computational tools. Significant algorithmic advances are needed in order to overcome the exorbitant costs that would be required by traditional numerical methods to solve such problems. SciML algorithms provide a viable path to aid in predictions for robust policy decisions relating to such complex physical systems. Indeed, techniques at the core of SciML have already had a real-world impact on public and private policy, with examples ranging from the ability to rapidly estimate ensembles of weather predictions [42, 29, 5] to the discovery of 200 million protein structures [24]. The process of scientific discovery is on the verge of a paradigm shift and it is here that the new discipline of SciML is emerging.

However, the unification of CSE and ML in SciML is not straightforward as the two research communities, while having significant overlap, do have distinct differences, for example in terms of software workflow, as well as in their educational backgrounds. Addressing these differences presents opportunities for new approaches to algorithmic development, education, funding, and collaboration, and will be discussed in the subsequent paragraphs.

Computational workflows. One barrier to future research in SciML is that CSE and ML have different software, hardware, and data ecosystem needs. Traditional ML is data-driven, centered around highly optimized automatic differentiation of batched linear algebraic kernels, e.g., computations that exploit homogeneous dense array patterns. ML workflows center around compute ecosystems that leverage massive GPU parallelism across easily accessible, large datasets. On the other hand, CSE is often driven by discretizing mathematical models, leading to complex linear and nonlinear systems, which are solved by methods that exploit the structure of the system (e.g., symmetry and locality). HPC workflows center around bulk synchronous parallel implementations of, e.g., massive-scale solvers and time steppers. As SciML sits at the intersection of these two disciplines, unique issues arise in marrying the different computing workflows. These issues come from structural changes in the hardware landscape (e.g., the rise of GPU computing), increasing energy and memory costs of modern computing, and challenges related to constructing sustainable computing environments for the multifaceted software workflows required by this naturally interdisciplinary field. Below, we delineate some of these challenges and note related opportunities.

The success of ML and SciML research, in conjunction with the more favorable FLOP/Watt cost characteristics of GPUs, has led to a significant proliferation of GPU-based systems. This structural change in the hardware landscape creates both challenges and opportunities for SciML and CSE research. One path forward is to port legacy CSE codes to this new GPU-dominated computing ecosystem. This was a major focus for the exascale computing project (ECP) [35] in the US, mirrored by similar initiatives in Europe. However, this has only been partially successful and significant amounts of CSE infrastructure are still (and will likely remain) CPU dependent. While the changing hardware landscape continues to create new opportunities to (re-)develop traditional numerical methods for deployment on GPUs (e.g., by exploiting locality, reformulating computations to leverage homogeneous array operations, or avoiding communication dependencies [1, 17, 27, 50, 51]) the major challenge will lie in providing infrastructure for efficiently interfacing GPU-based ML tools and CPU-based CSE codes. As several contributions at the BIRS workshop highlighted, much of modern SciML research (e.g., surrogates and multilevel/multifidelity approaches) is directed precisely at this challenge.

An important issue for all modern ML workflows concerns their increasing energy and memory costs, which often outstrip gains in hardware. It is well known that the data, energy and memory needs of state-of-the-art ML models (e.g., large language models) are growing at rates that will create issues in the future [23, 16]. This same issue will likely arise in SciML as its models and methods mature. In order to mitigate the coming burdens, it is essential to develop compute workflows that emphasize energy and memory efficiency. This can be achieved by pursuing sparse model representations, e.g., via neural network pruning [10]. Models that are designed to be efficient (e.g., independent of the grid dimension) are particularly favorable as the scales of simulations increase.

Mathematical guarantees. A challenge for achieving symbiosis arises from the different foci of CSE-based (physics-driven) and ML-based (data-driven) methods. The main focus of CSE is operationalizing numerical methods for the simulation of physical systems through the use of mathematical models. Examples of these numerical methods include the finite element [11, 34, 14, 38], finite difference [12], integral equation [37], and spectral methods [18] for solving partial differential equations (PDEs), alongside density functional theory and ab initio methods for quantum mechanical simulations [22, 41], and N-body simulations for particle systems [19]. The hallmark of these celebrated methods is the mathematical guarantees that have been rigorously established in their development. These mathematical models are generalizable and interpretable, and the associated numerical methods are equipped with rigorous theories of convergence, stability, consistency, and complexity [30].

In contrast, data-driven ML methods typically do not possess such guarantees. While universal approximation results are established for many ML methods, specific theories of rates of convergence and stability are much less common. Realizing convergent approximations is inhibited by the NP-hardness of the associated nonconvex optimization problems utilized to train ML methods [13]. Naïve adoption of ML methods thus introduces risk into important applications. It is important for future research in SciML to attempt to develop methods that are equipped with certain mathematical guarantees that are fundamental to CSE. Doing so will invariably require adapting techniques and theories from CSE to the ML setting or developing new mathematical tools. In addition, when such mathematical guarantees are not attainable, a possible symbiotic approach that combines SciML and CSE methods can be used to accelerate simulations while still maintaining asymptotic exactness. This can be achieved, for example, via rigorous a posteriori error estimation techniques [3], predictor-corrector methods and hybrid approaches, which have already been employed for forward prediction [40, 21] and Bayesian inverse problems [9, 8].

Reproducibility. A valuable opportunity to address the aforementioned challenges and develop trust in SciML is to ensure that results can be independently verified, confirming the validity and reliability of new findings. Reproducible and transparent open-source implementations also allow other researchers to build upon existing work, fostering collaborative advancements and the development of more robust and effective models and methods. To promote confidence in new research, leading voices have advocated for following strict scientific standards in reproducibility and verifiability [6, 25], such as following the Findable, Accessible, Interoperable, and Reusable (FAIR) data principles [48]. The community has responded by embracing benchmark datasets such as PDEArena [20], PDEBench framework [47], and Mechanical MNIST [31, 36] and conducting independent reviews of recent works [33]. We wish to encourage further datasets, protocols, and reviews, in particular, those that emphasize downstream tasks such as inference, prediction, and decision-making. This includes additional protocols for creating, sharing, and storing potentially large datasets that can constitute future community benchmarks in this rapidly developing field. It is important to have a clear view of the performance of new methods in these tasks as they are often used to motivate new research in SciML.

Education, collaboration, and funding. The background required to effectively conduct research in SciML is challenging because of the need to access a wide range of technologies, including mathematics, engineering, data science, machine learning, statistics, probability, and scientific computing. Moreover, if SciML is positioned to address decision-making for complex applications, a hierarchy of algorithms is required. These issues raise multiple questions: How do we educate students? How do we incorporate additional topics in already congested curricula? Can existing programs be extended, or should completely new degrees be designed? Can departments be organized to educate students within collaborative settings? And finally, what are the right and sustainable hiring strategies for faculty positions?

The success of ML has created significant enthusiasm among younger generations, which suggests that rethinking academic programs to foster this interest is timely. However, undergraduate engineering, science, mathematics, and computer science programs have little room to consider adding courses to address these needs. At the same time, simply replacing important fundamental mathematics courses with new courses on data-driven approaches is also not the right solution. Curricula need to be carefully modernized, while completely new undergraduate programs could also be considered by extracting cross-cutting elements from different educational programs. A third strategy could exclusively rely on graduate programs that build on standard undergraduate degrees to supplement the students' knowledge with key missing material. All of these options will require leveraging expertise across multiple departments and establishing close, multidisciplinary collaborations.

The required academic transformation will likely not happen immediately because of a shortage of educators in SciML. Instead, it will require an iterative and potentially multigenerational process, and academia will face challenges in finding suitable educators in the short term, regardless of how individual institutions decide on new curricula. Partnerships with industry and national laboratories through collaborations, internships, and post-doc programs could help accelerate the process. Of these options, national laboratories are particularly well-suited to help fill knowledge gaps in industrial applications and high-performance computing. Moreover, funding agencies can play a pivotal role in creating calls for proposals that encourage collaborative themes, as well as support for innovative educational infrastructure.

3. **Overview of the special issue.** This special issue was motivated by the BIRS workshop and features papers that cover a range of research contributions, including novel methodologies, theoretical analyses, applications to complex systems, and a review of state-of-the-art SciML methods. These works exemplify the symbiotic relation of CSE and ML, illustrating many exciting new research areas in the field of SciML. We list the papers appearing in the special issue below.

1. "GNEP based dynamic segmentation and motion estimation for neuromorphic imaging" by Harbir Antil and David Sayre.
2. "A system identification approach for non-intrusive reduced order modeling of radiation-induced photocurrents" by Pavel Bochev and Biliana Paskaleva.
3. "An over complete deep learning method for inverse problems" by Moshe Eliasof, Eldad Haber and Eran Treister.
4. "A practical existence theorem for reduced order models based on convolutional autoencoders" by Nicola Rares Franco and Simone Brugiapaglia.
5. "Hyper-differential sensitivity analysis with respect to model discrepancy: Posterior optimal solution sampling" by Joseph Hart and Bart van Bloemen Waanders.
6. "Stacked networks improve physics-informed training: Applications to neural networks and deep operator networks" by Amanda A. Howard, Sarah H. Murphy, Shady E. Ahmed and Panos Stinis.
7. "An operator learning perspective on parameter-to-observable maps" by Daniel Zhengyu Huang, Nicholas H. Nelsen and Margaret Trautner.
8. "Heterogeneous peridynamic neural operators: Discover biotissue constitutive law and microstructure from digital image correlation measurements" by Siavash Jafarzadeh, Stewart Silling, Lu Zhang, Colton Ross, Chung-Hao Lee, S. M. Rakibur Rahman, Shuodao Wang and Yue Yu.
9. "Multifidelity linear regression for scientific machine learning from scarce data" by Elizabeth Qian, Dayoung Kang, Vignesh Sella and Anirban Chaudhuri.
10. "Scientific machine learning for closure models in multiscale problems: A review" by Benjamin Sanderse, Panos Stinis, Romit Maulik and Shady E. Ahmed.
11. "Reduced basis approximations of parameterized dynamical partial differential equations via neural networks" by Peter Sentz, Kristian Beckwith, Eric C. Cyr, Luke N. Olson and Ravi Patel.
12. "Neural network approaches for parameterized optimal control" by Deepanshu Verma, Nick Winovich, Lars Ruthotto, and Bart van Bloemen Waanders
13. "Deep learning enhanced cost-aware multi-fidelity uncertainty quantification of a computational model for radiotherapy" by Piermario Vitullo, Nicola Rares Franco, and Paolo Zunino.
14. "Unsupervised physics-informed disentanglement of multimodal data" by Elise Walker, Nathaniel Trask, Carianne Martinez, Kookjin Lee, Jonas A. Actor, Sourav Saha, Troy Shilt, Daniel Vizoso, Remi Dingreville, and Brad L. Boyce.

Guest Editors:

Lianghao Cao
California Institute of Technology Pasadena CA, USA
lianghao@caltech.edu

Brendan Keith
Brown University, Providence RI, USA

brendan_keith@brown.edu

Lu Lu
Yale University, New Haven CT, USA
lu.lu@yale.edu

Thomas O'Leary-Roseberry
University of Texas at Austin, Austin TX, USA
tom.olearyroseberry@utexas.edu

REFERENCES

[1] A. Abdelfattah, V. Barra, N. Beams, R. Bleile, J. Brown, J.-S. Camier, R. Carson, N. Chalmers, V. Dobrev, Y. Dudouit, et al., Gpu algorithms for efficient exascale discretizations, *Parallel Computing*, **108** (2021), 102841.

[2] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat, et al., GPT-4 technical report, arXiv preprint, `arXiv:2303.08774`, 2023.

[3] M. Ainsworth and J. T. Oden, A posteriori error estimation in finite element analysis, *Computer Methods in Applied Mechanics and Engineering*, **142** (1997), 1-88.

[4] C. Badue, R. Guidolini, R. V. Carneiro, P. Azevedo, V. B. Cardoso, A. Forechi, L. Jesus, R. Berriel, T. M. Paixao, F. Mutz, et al., Self-driving cars: A survey, *Expert Systems with Applications*, **165** (2021), 113816.

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

[6] M. P. Brenner and P. Koumoutsakos, Machine learning and physical review fluids: An editorial perspective, *Physical Review Fluids*, **6** (2021), 070001.

[7] E. Calvello, N. B. Kovachki, M. E. Levine and A. M. Stuart, Continuum attention for neural operators, arXiv preprint, `arXiv:2406.06486`, 2024.

[8] L. Cao, T. O'Leary-Roseberry and O. Ghattas, Derivative-informed neural operator acceleration of geometric MCMC for infinite-dimensional Bayesian inverse problems, arXiv preprint, `arXiv:2403.08220`, 2024.

[9] L. Cao, T. O'Leary-Roseberry, P. K. Jha, J. T. Oden and O. Ghattas, Residual-based error correction for neural operator accelerated infinite-dimensional bayesian inverse problems, *Journal of Computational Physics*, **486** (2023), 112104, 33 pp.

[10] H. Cheng, M. Zhang and J. Q. Shi, A survey on deep neural network pruning: Taxonomy, comparison, analysis, and recommendations, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **46** (2024).

[11] R. Courant, Variational methods for the solution of problems of equilibrium and vibrations, *Lecture Notes in Pure and Applied Mathematics*, **49** (1943), 1-23.

[12] R. Courant, K. Friedrichs and H. Lewy, Über die partiellen differenzengleichungen der mathematischen physik, *Mathematische Annalen*, **100** (1928), 32-74.

[13] M. Danilova, P. Dvurechensky, A. Gasnikov, E. Gorbunov, S. Guminov, D. Kamzolov and I. Shibaev, Recent theoretical advances in non-convex optimization, In *High-Dimensional Optimization and Probability: With a View Towards Data Science*, Springer, 2022, 79-163.

[14] C. A. Felippa and R. W. Clough, The finite element method in solid mechanics, In *SIAM-AMS Proceedings*, American Mathematical Society, **2** (1970), 210-252.

[15] O. Ghattas and K. Willcox, Learning physics-based models from data: Perspectives from inverse problems and model reduction, *Acta Numerica*, **30** (2021), 445-554.

[16] A. Gholami, Z. Yao, S. Kim, C. Hooper, M. W. Mahoney and K. Keutzer, AI and memory wall, *IEEE Micro*, **44** (2024), 33-39.

[17] P. Ghysels and R. Synk, High performance sparse multifrontal solvers on modern gpus, *Parallel Computing*, **110** (2022), 102897, 11 pp.

[18] D. Gottlieb and S. A. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*, SIAM, 1977.

[19] L. Greengard and V. Rokhlin, A fast algorithm for particle simulations, *Journal of Computational Physics*, **73** (1987), 325-348.

[20] J. K. Gupta and J. Brandstetter, Towards multi-spatiotemporal-scale generalized PDE modeling, arXiv preprint, `arXiv:2209.15616`, 2022.

[21] Q. He, M. Perego, A. A. Howard, G. E. Karniadakis and P. Stinis, A hybrid deep neural operator/finite element method for ice-sheet modeling, *Journal of Computational Physics*, **492** (2023), 112428, 17 pp.

[22] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, *Physical Review*, **136** (1964), B864-B871.

[23] P. Jiang, C. Sonne, W. Li, F. You and S. You, Preventing the immense increase in the life-cycle energy and carbon footprints of llm-powered intelligent chatbots, *Engineering*, **40** (2024), 202-210.

[24] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al., Highly accurate protein structure prediction with alphafold, *Nature*, **596** (2021), 583-589.

[25] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang and L. Yang, Physics-informed machine learning, *Nature Reviews Physics*, **3** (2021), 422-440.

[26] B. Keith, T. O'Leary-Roseberry, L. Lu, S. Mishra and Z. Mao, Workshop report 23w5129 scientific machine learning, 2023.

[27] T. Kolev, P. Fischer, M. Min, J. Dongarra, J. Brown, V. Dobrev, T. Warburton, S. Tomov, M. S. Shephard, A. Abdelfattah, et al., Efficient exascale discretizations: High-order finite element methods, *The International Journal of High Performance Computing Applications*, **35** (2021), 527-552.

[28] J. N. Kutz and S. L. Brunton, Parsimony as the ultimate regularizer for physics-informed machine learning, *Nonlinear Dynamics*, **107** (2022), 1801-1817.

[29] R. Lam, A. Sanchez-Gonzalez, M. Willson, P. Wirnsberger, M. Fortunato, F. Alet, S. Ravuri, T. Ewalds, Z. Eaton-Rosen, W. Hu, et al., Learning skillful medium-range global weather forecasting, *Science*, **382** (2023), 1416-1421.

[30] P. D. Lax and R. D. Richtmyer, Survey of the stability of linear finite difference equations, *Comm. Pure and Appl. Math.*, **9** (1956), 267-293.

[31] E. Lejeune, Mechanical MNIST: A benchmark dataset for mechanical metamodels, *Extreme Mechanics Letters*, **36** (2020), 100659.

[32] D. Luo, T. O'Leary-Roseberry, P. Chen and O. Ghattas, Efficient PDE-constrained optimization under high-dimensional uncertainty using derivative-informed neural operators, arXiv preprint, `arXiv:2305.20053`, 2023.

[33] N. McGreivy and A. Hakim, Weak baselines and reporting biases lead to overoptimism in machine learning for fluid-related partial differential equations, *Nature Machine Intelligence*, **6** (2024), 1256-1269.

[34] R. J. Melosh, Basis for derivation of matrices for the direct stiffness method, *AIAA Journal*, **1** (1963), 1631-1637.

[35] P. Messina, The exascale computing project, *Computing in Science & Engineering*, **19** (2017), 63-67.

[36] S. Mohammadzadeh, P. Prachaseree and E. Lejeune, Investigating deep learning model calibration for classification problems in mechanics, *Mechanics of Materials*, **184** (2023), 104749.

[37] R. S. Nicholson, Some examples of the numerical solution of nonlinear integral equations, *Analytical Chemistry*, **37** (1965), 667-671.

[38] J. T. Oden, *Finite Elements of Nonlinear Continua*, Courier Corporation, 2006.

[39] J. T. Oden, The Third Pillar: The Computational Revolution of Science and Engineering, *Commemorative Lecture at the 34th Honda Prize Award Ceremony on the 18th November 2013*, 2013.

[40] V. Oommen, K. Shukla, S. Desai, R. Dingreville and G. E. Karniadakis, Rethinking materials simulations: Blending direct numerical simulations with neural operators, *npj Computational Materials*, **10** (2024), 145.

[41] R. G. Parr, D. P. Craig and I. G. Ross, Molecular orbital calculations of the lower excited electronic levels of benzene, configuration interaction included, *The Journal of Chemical Physics*, **18** (1950), 1561-1563.

[42] J. Pathak, S. Subramanian, P. Harrington, S. Raja, A. Chattopadhyay, M. Mardani, T. Kurth, D. Hall, Z. Li, K. Azizzadenesheli, P. Hassanzadeh, K. Kashinath and A. Anandkumar, *FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators*, Feb. 2022.

[43] D. A. Reed, R. Bajcsy, M. A. Fernandez, J.-M. Griffiths, R. D. Mott, J. Dongarra, C. R. Johnson, A. S. Inouye, W. Miner, M. K. Matzke, et al., Computational science: Ensuring america's

competitiveness, *United States. President's Information Technology Advisory Committee. National Coordination Office for Information Technology Research & Development*, 2005.

[44] R. D. Ruth, A canonical integration technique, *IEEE Trans. Nucl. Sci.*, **30** (1983), 2669-2671.

[45] B. Sanderse, P. Stinis, R. Maulik and S. E. Ahmed, Scientific machine learning for closure models in multiscale problems: A review, *Foundations of Data Science*, 2024.

[46] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, et al., Mastering the game of Go without human knowledge, *Nature*, **550** (2017), 354-359.

[47] M. Takamoto, T. Praditia, R. Leiteritz, D. MacKinlay, F. Alesiani, D. Pflüger and M. Niepert, PDEBench: An extensive benchmark for scientific machine learning, *Advances in Neural Information Processing Systems*, **35** (2022), 1596-1611.

[48] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al., The FAIR guiding principles for scientific data management and stewardship, *Scientific Data*, **3** (2016), 1-9.

[49] K. E. Willcox, O. Ghattas and P. Heimbach, The imperative of physics-based modeling and inverse theory in computational science, *Nature Computational Science*, **1** (2021), 166-168.

[50] A. Yesypenko and P.-G. Martinsson, GPU optimizations for the hierarchical Poincaré-Steklov scheme, In *International Conference on Domain Decomposition Methods*, Springer, 2022, 519-528.

[51] A. Yesypenko and P.-G. Martinsson, SlabLU: A two-level sparse direct solver for elliptic PDEs, *Advances in Computational Mathematics*, **50** (2024), 90, 28 pp.