A Multi-Industry Exploration of Model Flexibility and Performance Trade-offs in the Era of Artificial Intelligence and Advanced Computing

by

Caroline C. Warren

S.B., Mathematics and Computer Science, Economics, and Data Science
Massachusetts Institute of Technology (2023)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING IN COMPUTER SCIENCE, ECONOMICS, AND DATA SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2024

Authored by: Caroline C. Warren
Department of Electrical Engineering and Computer Science
May 17, 2024

Certified by: Neil C. Thompson
Research Scientist, MIT CSAIL, Thesis Supervisor

Accepted by: Katrina LaCurts
Chair, Master of Engineering Thesis Committee

A Multi-Industry Exploration of Model Flexibility and Performance Trade-offs in the Era of Artificial Intelligence and Advanced Computing

by

Caroline C. Warren

Submitted to the Department of Electrical Engineering and Computer Science
on May 17, 2024 in partial fulfillment of the requirements for the degree of

MASTER OF ENGINEERING IN COMPUTER SCIENCE, ECONOMICS, AND DATA SCIENCE

ABSTRACT

The evolution of advanced computing, driven by breakthroughs in artificial intelligence and large language models, presents significant opportunities for various industries. In this study, we analyze the trade-off between model performance and computational cost to understand industry-specific preferences and technology adoption dynamics. We construct a dataset of 150 published research papers that compare traditional machine learning, deep learning, and scientific computing models. Using both binary and relative comparison metrics, we assess improvements in performance and computational cost. We find that the healthcare industry prioritizes model accuracy over computational cost, with 40% of papers showing performance improvements but only 34.29% indicating cost efficiency. In contrast, the architecture industry demonstrates a significant focus on reducing computational costs, with 94.29% of papers reporting cost improvements but only 8.57% showing performance gains. The finance industry balances both aspects, with a preference for minimizing computational complexity, with 31.43% of papers showing performance improvements and 80% reporting cost reductions. We also find an exponential increase in publications relevant to this study over time, suggesting a rapidly evolving landscape in advanced computing.

Thesis supervisor: Neil C. Thompson
Title: Research Scientist, MIT CSAIL

Acknowledgments

This work is part of a broader project on the landscape of machine learning with the Director of FutureTech Dr. Neil Thompson, EECS PhD student Gabriel Filipe Manso Araujo, and fellow Masters student Haley Nakamura. Neil provided the direction for the project and Gabriel, Haley, and I worked to gather academic papers for our dataset, structure our dataset, and review papers. Using the data we collected, I performed the analysis and wrote this paper.

First, I would like to thank my thesis advisor Dr. Neil Thompson. Neil's feedback and direction during each stage of the project - from proposing a hypothesis to suggesting ways to refine my analysis has been invaluable. I feel incredibly fortunate to have had the chance to join Neil's lab and contribute to cutting-edge research in advanced computing within a supportive and collaborative environment.

I would also like to thank Gabriel and Haley for being incredible research partners over the past year and a half. Gabriel provided mentorship and insights that were essential for my progression as a researcher. I am especially thankful for the kindness that he has shown me during our time working together. Haley provided an innovative mindset that helped elevate the project with her ideas for streamlining data collection and paper reviewing processes through automation. I would like to thank her for being a wonderful research partner and friend.

To my friends, thank you for making my time at MIT truly unforgettable. The people are what make the place and I am forever grateful for the lifelong friendships that I have built at the institute. I would like to especially thank Devin, Helen, and my roommates Claire and Alice for their support over this past year.

Lastly, I would like to thank my family - Mom, Dad, and Peter for your unwavering love and support during my time at MIT. I would not be who I am today without you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The objective of this work is to bridge the gap between industry-specific applications and the advanced computing frontier by investigating where different industries believe the trade-off between performance and compute lies. Specifically, we aim to provide a comprehensive review of how the healthcare, architecture, and finance industries value model performance compared to flexibility. To do this, we analyze 150 published research papers across a wide variety of domains to extract data on the performance and computational cost of advanced computing models in practice. By understanding how different industries value performance and flexibility, we can make important predictions about the future of such industries in the age of artificial intelligence.

The remainder of this paper is organized as follows: Chapter 1 provides an introduction and reviews related works, Chapter 2 describes the methodology that was used for data extraction and presents the structure of the dataset we built. Chapter 3 explains how metrics were calculated using our dataset. Chapter 4 presents results and analysis. Finally, Chapter 5 summarizes our findings and discusses limitations of the current study as well as future research directions.

1.1 Trade-off Between Model Performance and Computational Cost

As the landscape of advanced computing evolves at a breakneck rate, firms across all industries are confronted with a pivotal question: how can they effectively incorporate new technologies into their business models? Recent breakthroughs in artificial intelligence and the widespread adoption of large language models highlight the potential that advanced computing has to impact a diverse set of industries. However, with great promise comes great challenges, as there are substantial computational, economic, and environmental costs associated with implementing and deploying such technologies [17]. When thinking about decisions that must be made relating to technology adoption, we face a conundrum: having more computing power tends to improve performance, and vice versa [2]. Discovering the optimal balance between computational cost and performance is an important step towards more widespread adoption. However, the real-world impact of these technologies varies significantly across different domains, making the situation more complex.

Understanding the trade-off between computational cost and performance is critical to understanding the implications for how industries strategize and compete. Many industries that have traditionally relied on high-performance computing models are now at a crossroads. The adoption of more computationally efficient models can lead to significant cost savings, faster decision-making processes, and improved scalability. However, these benefits must be weighed against the potential risks of reduced model accuracy and reliability. We hypothesize that each industry has a different optimal point of indifference between the two performance and computational cost and we seek to understand how the healthcare, architecture, and finance industries specifically view this trade-off.

For decades, scientific computing simulations have been the preferred method of modeling phenomena across a wide range of applications [19] Simulations rely on models built by experts in a field to predict probabilities and incorporate nuanced knowledge of physical systems to reach conclusions with a high degree of accuracy. These complex models perform extremely well but oftentimes the computational power required to initialize and run each simulation step can make applications impractical [18]. In terms of performance, simulations are commonly thought of as the gold standard of prediction models, a title that comes at the expense of computational cost. In more recent years, traditional machine learning has emerged as a leading technology in the pursuit towards predictions that are less computationally expensive traditional large-scale simulations. Rather than using a model carefully constructed by experts, machine learning models rely on the patterns found in historical data to make predictions rather than the physical laws that explain complex systems. In the early 2010s, deep learning, a subset of machine learning that uses artificial neural network architecture, emerged with the potential to strike a middle ground between performance and compute.

We are posed with a multi-objective optimization problem: how can we build models that maximize performance metrics and minimize computational costs? We utilize the idea of a Pareto frontier to describe the trade-off between performance and compute. When thinking about the performance and computational cost of different models, we can imagine each model corresponding to a coordinate on a 2-dimensional grid, with the x-axis representing the amount of computation used and the y-axis representing the performance of the model. When examining the data, we can then identify points that are Pareto optimal, meaning that there are no alternative points that make one metric (e.g. performance) better without making another metric (e.g. computational cost) worse. The set of all Pareto optimal points can then be plotted to create what is known as the Pareto frontier. Using our knowledge of the performance and computational cost of traditional machine learning, deep learning, and scientific computing techniques, the shape of a Pareto frontier that explains the trade-off between performance and compute can be hypothesized. Figure 1.1 describes our hypothesis.

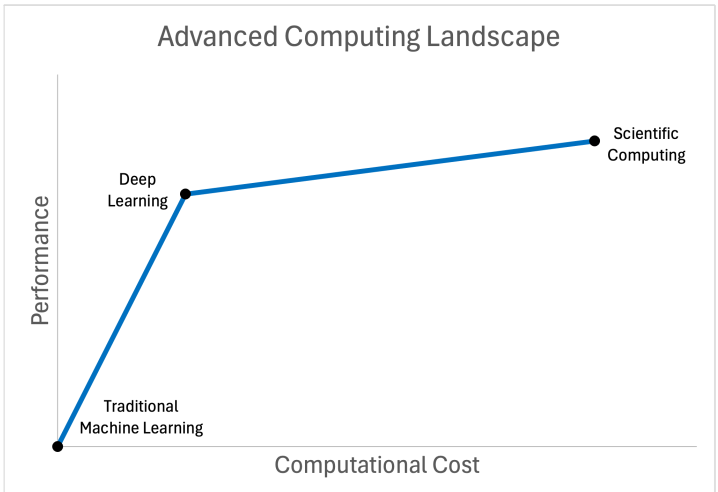


Figure 1.1: Hypothesized Pareto frontier of the advanced computing landscape (example of the predicted shape of the frontier, does not use real data to plot).

1.2 Related Work

Previous works provide overviews of the adoption of advanced computing technologies across a variety of industries and emphasize the importance of minimizing computational cost. Through the analysis of both the reports of individual models and surveys of adoption across a broader range of the landscape, we can understand what has been done and identify important extensions on the existing work in this field.

1.2.1 Adoption Across Firms

Many studies have found that advancements in technology are key drivers of economic growth [4]. Therefore, understanding how advanced technology adoption can be measured and encouraged is imperative for sustaining growth and planning for the future. Using evidence from the U.S. Census Bureau's 2018 Annual Business Survey, a 2020 paper from the National Bureau of Economic Research was able to categorize adoption at the firm level. Key takeaways from the study include the idea that the adoption of advanced computing exhibits a hierarchical pattern, with most firms that adopt AI or other advanced business technologies also using more widely diffused technologies and that very few firms are at the technology frontier [12].

1.2.2 Industry-Level Analysis

Previous works that focus on the industry-level analysis of the impact of advanced computing identify industries that have the potential to benefit most from advances in artificial intelligence. With data sources widely available, studies have found promising applications of machine learning and deep learning in smart cities, healthcare, and supply chain management [11]. However, past works do not investigate how different industries weigh the two objectives of maximizing performance and minimizing computational cost, with most studies only mentioning one or the other.

1.2.3 The Growing Importance of Computational Cost

Along with the clear time saving and money saving properties of computationally inexpensive models, recent works highlight additional motivations for minimizing the amount of computation that models require for satisfactory performance. In their 2022 paper, Thompson et al. study the importance of computing power from the lens of economically important applications. Despite economic theory assuming a power law relationship between inputs and outputs, they find that exponential increases in computing power are needed to get linear improvements in the domains of weather prediction, protein folding, and oil exploration [2]. In 2019, Strubell et al. quantified the environmental damage associated with running computationally expensive models, finding that the estimated carbon dioxide emissions from training one large Transformer model is nearly five times that of the estimated carbon dioxide emissions that a car emits over its lifetime [1].

Chapter 2

Data

2.1 Data Collection

While theoretical advances in computing are well documented, information about applications of such technologies are far less centralized. To investigate how various industries value the performance and flexibility of advanced computing models, we review published papers that compare traditional machine learning, deep learning, and scientific computing models. We use each paper as data point that represents a specific instance of model performance and computational cost within each industry. Of the papers reviewed in each domain, we chose to only include those that provided both performance and computational cost comparison metrics for both models.

Three different types of comparisons were made across the 150 papers used in this study. They are structured in the format Base $\rightarrow$ New, representing the switch from the original baseline model to a novel model. Papers are labeled as having compared traditional machine learning to deep learning (ML $\rightarrow$ DL), scientific computing to deep learning (SC $\rightarrow$ DL), or scientific computing to traditional machine learning (SC $\rightarrow$ ML). The distribution of papers between these three categories is satisfactory for performing further analysis, with a minimum of 29 papers in each Base $\rightarrow$ New category. The full count of papers is shown in Table 2.1.

2.2 Industry Identification

When considering different industries to prioritize collecting papers for, it was critical to pick industries that had the potential for enough data to be collected. Another major consideration was finding industries that historically have varying degrees of preference for

Industry	**Number of Papers**	**ML → DL**	**SC → DL**	**SC → ML**
Healthcare	35	28	5	2
Architecture	35	5	13	17
Finance	35	10	18	7
Other	45	8	34	3
Total	**150**	**51**	**70**	**29**

Table 2.1: Count of total papers and paper with each type of comparison broken down by industry.

the trade off between model performance and computational cost. For example, one industry may be willing to take more time to let models run in order for the most accurate result to be produced, while other industries may need real time results for the models to have any tangible impact on their work. With these considerations in mind, we chose to prioritize data collection for the healthcare, architecture, and finance industries. Below, we discuss further the decision to choose the healthcare, architecture, and finance industries.

2.2.1 Healthcare Industry

Strides in artificial intelligence have proven to be invaluable for the healthcare industry. Despite the limitations in data quality and high regulatory standards that slow the broad adoption of advanced computing in this industry, life-saving technologies have emerged for a wide-range of conditions [7]. Additionally, the healthcare industry has the second highest sectoral diffusion rate for all business technologies, with 14% of firms indicating the use of at least one business technology, behind only the manufacturing industry at 15% [12].

Accuracy, reliability, and privacy are all critical metrics to prioritize in the healthcare industry due to the direct impact that models have on patient outcomes and regulatory requirements. Given the nature of the industry, we hypothesize the healthcare industry may favor higher computational costs if model accuracy and reliability is improved, even if only slightly. An incorrect prediction in this industry can have a very high cost to both the patient and the healthcare provider, justifying an investment in more advanced computational resources. For example, it is critical for healthcare providers to minimize false negatives when using a disease prediction algorithm and it is usually reasonable for patients to wait hours or even days for results so long as they are accurate.

Within the healthcare industry, 35 total papers with relevant data for this work were found. Table 2.2 shows a breakdown of the subdomains of which the papers belong.

Healthcare Subdomains	Paper Count
Disease prediction	23
Clinical Data	9
Medical Imaging	3
Total	**35**

Table 2.2: Breakdown of the number of papers in the dataset from the healthcare industry by subdomain.

The most frequent subdomain is disease prediction, representing some of the most promising applications of advanced computing in healthcare. Studies have found that the use of machine learning-based computational models for disease prediction can reduce the time and resources required for analysis significantly [45]. .

While some studies have found that artificial intelligence models can outperform medical professionals [21], these conclusions are not universally accepted. Other research suggests that most models have lower accuracy metrics and should support serve as support for medical professionals rather than as substitutes [3].

2.2.2 Architecture Industry

Despite its title as one of the oldest industries in the world and its origins tracing back to the Neolithic period (10000 BC), the architecture industry has benefited significantly from recent advancements in computing. The use of scientific computing and simulations is well documented within architecture, with large simulations for structural integrity, energy efficiency modeling, and generative design algorithms commonplace across firms in the industry [22]. However, the rise of machine learning has proven to be particularly beneficial in architecture for understanding interactions between systems, environments, and inhabitants, and across different disciplines by observing recurring events in a more precise, efficient, and innovative way [14].

We hypothesize that researchers in the architecture industry value model performance and computational complexity to a similar degree. It is crucial for models to operate quickly and accurately, enabling timely and effective decision-making in design and construction processes. While minor errors can occasionally be tolerated, the overarching goal is to ensure that the benefits and cost savings realized over the lifespan of a building outweigh the initial computational investments. This balance is vital for driving innovation while maintaining economic feasibility in architectural practices.

Within the architecture industry, 35 total papers with relevant data for this work were found. Table 2.3 shows a breakdown of the subdomains of which the papers belong.

Architecture Subdomains	**Paper Count**
Building Design	31
Landscape Architecture	4
Total	**35**

Table 2.3: Breakdown of the number of papers in the dataset from the architecture industry by subdomain.

The most frequent subdomain is building design, a category that encompasses building energy assessments, the design and construction of building structures, and project planning. Of the 31 building design papers represented in our dataset, 17 (54.8%) are SC $\rightarrow$ ML comparisons, and 11 (35.5%) are SC $\rightarrow$ DL comparisons, suggesting that machine learning models may be well on their way to replacing traditional simulations in an industry that has historically relied very heavily on them.

2.2.3 Finance Industry

Advanced computing has a rich history in the finance industry, dating back to the 1960s when mainframe computers were used to automate banking operations and manage large datasets. The 1980s and 1990s saw the development of algorithmic trading and quantitative finance, where complex mathematical models were used to predict market trends and optimize investment strategies [20]. From the early 2000s to current date, computing in the industry has been characterized by the integration of big data analytics and artificial intelligence, which have further improved the ability to accurately and cost-effectively analyze market behaviors and manage financial risks [16].

In areas of the finance industry such as algorithmic trading, fraud detection, and asset pricing, both the speed and accuracy metrics of models are critical. Financial models require rapid processing of large volumes of data and often run continuously, demanding substantial computational resources. Despite the desire for high accuracy, we hypothesize that researchers in the finance industry value computational complexity more than the healthcare or architecture industries, since milliseconds can equate to significant financial differences in many model applications.

Within the finance industry, 35 total papers with relevant data for this work were found. Table 2.4 shows a breakdown of the subdomains of which the papers belong.

Finance Subdomains	**Paper Count**
Markets	25
Insurance	6
Econometrics	4
Total	**35**

Table 2.4: Breakdown of the number papers in the dataset from the finance industry by subdomain.

The most frequent subdomain is markets, which encompasses studies related to predicting various characteristics of financial markets. Of the papers in this subdomain, 25 (64%) relate to derivatives pricing and volatility modeling, an area financial engineering that has historically relied heavily on computationally expensive models to solve stochastic partial differential equations and costly simulations to mimic the results of such models. The potential for faster derivatives pricing via relatively accurate machine learning models has led to an influx of buzz around potential applications within this space.

2.3 Other Domains

In order to leverage more available data and create a benchmark to which we can compare industry-specific results, an additional 45 papers were gathered that do not fall under the healthcare, architecture, or finance domains.

The choice to include papers from other domains is essential for establishing a comprehensive baseline for comparison, allowing us to understand the performance and applicability of advanced computing models across a broader spectrum of industries. By incorporating data from diverse fields, we can better assess the generalizability of our findings and better understand industry-specific nuances. This holistic approach ensures that our analysis is robust and reflective of the wide-ranging impact of advanced computing technologies.

Table 2.5 provides a summary of the additional 45 papers, breaking down the count of papers in our database by domain and specifying the subdomains.

Domain	**Subdomains Represented**	**Paper Count**
Sustainability	Climate Models, Weather Forecasting, Water Supply	15
Natural Sciences	Particle Physics, Chemistry, Cosmology	16
Computer Science	Large-scale Simulations, Computer Vision, NLP	11
Miscellaneous	Agriculture, Safety, Material Design	3
Total		**45**

Table 2.5: Breakdown of the number of papers from other domains represented in the dataset.

Chapter 3

Methods

3.1 Data Classification

Throughout the data collection and paper review phases of this project, we noticed that the data we needed was reported in various forms across different studies. Some studies made direct comparisons, for example comparing a singular traditional machine learning model to a singular deep learning model, while others used more complex methods involving full simulations versus surrogate models or several models compared within each of the scientific computing, machine learning, and deep learning fields. The goal of our classification strategy was to extract useful metrics from as many papers as possible, without sacrificing the accuracy of our analysis. In order to have data that is comparable across papers, we needed to design a framework for standardizing and comparing data.

Overall, we focused on determining how each study as a whole aligned or diverged in the context of computational cost versus model performance trade-offs. This approach enabled us to systematically categorize the papers, streamline our analysis, and generate insights into how different models are evaluated and chosen based on industry-specific criteria. Below we review how metrics were standardized to be compared across studies, how model comparison metrics were calculated, and our methodology for extracting relevant data for papers that made a singular model comparison. For further details on the special cases of model comparisons that we encountered and how such papers were classified, refer to Appendix A.

3.1.1 Binary Comparisons

For each of the 150 papers in our dataset, we were able to identify which model performed better and which model used more computational cost, even if specific cost and performance metrics were not provided. Using a binary comparison metric allowed us to mark more papers as relevant in our analysis than if we had only focused on making relative comparisons. Many of the papers that we reviewed only reported one of the metrics that we were looking for as a numerical comparison, but still touched on the models relative performance for the other metric. Binary comparisons allow us to include such models in our analysis. Table 3.1 describes how papers were labeled, with each paper getting one label to describe model performance and one label to describe cost.[1]

In Comparison to the Base Model, the New Model	**Label**
Has Better Performance	+1
Has Equal Performance	0
Has Worse Performance	-1
Uses More Computational Power	+1
Uses Equal Computational Power	0
Uses Less Computational Power	-1

Table 3.1: Description of Binary Classification Labels.

3.1.2 Relative Comparisons

After determining the model that had better performance the model that used less computational power in each paper, we sought out quantitative metrics to analyze the magnitude of model improvements. Of the papers in our dataset, 71 provided quantitative comparison metrics for both model performance and computational cost. We refer to these papers as having relative comparisons. Table 3.2 summarizes the number of papers within each domain that we were able to extract the metrics necessary for a relative comparison.

After determining which model performed better and which model used less computational power in each paper, we then sought out quantitative metrics to analyze the magnitude of model improvements. Across the papers that we reviewed, several quantitative metrics were used to measure performance and computational cost.

For each of the papers with relative comparisons in our dataset, we were able to classify the performance and cost metrics and label the papers as belonging to one of two sets: larger numbers indicating better performance (P_L) or smaller numbers indicating better performance (P_S). Similarly, all cost metrics that were reported could be partitioned into two

[1]It is important to note that favorable models will have a +1 label for performance and a -1 label for compute.

Industry	Binary Comparisons	Relative Comparisons
Healthcare	35	15
Architecture	35	13
Finance	35	17
Other	45	26
Total	**150**	**71**

Table 3.2: Count of papers broken down by level of data extracted (binary vs relative).

categories: larger numbers indicating a higher computational cost (C_L) or smaller numbers higher computational cost (C_S). In Table 3.3, we present the most commonly reported metrics that belong in each set.

Paper Classification	Metrics
P_L	Accuracy, AUC, F1-score, R^2
P_S	MSE, RMSE, MAE, MAPE
C_L	Training Time, Execution Time, Simulation Runs
C_S	Computation Speedups

Table 3.3: Commonly reported model performance and computational cost metrics in each paper classification set.

Using the performance and cost, we were able to calculate relative performance and relative cost metrics. When comparing cost metrics, specifically when comparing traditional machine learning or deep learning to scientific computing, it is not uncommon to see improvements of several orders of magnitude. To allow us to make a more meaningful comparison between the trade-offs between performance and cost, we chose to calculate our metrics by taking the logarithm of the ratio of the models improvement. This calculation varied slightly depending on which relative metric was being computed and the paper classification of the source. Equations 3.1, 3.2, 3.3, and 3.4 explain how calculations were made for each paper classification.

$$Paper_i \in P_L : RelPref_i = log_{10}(\frac{PerformanceMetric_{New_i}}{PerformanceMetric_{Base_i}}) \tag{3.1}$$

$$Paper_i \in P_S : RelPref_i = log_{10}(\frac{PerformanceMetric_{Base_i}}{PerformanceMetric_{New_i}}) \tag{3.2}$$

$$Paper_i \in C_L : RelCost_i = log_{10}(\frac{CostMetric_{Base_i}}{CostMetric_{New_i}}) \tag{3.3}$$

$$Paper_i \in C_S : RelCost_i = log_{10}(\frac{CostMetric_{New_i}}{CostMetric_{Base_i}}) \tag{3.4}$$

Chapter 4

Results

4.1 Performance vs Compute Frontiers

Using both the binary and relative comparison data, we perform several analyses to examine the relationship between model performance and computational cost across our three selected industries.

4.1.1 Binary Data

To compare binary data points, we calculated the percentage of total papers within each domain and Base → New model comparison category that reported a performance improvement. We performed the same calculation for computational cost improvements. Tables 4.1 and 4.2 summarize our findings.

Industry	ML → DL	SC → DL	SC → ML	All Base → New
Healthcare	46.43%	20.00%	0.00%	40.00%
Architecture	40.00%	0.00%	5.88%	8.57%
Finance	80.00%	5.56%	28.57%	31.43%
Other	87.50%	14.71%	0.00%	26.67%
All Domains	**58.52%**	**10.00%**	**10.34%**	**26.67%**

Table 4.1: Percent of papers that showed performance improvements after switching from the Base model to the New model.

We see that papers in our dataset that compared traditional machine learning to scientific computing reported an improvement in computational cost. However, when looking at other

Industry	ML → DL	SC → DL	SC → ML	All Base → New
Healthcare	25.00%	60.00%	100.00%	34.29%
Architecture	60.00%	100.00%	100.00%	94.29%
Finance	40.00%	94.44%	100.00%	80.00%
Other	37.50%	100.00%	100.00%	88.89%
All Domains	**33.33%**	**95.71%**	**100.00%**	**75.33%**

Table 4.2: Percent of papers that showed computational cost improvements after switching from the Base to the New model.

base model and new model pairs, we see distinct trends across industries. For example, the healthcare industry had significantly fewer new models with improved computational cost metrics across both ML → DL comparisons and SC → DL comparisons. This suggests that the healthcare industry may prioritize reducing computational cost less than the other industries that we studied. For example, the finance industry showed significantly better rates of computational cost improvement than the healthcare industry. In terms of performance improvements, the healthcare industry has the highest improvement rate across all Base → New models at 40%. The finance industry also showed a significant percentage of papers reporting performance improvements (31.43%), though not as high as the healthcare industry. The architecture industry, heavily reliant on machine learning-based surrogate models, saw low performance improvements compared to the rest of the industries.

These results corroborate our hypothesis that researchers in the healthcare industry tend to favor computationally complex models with extremely high accuracy over models that are significantly less expensive but slightly less accurate while the opposite is true in the finance industry.

Figure 4.1 shows the percentage of papers that demonstrated performance improvements after switching from the Base model to the New model while Figure 4.2 illustrates the percentage of papers that observed computational cost improvements when switching models. These graphs helps visualize how often new alternative computational models outperform the baseline across different comparisons (ML → DL, SC → DL, ML → SC) across different industries.

Figure 4.3 captures both the performance and cost dimensions at the same time by plotting the percent improvements for both performance and cost metrics on the same graph. This helps us visualize where the performance-compute trade-off point may lie for each industry.

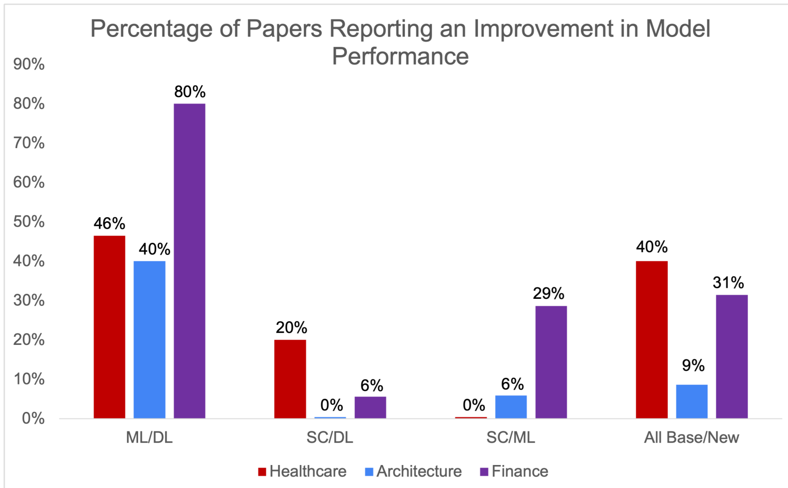


Figure 4.1: Percent of papers that showed performance improvements after switching from the Base model to the New model.

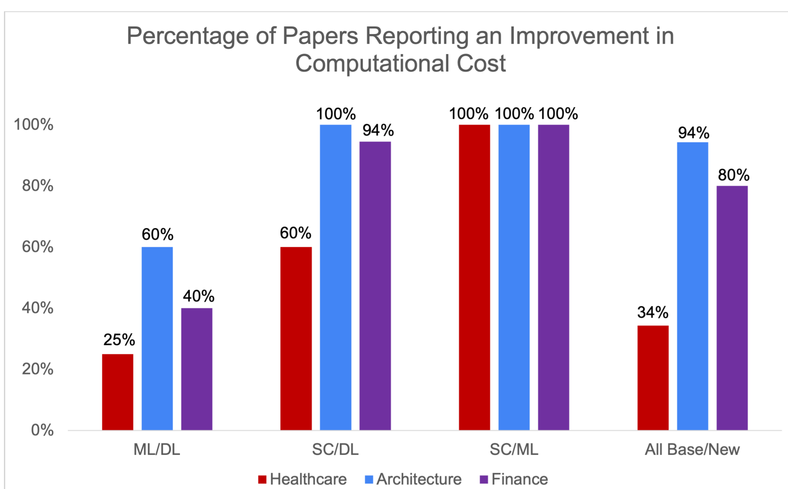


Figure 4.2: Percent of papers that showed computational cost improvements after switching from the Base model to the New model.

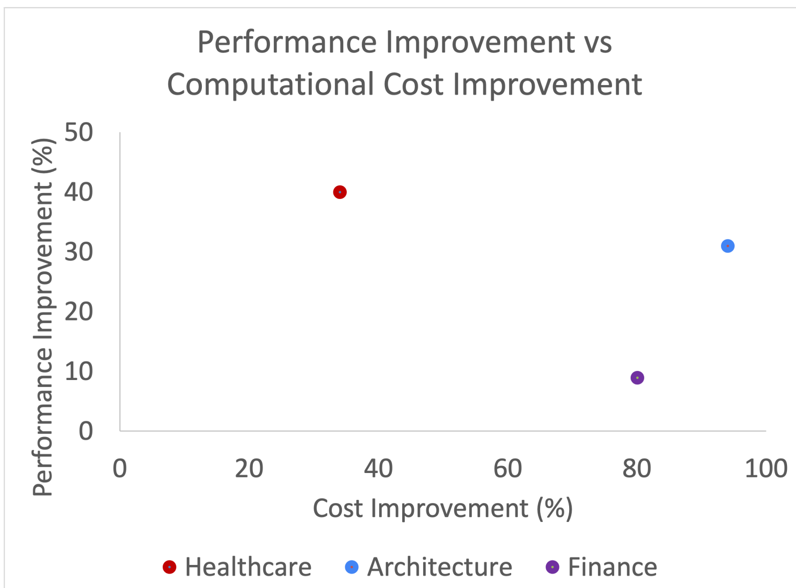


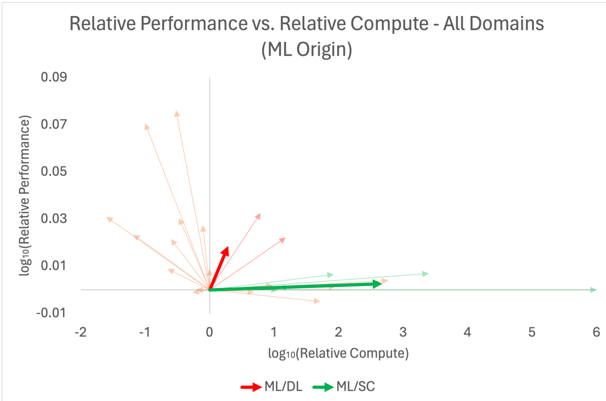
Figure 4.3: Performance and Computational Cost Percent Improvements by Industry.

4.1.2 Relative Data

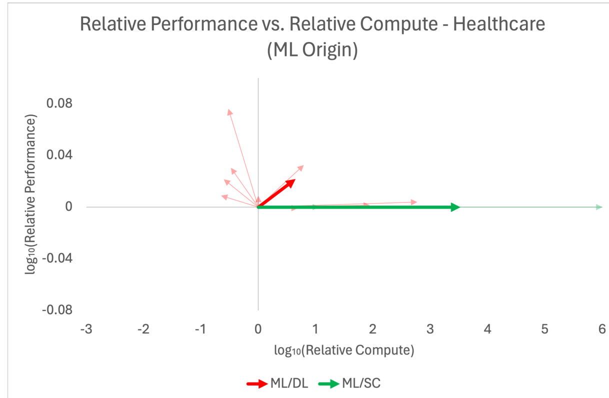
Using the idea illustrated in Figure 1.1, we generate model performance vs computational cost frontiers for each of our industries of interest. Each graph plots the relative performance and relative compute metrics for the specified Base $\rightarrow$ New categories. Metrics were calculated using Equations 3.1, 3.2, 3.3, and 3.4.

When plotting the data-points, we use arrows from the Base model metric to the New model metric to represent the magnitude and direction of performance and cost improvements. For example, a paper i with a DL $\rightarrow$ SC relative comparison in which the scientific computing model was preferred in terms of performance ($log_{10}(RelPref_i) \geq 0$) and had a higher computational cost ($log_{10}(RelComp_i) \geq 0$) would be represented by an arrow pointing to the first quadrant of the graph.

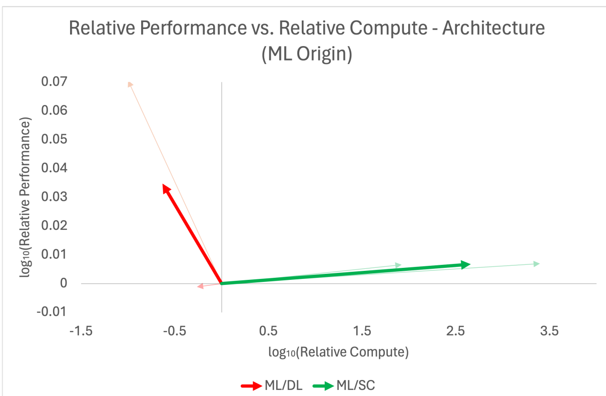
Plots were generated for all three of the potential base models, with the base model representing coordinate (0,0) on their respective graphs. For example, the charts labeled "ML Origin" have data points of the form ML $\rightarrow$ DL and ML $\rightarrow$ SC plotted. The relative accuracy/computational cost of either a scientific computing or a deep learning model is compared to that of machine learning, and each arrow pointing away from the origin represents both the magnitude and the direction of the improvement metrics. The performance of models from individual papers are represented by the thin, semitransparent arrows. The thicker, opaque arrows represent the arithmetic mean of all Base $\rightarrow$ New performance and cost values within the specified domain(s). In the plots, arrows ($\rightarrow$) are replaced with slashes (/).



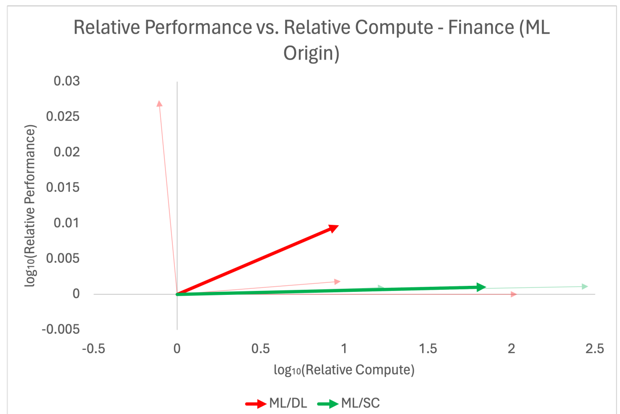
(a) All domains (n=25)



(b) Healthcare (n=11)



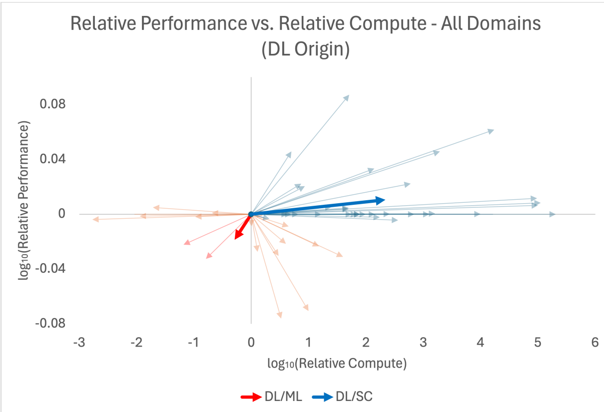
(c) Architecture (n=4)



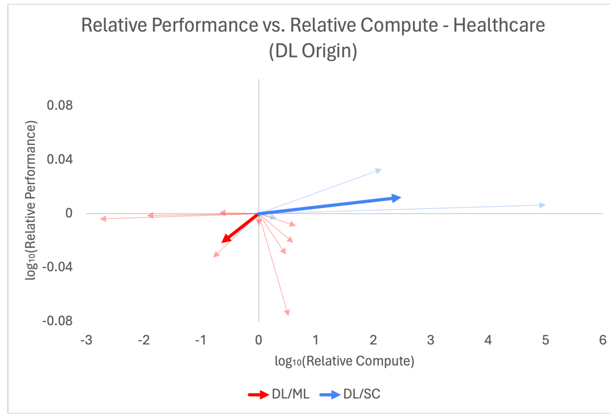
(d) Finance (n=5)

Figure 4.4: Relative performance vs relative compute plot for different domains with machine learning as the base model.

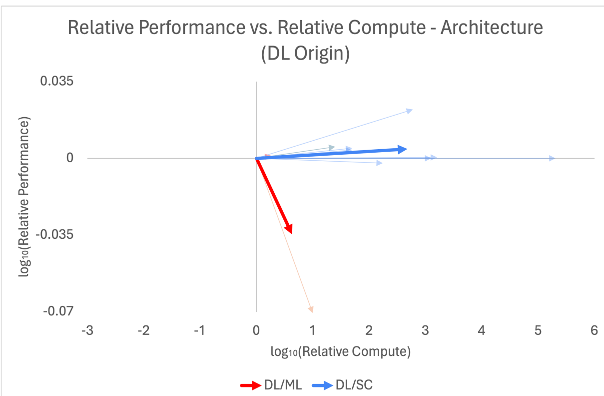
Figure 4.4 shows relationships that support our hypothesis that compared to traditional machine learning, both deep learning and scientific computing have better performance and use more computational cost. The one exception to this is shown in the architecture industry. Figure 4.4c shows that the papers that compared deep learning to machine learning in the within this domain have an unexpected relationship, with deep learning models being, on average, less computationally expensive than traditional machine learning. After taking a closer look at the data, we see that there are only two papers with relative comparisons between ML and DL in the architecture domain. A small sample size may explain why we see this unexpected relationship.



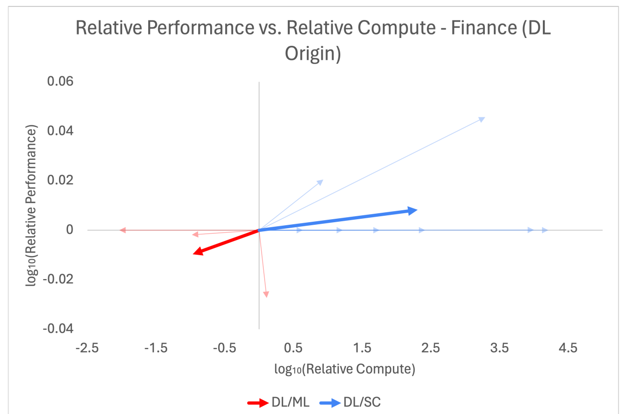
(a) All domains (n=54)



(b) Healthcare (n=12)



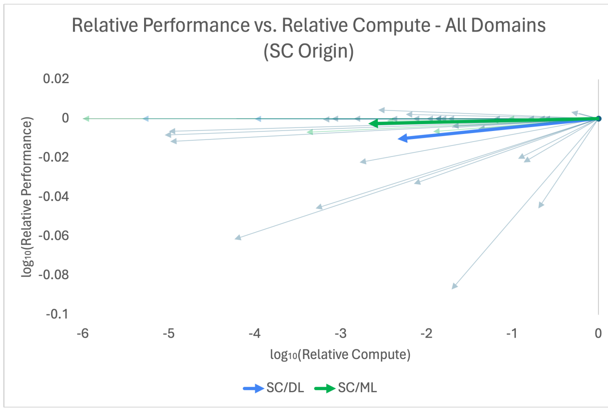
(c) Architecture (n=10)



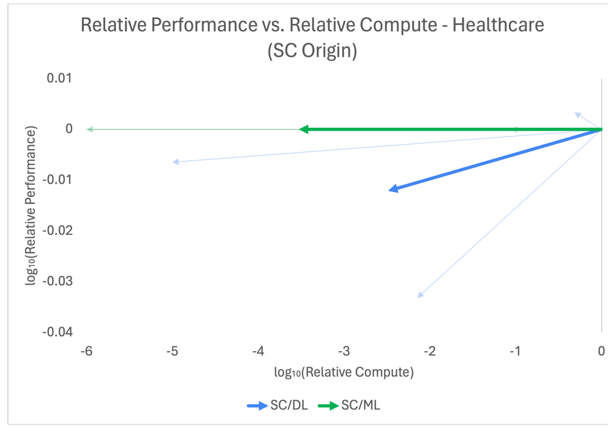
(d) Finance (n=11)

Figure 4.5: Relative performance vs relative compute plot for different domains with deep learning as the base model.

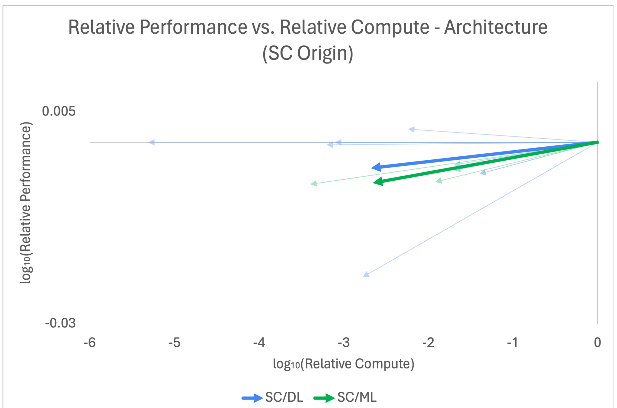
The relationships shown in Figure 4.5 support our hypothesis that compared to deep learning, machine learning has worse performance but is less computationally demanding and scientific computing has better performance but has higher computational cost. Again, the only counterexample to this claim is comparing machine learning to deep learning in the architecture industry. Since ML → DL and DL → ML comparisons are equal in magnitude and opposite in direction, it makes sense that we see the same relationship as in Figure 4.4.



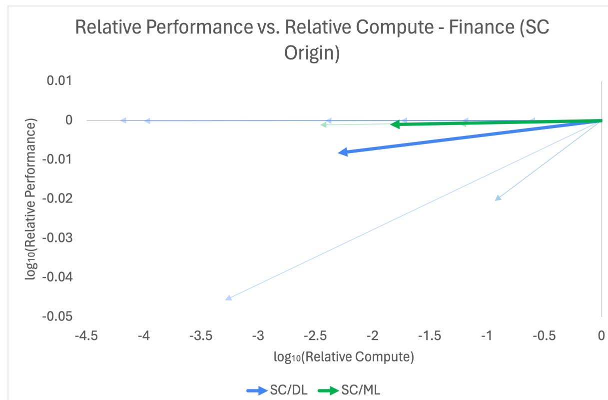
(a) All domains (n=41)



(b) Healthcare (n=5)



(c) Architecture (n=10)



(d) Finance (n=10)

Figure 4.6: Relative performance vs relative compute plot for different domains with scientific computing as the base model.

Figure 4.6 supports our hypothesis that traditional machine learning models and deep learning models both significantly reduce computational cost when compared to scientific computing. All eight of the thicker, opaque arrows shown in the figure lie in the third quadrant of the performance vs compute graphs. This suggests that the trend is both observable and consistent across all industries.

4.2 Progression of Model Improvements Over Time

To better understand model improvements and the rate at which advanced computing techniques are adopted across industries, we perform analysis our data segmented by the year of publication of each other papers in our dataset. When looking for a way to effectively split our dataset into two time buckets, we prioritized choosing a year that would partition the data into two buckets that were relatively equal in size. We also wanted the year we chose to mark a significant milestone in the landscape of advanced computing. Keeping this in mind, we chose to split our data into two categories: papers published in 2018 and before and papers published in 2019 and after. Table 4.3 shows that this offers not only an approximately even split between the number of papers across all domains, but within each industry there are enough papers within each time bucket for meaning for analysis.

Industry	**2000-2018**	**2019-2024**
Healthcare	19	16
Architecture	26	9
Finance	10	25
Other	45	26
Total	**70**	**80**

Table 4.3: Data split between two time buckets.

4.2.1 Exponentially Increasing Publications

As new computing techniques, specifically traditional machine learning and deep learning models, perform better and become more widespread, more and more publications that are relevant to this study are released each year. In an analysis of the landscape of machine learning in architectural design, Papasotiriou finds that there has been an exponential increase in the number of publications in the field since the early 2000s [14]. We reach the same conclusion when analyzing the papers in our dataset. Figure 4.7 presents the distribution of papers in the dataset published during each three-year period, providing insights into trends over time regarding the publication and possibly the evolution of computational models in research.

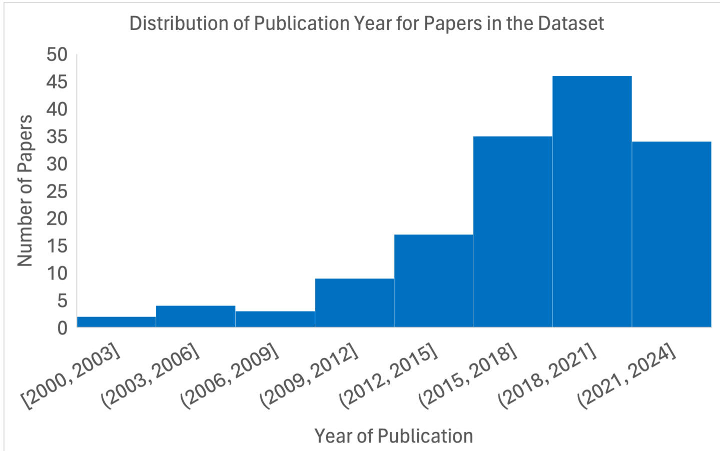


Figure 4.7: Number of papers in the dataset published during each three year period.

4.2.2 Model Improvements

Figures 4.8 and 4.9 show the percentage of papers that reported improvements in performance and computational cost, respectively, from the Base to the New model across different model comparisons over time. These graphs help in understanding how advancements computing have contributed to better performance and efficiency across years and which industries are benefiting the most.

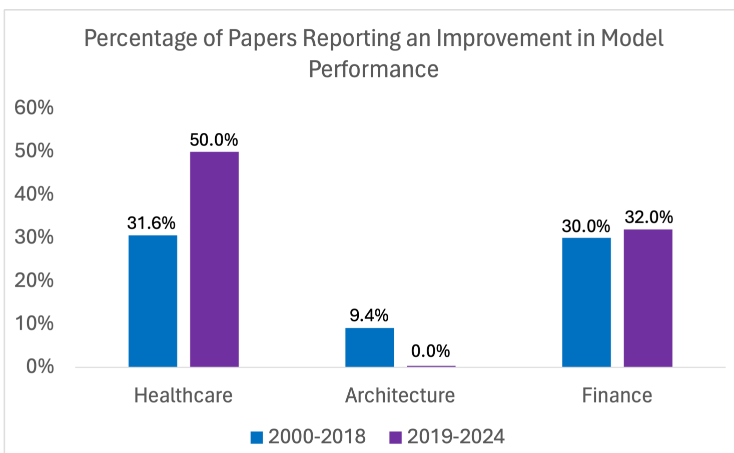


Figure 4.8: Percent of papers that showed performance improvements from the Base model to the New model across all Base → New combinations.

We see that the healthcare industry has shown significant improvements since 2019 in both model performance and computational cost while the architecture and finance industries are lagging behind in both metrics. This suggests that researchers in the healthcare industry may be prioritizing both performance and cost metrics in recent works and the architecture

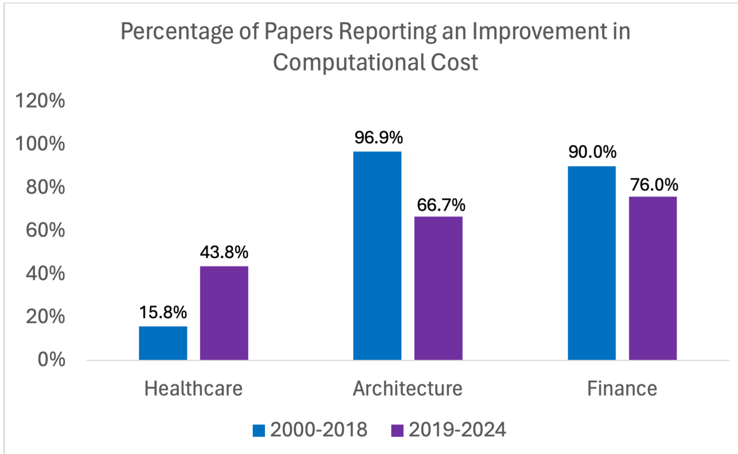


Figure 4.9: Percent of papers that showed computational cost improvements from the Base model to the New model across all Base $\rightarrow$ New combinations.

and finance industries may have a clearer focus on one metric over the other.

Chapter 5

Conclusion

The findings of the study reveal distinct trends in how different industries prioritize model performance versus computational cost. We hypothesize what the performance vs computational cost trade-off curve looks like when comparing traditional machine learning, deep learning, and scientific computing and find results that are consistent with our postulate.

We find that the healthcare industry favors computationally complex models with extremely high accuracy over models that demand significantly less computational power but are slightly less accurate. 40% of the papers reviewed in the finance industry showed performance improvements when switching from the Base model to the New model but only 34.29% of the papers showing improvements in computational cost. In the architecture industry, we saw significant strides in reducing computational cost, with 94.29% of the papers reviewed showing improvements, while performance metrics lagged behind with only 8.57% of the papers reviewed showing improvements. We note that, compared to the other industries reviewed, architecture relies significantly more on scientific computing, which can explain the focus on reducing cost at the expense of performance. In the finance industry, we find that speed is often prioritized over marginal gains in model performance, with only 31.43% of the papers reviewed showed performance improvements while an impressive 80% of the papers showed improvements in computational cost.

Finally, we saw that the number of publications relevant to this study has increased exponentially with time, suggesting that the landscape of advanced computing is constantly evolving, as is the performance vs computational cost trade-off curve.

5.1 Future Work

The methodology that we present in this project gives rise to several directions for future work. One promising direction is applying the framework of our analysis to additional industries beyond those analyzed in this study. While this research focused on healthcare, architecture, and finance, other sectors such as manufacturing, transportation, and energy could benefit from a similar analysis. The end goal for this work is to have a comprehensive description of the entire landscape of machine learning, which will require the analysis of papers across all domains.

Another important direction for future work is to perform a similar analysis on additional models and computation techniques. For example, quantum computing and other emerging technologies have the potential to significantly change the trade-offs between performance and computational cost that we have found in this study. Additionally, performing longitudinal studies to examine the evolution of these trade-offs over time, can provide valuable insights as to the extent of the impact that each advancement in computing has on individual domains and the landscape as a whole.

Appendix A

Special Cases of Model Comparisons

A.1 Machine Learning-Based Surrogate Models

A machine learning surrogate model is a type of model used across several domains to approximate complex, computationally expensive simulations or functions. The goal of a surrogate model is to provide an efficient alternative that closely mimics the behavior of the original model but at a significantly reduced computational cost [15]. Figure 3.1 provides a visual representation of a surrogate model embedded into a simulation.

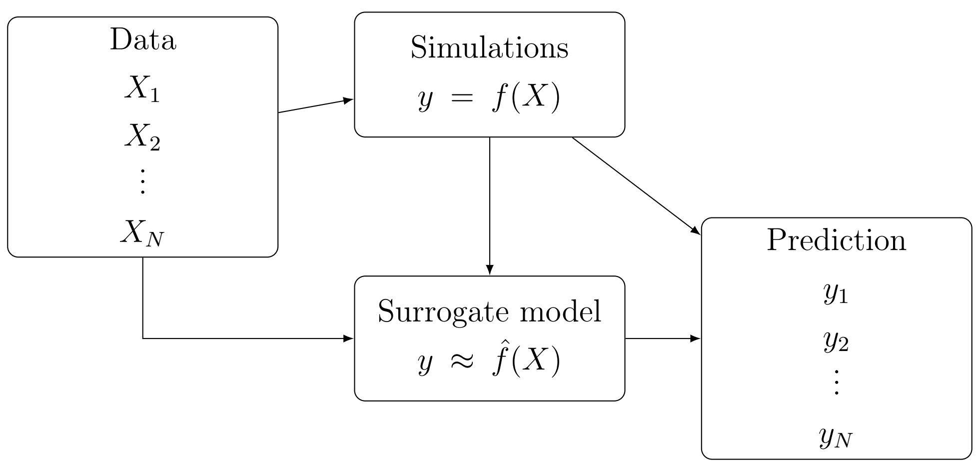


Figure A.1: Pipeline for enhancing simulations with machine learning and deep learning-based surrogate models. Computationally expensive steps of the full simulation (f) are replaced with approximate predictions generated by the surrogate ($\hat{f}$).

Surrogate models are often used when the relationship between input and output data is unknown or not observable in the real world. As a result, it is common practice for surrogate models to be trained on and tested against data that is generated from the original full simulation model. This results in performance metrics for the new model that are reported relative to the baseline model, not absolute comparisons to actual labels.

Having knowledge of surrogate models significantly aided our search for relevant papers for this study. Because the primary goal of a surrogate model is to find a better balance between computational cost and model performance, many of the papers reviewed that implemented surrogate models included all of the data points that we were looking to extract. When analyzing surrogate model papers, we chose to think of them as papers that show "DL enhancing SC" rather than "DL outperforming SC". However, for the purposes of our dataset, we classified papers that compared the performance of a surrogate model to a full simulation as SC → "Surrogate model type". For example, a paper that investigates using a SVM-based surrogate model for down-scaling the computational resources needed to run part of a climate simulation would be classified as SC → ML.

A.2 Papers with Multiple Comparisons

Oftentimes, papers that are specifically focused on improving one of the two metrics will hold the less relevant metric constant and report the size of the improvement for the other metric. For example, a paper may set an accuracy threshold that it requires each model to meet and reports the amount of computing power necessary for each model to reach the threshold. This is particularly common in SC → DL and SC → ML papers that compare a machine learning enhanced surrogate model to a full simulation. For papers with metrics reported in this fashion, the binary value of the metric that was held constant was recorded as 0 and the main metric for improvement comparison was recorded as normal.

Many of the papers we reviewed compared more than two different models. For example rather than comparing a support vector machine (SVM) machine learning model to a Long Short Term Memory (LSTM) deep learning model, a paper may compare SVM, logistic regression, LSTM, and CNN models, resulting in several different ML → DL comparisons for one paper. To weigh each paper evenly in our analysis regardless of the number of models compared, we chose to pick one model from each of the base and new techniques. We chose to use the model with the highest accuracy within each, breaking ties by choosing the model that is less computationally expensive as necessary.

References

[1] E. Strubell, A. Ganesh, and A. McCallum, *Energy and Policy Considerations for Deep Learning in NLP*, arXiv:1906.02243 [cs], Jun. 2019.

[2] N. C. Thompson, S. Ge, and G. F. Manso, *The Importance of (Exponentially More) Computing Power*, arXiv:2206.14007 [cs], Jun. 2022.

[3] I. Kassandra Pereira Belfort, I. Danielli Borges De Sousa, T. Silva Rodrigues, A. Paula Cunha, V. Edilene Duarte Martins, S. Cristina Moutinho Monteiro, and A. Kardec Duailibe Barros, "COMPUTER MODELS FOR DISEASE PREDICTION," en, *International Journal for Innovation Education and Research*, vol. 8, no. 1, pp. 275–284, Jan. 2020, ISSN: 2411-2933, 2411-3123.

[4] J. Broughel and A. Thierer, "Technological Innovation and Economic Growth: A Brief Report on the Evidence," en,

[5] R. M. Solow, "A Contribution to the Theory of Economic Growth," en, *The Quarterly Journal of Economics*, vol. 70, no. 1, p. 65, Feb. 1956, ISSN: 00335533.

[6] V. Pai and S. Chandra, "Exploring Factors Influencing Organizational Adoption of Artificial Intelligence (AI) in Corporate Social Responsibility (CSR) Initiatives," *Pacific Asia Journal of the Association for Information Systems*, vol. 14, no. 5, Sep. 2022, ISSN: 1943-7544.

[7] C. Bauer and A. Thamm, "Six Areas of Healthcare Where AI Is Effectively Saving Lives Today," en, in *Digitalization in Healthcare: Implementing Innovation and Artificial Intelligence*, ser. Future of Business and Finance, P. Glauner, P. Plugmann, and G. Lerzynski, Eds., Cham: Springer International Publishing, 2021, pp. 245–267, ISBN: 978-3-030-65896-0.

[8] H. Varian, "Artificial Intelligence, Economics, and Industrial Organization," en,

[9] T. Babina, A. Fedyk, A. He, and J. Hodson, "Artificial intelligence, firm growth, and product innovation," en, *Journal of Financial Economics*, vol. 151, p. 103 745, Jan. 2024, ISSN: 0304405X.

[10] "Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning," en, *Iraqi Journal for Computer Science and Mathematics*, pp. 115–123, Jan. 2022, ISSN: 2788-7421.

[11] O. Ameri Sianaki, A. Yousefi, A. Tabesh, and M. Mahdavi, "Machine Learning Applications: The Past and Current Research Trend in Diverse Industries," en, *Inventions*, vol. 4, no. 1, p. 8, Feb. 2019, ISSN: 2411-5134.

[12] N. Zolas, Z. Kroff, E. Brynjolfsson, K. McElheran, D. Beede, C. Buffington, N. Goldschlag, L. Foster, and E. Dinlersoz, "Advanced Technologies Adoption and Use by U.S. Firms: Evidence from the Annual Business Survey," en, w28290, Dec. 2020, Place: Cambridge, MA Publisher: National Bureau of Economic Research Report Number: w28290.

[13] S. Dutta, "An overview on the evolution and adoption of deep learning applications used in the industry," en, *WIREs Data Mining and Knowledge Discovery*, vol. 8, no. 4, e1257, Jul. 2018, ISSN: 1942-4787, 1942-4795.

[14] T. Papasotiriou, "Identifying the Landscape of Machine Learning-Aided Architectural Design - A Term Clustering and Scientometrics Study," pp. 815–824, 2019, Conference Name: CAADRIA 2019: Intelligent & Informed Place: Wellington, New Zealand.

[15] S. E. Davis, S. Cremaschi, and M. R. Eden, "Efficient Surrogate Model Development: Optimum Model Form Based on Input Function Characteristics," en, in *Computer Aided Chemical Engineering*, vol. 40, Elsevier, 2017, pp. 457–462, ISBN: 978-0-444-63965-3.

[16] S. A. Zenios, "High-performance computing in finance: The last 10 years and the next," en, *Parallel Computing*, vol. 25, no. 13-14, pp. 2149–2175, Dec. 1999, ISSN: 01678191.

[17] *AI and compute*, en-US.

[18] C. O. de Burgh-Day and T. Leeuwenburg, "Machine learning for numerical weather and climate modelling: A review," English, *Geoscientific Model Development*, vol. 16, no. 22, pp. 6433–6477, Nov. 2023, Publisher: Copernicus GmbH, ISSN: 1991-959X.

[19] *Pulling Clouds into Focus – Oak Ridge Leadership Computing Facility.*

[20] I. V. D. Srihith, L. Rajitha, K. Thriveni, A. D. Donald, and P. Blessy, "Trading on Autopilot: The Rise of Algorithmic Trading," en, *Communication and Technology*, vol. 3, no. 4, 2023.

[21] A. S. Huang, K. Hirabayashi, L. Barna, D. Parikh, and L. R. Pasquale, "Assessment of a Large Language Model's Responses to Questions and Cases About Glaucoma and Retina Management," *JAMA Ophthalmology*, e236917, Feb. 2024, ISSN: 2168-6165.

[22] N. Bredella, "Simulation and Architecture: Mapping Building Information Modeling," en, *NTM Zeitschrift für Geschichte der Wissenschaften, Technik und Medizin*, vol. 27, no. 4, pp. 419–441, Dec. 2019, ISSN: 1420-9144.

[23] A. Lg and E. At, "Using Three Machine Learning Techniques for Predicting Breast Cancer Recurrence," en, *Journal of Health & Medical Informatics*, vol. 04, no. 02, 2013, ISSN: 21577420.

[24] D. Delen, G. Walker, and A. Kadam, "Predicting breast cancer survivability: A comparison of three data mining methods," *Artificial Intelligence in Medicine*, vol. 34, no. 2, pp. 113–127, Jun. 2005, ISSN: 0933-3657.

[25] S. Malik, R. Khadgawat, S. Anand, and S. Gupta, "Non-invasive detection of fasting blood glucose level via electrochemical measurement of saliva," *SpringerPlus*, vol. 5, no. 1, p. 701, May 2016, ISSN: 2193-1801.

[26] S. Aneja and S. Lal, "Effective asthma disease prediction using naive Bayes - Neural network fusion technique," en, in *2014 International Conference on Parallel, Distributed and Grid Computing*, Solan, India: IEEE, Dec. 2014, pp. 137–140, ISBN: 978-1-4799-7683-6 978-1-4799-7682-9 978-1-4799-7681-2.

[27] L. Tapak, H. Mahjub, O. Hamidi, and J. Poorolajal, "Real-Data Comparison of Data Mining Methods in Prediction of Diabetes in Iran," en, *Healthcare Informatics Research*, vol. 19, no. 3, p. 177, 2013, ISSN: 2093-3681, 2093-369X.

[28] N. C. Long, P. Meesad, and H. Unger, "A highly accurate firefly based algorithm for heart disease prediction," *Expert Systems with Applications*, vol. 42, no. 21, pp. 8221–8231, Nov. 2015, ISSN: 0957-4174.

[29] S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," en, in *2008 IEEE/ACS International Conference on Computer Systems and Applications*, Doha, Qatar: IEEE, Mar. 2008, pp. 108–115, ISBN: 978-1-4244-1967-8 978-1-4244-1968-5.

[30] Z.-H. Tang, J. Liu, F. Zeng, Z. Li, X. Yu, and L. Zhou, "Comparison of Prediction Model for Cardiovascular Autonomic Dysfunction Using Artificial Neural Network and Logistic Regression Analysis," en, *PLoS ONE*, vol. 8, no. 8, R. K. Hills, Ed., e70571, Aug. 2013, ISSN: 1932-6203.

[31] J. Kim, J. Lee, and Y. Lee, “Data-Mining-Based Coronary Heart Disease Risk Prediction Model Using Fuzzy Logic and Decision Tree,” en, *Healthcare Informatics Research*, vol. 21, no. 3, p. 167, 2015, ISSN: 2093-3681, 2093-369X.

[32] N. Bhatla and K. Jyoti, “An Analysis of Heart Disease Prediction using Different Data Mining Techniques,” en, *International Journal of Engineering Research*, vol. 1, no. 8, 2012.

[33] K. Thenmozhi and P. Deepika, “Heart Disease Prediction Using Classification with Different Decision Tree Techniques,” en, vol. 2, no. 6, 2014.

[34] C.-Y. Hung, W.-C. Chen, P.-T. Lai, C.-H. Lin, and C.-C. Lee, “Comparing deep neural network and other machine learning algorithms for stroke prediction in a large-scale population-based electronic medical claims database,” en, in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Seogwipo: IEEE, Jul. 2017, pp. 3110–3113, ISBN: 978-1-5090-2809-2.

[35] Ö. Eskidere, F. Ertaş, and C. Hanilçi, “A comparison of regression methods for remote tracking of Parkinson’s disease progression,” en, *Expert Systems with Applications*, vol. 39, no. 5, pp. 5523–5528, Apr. 2012, ISSN: 09574174.

[36] M. Islam, C.-C. Wu, T. N. Poly, and C. Yang, “Applications of Machine Learning in Fatty Live Disease Prediction,” en,

[37] P. Lu, S. Guo, H. Zhang, Q. Li, Y. Wang, Y. Wang, and L. Qi, “Research on Improved Depth Belief Network-Based Prediction of Cardiovascular Diseases,” en, *Journal of Healthcare Engineering*, vol. 2018, pp. 1–9, 2018, ISSN: 2040-2295, 2040-2309.

[38] R. Ani, G. Sasi, U. R. Sankar, and O. S. Deepa, “Decision support system for diagnosis and prediction of chronic renal failure using random subspace classification,” en, in *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Jaipur, India: IEEE, Sep. 2016, pp. 1287–1292, ISBN: 978-1-5090-2029-4.

[39] B. Moradi, M. Aghapour, and A. Shirbandi, “Compare of Machine Learning and Deep Learning Approaches for Human Activity Recognition,” in *2022 30th International Conference on Electrical Engineering (ICEE)*, ISSN: 2642-9527, May 2022, pp. 592–596.

[40] H. Daoud and M. A. Bayoumi, “Efficient Epileptic Seizure Prediction Based on Deep Learning,” en, *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 5, pp. 804–813, Oct. 2019, ISSN: 1932-4545, 1940-9990.

[41] T. Ayer, J. Chhatwal, O. Alagoz, C. E. Kahn, R. W. Woods, and E. S. Burnside, "Comparison of Logistic Regression and Artificial Neural Network Models in Breast Cancer Risk Estimation," *RadioGraphics*, vol. 30, no. 1, pp. 13–22, Jan. 2010, Publisher: Radiological Society of North America, ISSN: 0271-5333.

[42] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease Prediction by Machine Learning Over Big Data From Healthcare Communities," en, *IEEE Access*, vol. 5, pp. 8869–8879, 2017, ISSN: 2169-3536.

[43] W. Wang, J. Lee, F. Harrou, and Y. Sun, "Early Detection of Parkinson's Disease Using Deep Learning and Machine Learning," en, *IEEE Access*, vol. 8, pp. 147 635–147 646, 2020, ISSN: 2169-3536.

[44] A. Abrol, Z. Fu, M. Salman, R. Silva, Y. Du, S. Plis, and V. Calhoun, "Deep learning encodes robust discriminative neuroimaging representations to outperform standard machine learning," en, *Nature Communications*, vol. 12, no. 1, p. 353, Jan. 2021, ISSN: 2041-1723.

[45] P. Du, X. Zhu, and J.-X. Wang, "Deep learning-based surrogate model for three-dimensional patient-specific computational fluid dynamics," *Physics of Fluids*, vol. 34, no. 8, p. 081 906, Aug. 2022, ISSN: 1070-6631.

[46] M. Pop, M. Sermesant, O. Camara, X. Zhuang, S. Li, A. Young, T. Mansi, and A. Suinesiaputra, Eds., *Statistical Atlases and Computational Models of the Heart. Multi-Sequence CMR Segmentation, CRT-EPiggy and LV Full Quantification Challenges: 10th International Workshop, STACOM 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Revised Selected Papers* (Lecture Notes in Computer Science), en. Cham: Springer International Publishing, 2020, vol. 12009, ISBN: 978-3-030-39073-0 978-3-030-39074-7.

[47] W. A. Pruett and R. L. Hester, "The Creation of Surrogate Models for Fast Estimation of Complex Model Outcomes," en, *PLOS ONE*, speex, vol. 11, no. 6, J. Boltze, Ed., e0156574, Jun. 2016, ISSN: 1932-6203.

[48] B. Tran, B. Xue, and M. Zhang, "Genetic programming for feature construction and selection in classification on high-dimensional data," en, *Memetic Computing*, vol. 8, no. 1, pp. 3–15, Mar. 2016, ISSN: 1865-9284, 1865-9292.

[49] Y. Zhang and S. Wang, "Detection of Alzheimer's disease by displacement field and machine learning," *PeerJ*, vol. 3, e1251, Sep. 2015, ISSN: 2167-8359.

[50] C. Mansoor, S. K. Chettri, and H. Naleer, “Efficient Prediction Model for Cardiovascular Disease Using Deep Learning Techniques,” en, *Migration Letters*, vol. 20, no. S13, pp. 449–459, Dec. 2023, ISSN: 1741-8992, 1741-8984.

[51] S. Purushotham, C. Meng, Z. Che, and Y. Liu, “Benchmarking deep learning models on large healthcare datasets,” *Journal of Biomedical Informatics*, vol. 83, pp. 112–134, Jul. 2018, ISSN: 1532-0464.

[52] K. Davagdorj, J.-W. Bae, V.-H. Pham, N. Theera-Umpon, and K. H. Ryu, “Explainable Artificial Intelligence Based Framework for Non-Communicable Diseases Prediction,” en, *IEEE Access*, vol. 9, pp. 123 672–123 688, 2021, ISSN: 2169-3536.

[53] Y. X. Zhao, H. Yuan, and Y. Wu, “Prediction of Adverse Drug Reaction using Machine Learning and Deep Learning Based on an Imbalanced Electronic Medical Records Dataset,” en, in *2021 5th International Conference on Medical and Health Informatics*, Kyoto Japan: ACM, May 2021, pp. 17–21, ISBN: 978-1-4503-8984-6.

[54] R. C. Myers, F. Augustin, J. Huard, and C. M. Friedrich, “Using machine learning surrogate modeling for faster QSP VP cohort generation,” en, *CPT: Pharmacometrics & Systems Pharmacology*, vol. 12, no. 8, pp. 1047–1059, Aug. 2023, ISSN: 2163-8306, 2163-8306.

[55] T. Dang, A. V. Luong, A. W. C. Liew, J. McCall, and T. T. Nguyen, “Ensemble of deep learning models with surrogate-based optimization for medical image segmentation,” en, in *2022 IEEE Congress on Evolutionary Computation (CEC)*, Padua, Italy: IEEE, Jul. 2022, pp. 1–8, ISBN: 978-1-66546-708-7.

[56] *Surrogate models based on machine learning methods for parameter estimation of left ventricular myocardium*, en.

[57] F. Leong, C. Y. Lai, S. F. Khosroshahi, L. He, S. de Lusignan, T. Nanayakkara, and M. Ghajari, “A Surrogate Model Based on a Finite Element Model of Abdomen for Real-Time Visualisation of Tissue Stress during Physical Examination Training,” eng, *Bioengineering (Basel, Switzerland)*, vol. 9, no. 11, p. 687, Nov. 2022, ISSN: 2306-5354.

[58] D. Buscombe and A. Ritchie, “Landscape Classification with Deep Neural Networks,” en, *Geosciences*, vol. 8, no. 7, p. 244, Jul. 2018, ISSN: 2076-3263.

[59] N. Limei, W. Dongfan, and Z. Bo, “Landscape image recognition and analysis based on deep learning algorithm,” en, *Journal of Intelligent & Fuzzy Systems*, pp. 1–11, Mar. 2024, ISSN: 10641246, 18758967.

[60] L. Zhang, “Application of Lightweight Deep Learning Model in Landscape Architecture Planning and Design,” en, *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 10, 2023, ISSN: 21565570, 2158107X.

[61] S. Yousif and D. Bolojan, “Deep Learning-Based Surrogate Modeling for Performance-Driven Generative Design Systems,” en, Sydney, Australia, 2022, pp. 363–372.

[62] S. Asadi, S. S. Amiri, and M. Mottahedi, “On the development of multi-linear regression analysis to assess energy consumption in the early stages of building design,” en, *Energy and Buildings*, vol. 85, pp. 246–255, Dec. 2014, ISSN: 03787788.

[63] B. Eisenhower, Z. O’Neill, S. Narayanan, V. A. Fonoberov, and I. Mezić, “A methodology for meta-model based optimization in building energy models,” en, *Energy and Buildings*, vol. 47, pp. 292–301, Apr. 2012, ISSN: 03787788.

[64] S. Tarabishy, S. Psarras, M. Kosicki, and M. Tsigkari, “Deep learning surrogate models for spatial and visual connectivity,” en, *International Journal of Architectural Computing*, vol. 18, no. 1, pp. 53–66, Mar. 2020, Publisher: SAGE Publications, ISSN: 1478-0771.

[65] F. Ritter, P. Geyer, and A. Borrmann, “Simulation-based Decision-making in Early Design Stages,” en, 2015.

[66] P. Geyer and A. Schlüter, “Automated metamodel generation for Design Space Exploration and decision-making – A novel method supporting performance-oriented building design and retrofitting,” *Applied Energy*, vol. 119, pp. 537–556, Apr. 2014, ISSN: 0306-2619.

[67] R. E. Edwards, J. New, L. E. Parker, B. Cui, and J. Dong, “Constructing large scale surrogate models from big data and artificial intelligence,” *Applied Energy*, vol. 202, pp. 685–699, Sep. 2017, ISSN: 0306-2619.

[68] I. Korolija, Y. Zhang, L. Marjanovic-Halburd, and V. I. Hanby, “Regression models for predicting UK office building energy consumption from heating and cooling demands,” *Energy and Buildings*, vol. 59, pp. 214–227, Apr. 2013, ISSN: 0378-7788.

[69] J. C. Lam, K. K. W. Wan, D. Liu, and C. L. Tsang, “Multiple regression models for energy use in air-conditioned office buildings in different climates,” *Energy Conversion and Management*, vol. 51, no. 12, pp. 2692–2697, Dec. 2010, ISSN: 0196-8904.

[70] J. S. Hygh, J. F. DeCarolis, D. B. Hill, and S. Ranji Ranjithan, “Multivariate regression as an energy assessment tool in early building design,” *Building and Environment*, vol. 57, pp. 165–175, Nov. 2012, ISSN: 0360-1323.

[71] T. Catalina, V. Iordache, and B. Caracaleanu, "Multiple regression model for fast prediction of the heating energy demand," *Energy and Buildings*, vol. 57, pp. 302–312, Feb. 2013, ISSN: 0378-7788.

[72] F. Ascione, N. Bianco, C. De Stasio, G. M. Mauro, and G. P. Vanoli, "Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach," *Energy*, vol. 118, pp. 999–1017, Jan. 2017, ISSN: 0360-5442.

[73] Z. Romani, A. Draoui, and F. Allard, "Metamodeling the heating and cooling energy needs and simultaneous building envelope optimization for low energy building design in Morocco," *Energy and Buildings*, vol. 102, pp. 139–148, Sep. 2015, ISSN: 0378-7788.

[74] A. Rackes, A. P. Melo, and R. Lamberts, "Naturally comfortable and sustainable: Informed design guidance and performance labeling for passive commercial buildings in hot climates," *Applied Energy*, vol. 174, pp. 256–274, Jul. 2016, ISSN: 0306-2619.

[75] I. Jaffal, C. Inard, and C. Ghiaus, "Fast method to predict building heating demand based on the design of experiments," *Energy and Buildings*, vol. 41, no. 6, pp. 669–677, Jun. 2009, ISSN: 0378-7788.

[76] Y.-J. Kim, "Comparative study of surrogate models for uncertainty quantification of building energy model: Gaussian Process Emulator vs. Polynomial Chaos Expansion," *Energy and Buildings*, vol. 133, pp. 46–58, Dec. 2016, ISSN: 0378-7788.

[77] S. Singaravel, J. Suykens, and P. Geyer, "Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction," *Advanced Engineering Informatics*, vol. 38, pp. 81–90, Oct. 2018, ISSN: 1474-0346.

[78] I. Jaffal and C. Inard, "A metamodel for building energy performance," *Energy and Buildings*, vol. 151, pp. 501–510, Sep. 2017, ISSN: 0378-7788.

[79] P. Das, C. Shrubsole, B. Jones, I. Hamilton, Z. Chalabi, M. Davies, A. Mavrogianni, and J. Taylor, "Using probabilistic sampling-based sensitivity analyses for indoor air quality modelling," *Building and Environment*, vol. 78, pp. 171–182, Aug. 2014, ISSN: 0360-1323.

[80] A. Tsanas and A. Xifara, "Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools," *Energy and Buildings*, vol. 49, pp. 560–567, Jun. 2012, ISSN: 0378-7788.

[81] S. Papadopoulos and E. Azar, "Integrating building performance simulation in agent-based modeling using regression surrogate models: A novel human-in-the-loop energy modeling approach," *Energy and Buildings*, vol. 128, pp. 214–223, Sep. 2016, ISSN: 0378-7788.

[82] L. Magnier and F. Haghighat, "Multiobjective optimization of building design using TRNSYS simulations, genetic algorithm, and Artificial Neural Network," *Building and Environment*, vol. 45, no. 3, pp. 739–746, Mar. 2010, ISSN: 0360-1323.

[83] G. M. Stavrakakis, P. L. Zervas, H. Sarimveis, and N. C. Markatos, "Optimization of window-openings design for thermal comfort in naturally ventilated buildings," *Applied Mathematical Modelling*, vol. 36, no. 1, pp. 193–211, Jan. 2012, ISSN: 0307-904X.

[84] X. Chen and H. Yang, "A multi-stage optimization of passively designed high-rise residential buildings in multiple building operation scenarios," *Applied Energy*, vol. 206, pp. 541–557, Nov. 2017, ISSN: 0306-2619.

[85] E. E. Aydın, O. Dursun, I. Chatzikonstantinou, and B. Ekici, "Optimisation of energy consumption and daylighting using building performance surrogate model," en,

[86] E. Asadi, M. G. d. Silva, C. H. Antunes, L. Dias, and L. Glicksman, "Multi-objective optimization for building retrofit: A model using genetic algorithm and artificial neural network and an application," *Energy and Buildings*, vol. 81, pp. 444–456, Oct. 2014, ISSN: 0378-7788.

[87] T. Østergård, R. L. Jensen, and S. E. Maagaard, "A comparison of six metamodeling techniques applied to building performance simulations," *Applied Energy*, vol. 211, pp. 89–103, Feb. 2018, ISSN: 0306-2619.

[88] G. Zemella, D. De March, M. Borrotti, and I. Poli, "Optimised design of energy efficient building façades via Evolutionary Neural Networks," *Energy and Buildings*, vol. 43, no. 12, pp. 3297–3302, Dec. 2011, ISSN: 0378-7788.

[89] L. Van Gelder, H. Janssen, and S. Roels, "Metamodelling in robust low-energy dwelling design," eng, in *2nd Central European Symposium on Building Physics*, ISSN: 978-3-85437-321-6, Vienna University of Technology; Vienna, Sep. 2013, pp. 93–99.

[90] J.-S. Chou and D.-K. Bui, "Modeling heating and cooling loads by artificial intelligence for energy-efficient building design," *Energy and Buildings*, vol. 82, pp. 437–446, Oct. 2014, ISSN: 0378-7788.

[91] P. Symonds, J. Taylor, Z. Chalabi, and M. Davies, “Performance of Neural Networks vs. Radial Basis Functions When Forming a Metamodel for Residential Buildings,” en, *International Journal of Civil and Environmental Engineering*, vol. 9, no. 12, pp. 1594–1598, Nov. 2015.

[92] S. Yang, W. Tian, E. Cubi, Q. Meng, Y. Liu, and L. Wei, “Comparison of Sensitivity Analysis Methods in Building Energy Assessment,” *Procedia Engineering*, The 8th international cold climate HVAC Conference, vol. 146, pp. 174–181, Jan. 2016, ISSN: 1877-7058.

[93] R. Tkachenko, I. Izonin, N. Kryvinska, V. Chopyak, N. Lotoshynska, and D. Danyluk, “Piecewise-linear Approach for Medical Insurance Costs Prediction using SGTM Neural-Like Structure,” en,

[94] Y.-L. Grize, W. Fischer, and C. Lützelschwab, “Machine learning applications in non-life insurance,” en, *Applied Stochastic Models in Business and Industry*, vol. 36, no. 4, pp. 523–537, 2020, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asmb.2543, ISSN: 1526-4025.

[95] N. Zhang and H. Xu, “Fairness of Ratemaking for Catastrophe Insurance: Lessons from Machine Learning,” *Information Systems Research*, Jan. 2023, Publisher: INFORMS, ISSN: 1047-7047.

[96] D. Ersan, C. Nishioka, and A. Scherp, “Comparison of machine learning methods for financial time series forecasting at the examples of over 10 years of daily and hourly data of DAX 30 and S&P 500,” en, *Journal of Computational Social Science*, vol. 3, no. 1, pp. 103–133, Apr. 2020, ISSN: 2432-2725.

[97] J. Sirignano and R. Cont, *Universal features of price formation in financial markets: Perspectives from Deep Learning*, en, arXiv:1803.06917 [q-fin, stat], Mar. 2018.

[98] M. Jaggi, P. Mandal, S. Narang, U. Naseem, and M. Khushi, “Text Mining of Stocktwits Data for Predicting Stock Prices,” *Applied System Innovation*, vol. 4, p. 13, Feb. 2021.

[99] Y. Jiang, “Application and Comparison of Multiple Machine Learning Models in Finance,” en, *Scientific Programming*, vol. 2022, e9613554, Mar. 2022, Publisher: Hindawi, ISSN: 1058-9244.

[100] C. Xu, J. Li, B. Feng, and B. Lu, “A Financial Time-Series Prediction Model Based on Multiplex Attention and Linear Transformer Structure,” en, *Applied Sciences*, vol. 13, no. 8, p. 5175, Apr. 2023, ISSN: 2076-3417.

[101] A. Sridi and P. Bilokon, "Applying Deep Learning to Calibrate Stochastic Volatility Models," en, *SSRN Electronic Journal*, 2023, ISSN: 1556-5068.

[102] C. Bayer and B. Stemper, "Deep calibration of rough stochastic volatility models," en,

[103] X. Hui, B. Sun, I. SenGupta, Y. Zhou, and H. Jiang, "Stochastic volatility modeling of high-frequency CSI 300 index and dynamic jump prediction driven by machine learning," en, *Electronic Research Archive*, vol. 31, no. 3, pp. 1365–1386, 2023, arXiv:2204.02891 [q-fin], ISSN: 2688-1594.

[104] W. E, J. Han, and A. Jentzen, "Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations," en, *Communications in Mathematics and Statistics*, vol. 5, no. 4, pp. 349–380, Dec. 2017, arXiv:1706.04702 [cs, math, stat], ISSN: 2194-6701, 2194-671X.

[105] J. Fan and L. Mancini, "Option Pricing with Model-guided Nonparametric Methods," en,

[106] S. Liu, C. W. Oosterlee, and S. M. Bohte, "Pricing options and computing implied volatilities using neural networks," en, *Risks*, vol. 7, no. 1, p. 16, Feb. 2019, arXiv:1901.08943 [cs, math, q-fin], ISSN: 2227-9091.

[107] J. D. Spiegeleer, D. B. Madan, S. Reyners, and W. Schoutens, "Machine Learning for Quantitative Finance: Fast Derivative Pricing, Hedging and Fitting," en,

[108] B. Horvath, A. Muguruza, and M. Tomas, *Deep Learning Volatility*, en, arXiv:1901.09647 [q-fin], Aug. 2019.

[109] A. Hirsa, T. Karatas, and A. Oskoui, *Supervised Deep Neural Networks (DNNs) for Pricing/Calibration of Vanilla/Exotic Options Under Various Different Processes*, en, arXiv:1902.05810 [cs, q-fin], Feb. 2019.

[110] S. Liu, A. Borovykh, L. A. Grzelak, and C. W. Oosterlee, "A neural network-based framework for financial model calibration," en, *Journal of Mathematics in Industry*, vol. 9, no. 1, p. 9, Dec. 2019, ISSN: 2190-5983.

[111] S. Becker, P. Cheridito, and A. Jentzen, *Deep optimal stopping*, en, arXiv:1804.05394 [cs, math], Jan. 2020.

[112] T. Wang, R. Beard, J. Hawkins, and R. Chandra, *Recursive deep learning framework for forecasting the decadal world economic outlook*, en, arXiv:2301.10874 [cs], Jan. 2023.

[113] Y. Chen and J. W. L. Wan, *Deep Neural Network Framework Based on Backward Stochastic Differential Equations for Pricing and Hedging American Options in High Dimensions*, en, arXiv:1909.11532 [cs, q-fin, stat], Sep. 2019.

[114] V. Valaitis and A. T. Villa, "A Machine Learning Projection Method for Macro-Finance Models," en,

[115] J. Han, A. Jentzen, and W. E, "Solving high-dimensional partial differential equations using deep learning," *Proceedings of the National Academy of Sciences*, vol. 115, no. 34, pp. 8505–8510, Aug. 2018, Publisher: Proceedings of the National Academy of Sciences.

[116] H. Chen, A. Didisheim, and S. Scheidegger, "Deep Surrogates for Finance: With an Application to Option Pricing," en,

[117] S. Calcetero-Vanegas, A. L. Badescu, and X. S. Lin, *Effective a Posteriori Ratemaking with Large Insurance Portfolios via Surrogate Modeling*, en, arXiv:2211.06568 [q-fin, stat], May 2023.

[118] S. Yang, "Surrogate Model Assisted Nested Simulation with Applications to Variable Annuity Portfolio Valuation and Hedging," English, ISBN: 9798698548485, Ph.D. University of Toronto (Canada), Canada – Ontario, CA, 2020.

[119] M. Tomas, "Pricing and calibration of stochastic models via neural networks," en,

[120] A. Lazar, "Income prediction via support vector machine," en, in *2004 International Conference on Machine Learning and Applications, 2004. Proceedings.*, Louisville, Kentucky, USA: IEEE, 2004, pp. 143–149, ISBN: 978-0-7803-8823-9.

[121] H. M. Ezzat, "Estimating Behavioral Agent-Based Models for Financial Markets through Machine Learning Surrogates," en, *International Journal of Multidisciplinary Studies on Management, Business, and Economy*, vol. 5, no. 1, pp. 1–13, Jun. 2022, ISSN: 2735-5446.

[122] S. E. Rømer, "Empirical analysis of rough and classical stochastic volatility models to the SPX and VIX markets," en, *Quantitative Finance*, vol. 22, no. 10, pp. 1805–1838, Oct. 2022, ISSN: 1469-7688, 1469-7696.

[123] W. A. McGhee, "An Artificial Neural Network Representation of the SABR Stochastic Volatility Model," en,

[124] S. Liu, L. A. Grzelak, and C. W. Oosterlee, "The Seven-League Scheme: Deep Learning for Large Time Step Monte Carlo Simulations of Stochastic Differential Equations," en, *Risks*, vol. 10, no. 3, p. 47, Mar. 2022, Number: 3 Publisher: Multidisciplinary Digital Publishing Institute, ISSN: 2227-9091.

[125] J. R. Bertini, S. Ferreira Batista, M. A. Funcia, L. O. Mendes da Silva, A. A. S. Santos, and D. J. Schiozer, "A comparison of machine learning surrogate models for net present value prediction from well placement binary data," *Journal of Petroleum Science and Engineering*, vol. 208, p. 109 208, Jan. 2022, ISSN: 0920-4105.

[126] X. Chen, R. Chen, Q. Wan, R. Xu, and J. Liu, "An improved data-free surrogate model for solving partial differential equations using deep neural networks," en, *Scientific Reports*, vol. 11, no. 1, p. 19 507, Sep. 2021, ISSN: 2045-2322.

[127] W. Chung and J. H. Lee, "Input–Output Surrogate Models for Efficient Economic Evaluation of Amine Scrubbing CO2 Capture Processes," *Industrial & Engineering Chemistry Research*, vol. 59, no. 42, pp. 18 951–18 964, Oct. 2020, Publisher: American Chemical Society, ISSN: 0888-5885.

[128] T. Weber, A. Corotan, B. Hutchinson, B. Kravitz, and R. Link, "Technical note: Deep learning for creating surrogate models of precipitation in Earth system models," English, *Atmospheric Chemistry and Physics*, vol. 20, no. 4, pp. 2303–2317, Feb. 2020, Publisher: Copernicus GmbH, ISSN: 1680-7316.

[129] N. Shirzadi, A. Nizami, M. Khazen, and M. Nik-Bakht, "Medium-Term Regional Electricity Load Forecasting through Machine Learning and Deep Learning," en, *Designs*, vol. 5, no. 2, p. 27, Jun. 2021, Number: 2 Publisher: Multidisciplinary Digital Publishing Institute, ISSN: 2411-9660.

[130] L. Schröder, N. Krasimirov Dimitrov, D. R. Verelst, and J. A. Sørensen, "Wind turbine site-specific load estimation using artificial neural networks calibrated by means of high-fidelity load simulations," en, *Journal of Physics: Conference Series*, vol. 1037, p. 062 027, Jun. 2018, ISSN: 1742-6588, 1742-6596.

[131] M. Steininger, D. Abel, K. Ziegler, A. Krause, H. Paeth, and A. Hotho, *Deep Learning for Climate Model Output Statistics*, en, arXiv:2012.10394 [physics], Dec. 2020.

[132] X. Huang, *Deep-learning based climate downscaling using the super-resolution method: A case study over the western US*, en, Sep. 2020.

[133] S. Park, K. Singh, A. Nellikkattil, E. Zeller, T. D. Mai, and M. Cha, "Downscaling Earth System Models with Deep Learning," en, in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Washington DC USA: ACM, Aug. 2022, pp. 3733–3742, ISBN: 978-1-4503-9385-0.

[134] T. Bolton and L. Zanna, "Applications of Deep Learning to Ocean Data Inference and Subgrid Parameterization," en, *Journal of Advances in Modeling Earth Systems*, vol. 11, no. 1, pp. 376–399, 2019, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2018MS0 ISSN: 1942-2466.

[135] A. Nikolaev, I. Richter, and P. Sadowski, "Deep Learning for Climate Models of the Atlantic Ocean," en,

[136] M. Andrychowicz, L. Espeholt, D. Li, S. Merchant, A. Merose, F. Zyda, S. Agrawal, and N. Kalchbrenner, *Deep Learning for Day Forecasts from Sparse Observations*, arXiv:2306.06079 [physics], Jul. 2023.

[137] P. Grönquist, C. Yao, T. Ben-Nun, N. Dryden, P. Dueben, S. Li, and T. Hoefler, "Deep learning for post-processing ensemble weather forecasts," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 379, no. 2194, p. 20 200 092, Feb. 2021, Publisher: Royal Society.

[138] S. Rasp, M. S. Pritchard, and P. Gentine, "Deep learning to represent sub-grid processes in climate models," *Proceedings of the National Academy of Sciences*, vol. 115, no. 39, pp. 9684–9689, Sep. 2018, arXiv:1806.04731 [physics, stat], ISSN: 0027-8424, 1091-6490.

[139] V. M. Johnson and L. L. Rogers, "Accuracy of Neural Network Approximators in Simulation-Optimization," EN, *Journal of Water Resources Planning and Management*, vol. 126, no. 2, pp. 48–56, Mar. 2000, Publisher: American Society of Civil Engineers, ISSN: 0733-9496.

[140] S.-T. Khu and M. G. F. Werner, "Reduction of Monte-Carlo simulation runs for uncertainty estimation in hydrological modelling," English, *Hydrology and Earth System Sciences*, vol. 7, no. 5, pp. 680–692, Oct. 2003, Publisher: Copernicus GmbH, ISSN: 1027-5606.

[141] S.-T. Khu, D. Savic, Y. Liu, and H. Madsen, "A fast Evolutionary-based Meta-Modelling Approach for the Calibration of a Rainfall-Runoff Model," *International Congress on Environmental Modelling and Software*, Jul. 2004.

[142] D. R. Broad, G. C. Dandy, and H. R. Maier, "Water Distribution System Optimization Using Metamodels," EN, *Journal of Water Resources Planning and Management*, vol. 131, no. 3, pp. 172–180, May 2005, Publisher: American Society of Civil Engineers, ISSN: 0733-9496.

[143] K. Kobayashi, M. Okumura, H. Nakamura, M. Itakura, M. Machida, and M. W. D. Cooper, "Machine learning molecular dynamics simulations toward exploration of high-temperature properties of nuclear fuel materials: Case study of thorium dioxide," eng, *Scientific Reports*, vol. 12, no. 1, p. 9808, Jun. 2022, ISSN: 2045-2322.

[144] A. Edelen, N. Neveu, M. Frey, Y. Huber, C. Mayes, and A. Adelmann, "Machine learning for orders of magnitude speedup in multiobjective optimization of particle accelerator systems," en, *Physical Review Accelerators and Beams*, vol. 23, no. 4, p. 044 601, Apr. 2020, ISSN: 2469-9888.

[145] M. Paganini, L. de Oliveira, and B. Nachman, "Accelerating Science with Generative Adversarial Networks: An Application to 3D Particle Showers in Multilayer Calorimeters," *Physical Review Letters*, vol. 120, no. 4, p. 042 003, Jan. 2018, Publisher: American Physical Society.

[146] D. Kochkov, J. A. Smith, A. Alieva, Q. Wang, M. P. Brenner, and S. Hoyer, "Machine learning–accelerated computational fluid dynamics," *Proceedings of the National Academy of Sciences*, vol. 118, no. 21, e2101784118, May 2021, Publisher: Proceedings of the National Academy of Sciences.

[147] R. Hu, F. Fang, C. C. Pain, and I. M. Navon, "Rapid spatio-temporal flood prediction and uncertainty quantification using a deep learning method," *Journal of Hydrology*, vol. 575, pp. 911–920, Aug. 2019, ISSN: 0022-1694.

[148] B. Moseley, T. Nissen-Meyer, and A. Markham, "Deep learning for fast simulation of seismic waves in complex media," English, *Solid Earth*, vol. 11, no. 4, pp. 1527–1549, Aug. 2020, Publisher: Copernicus GmbH, ISSN: 1869-9510.

[149] E. Buhmann, S. Diefenbacher, E. Eren, F. Gaede, G. Kasieczka, A. Korol, and K. Krüger, "Getting High: High Fidelity Simulation of High Granularity Calorimeters with High Speed," *Computing and Software for Big Science*, vol. 5, no. 1, p. 13, Dec. 2021, arXiv:2005.05334 [hep-ex, physics:hep-ph, physics:physics], ISSN: 2510-2036, 2510-2044.

[150] A. Ghosh, X. Ju, B. Nachman, and A. Siodmok, "Towards a Deep Learning Model for Hadronization," *Physical Review D*, vol. 106, no. 9, p. 096 020, Nov. 2022, arXiv:2203.12660 [hep-ex, physics:hep-ph, physics:physics], ISSN: 2470-0010, 2470-0029.

[151] T. Takhtaganov, Z. Lukić, J. Müller, and D. Morozov, "Cosmic Inference: Constraining Parameters with Observations and a Highly Limited Number of Simulations," en, *The Astrophysical Journal*, vol. 906, no. 2, p. 74, Jan. 2021, Publisher: The American Astronomical Society, ISSN: 0004-637X.

[152] P. Harrington, M. Mustafa, M. Dornfest, B. Horowitz, and Z. Lukić, "Fast, High-fidelity Ly Forests with Convolutional Neural Networks," en, *The Astrophysical Journal*, vol. 929, no. 2, p. 160, Apr. 2022, Publisher: The American Astronomical Society, ISSN: 0004-637X.

[153] B. W. (W. Parker, "An automatic, multi-fidelity framework for optimizing the performance of super-cavitating hydrofoils using Gaussian process regression and Bayesian optimization," eng, Accepted: 2018-10-22T18:46:17Z, Thesis, Massachusetts Institute of Technology, 2018.

[154] D. Kochkov, J. A. Smith, A. Alieva, Q. Wang, M. P. Brenner, and S. Hoyer, "Machine learning–accelerated computational fluid dynamics," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, no. 21, e2101784118, May 2021, ISSN: 0027-8424.

[155] A. Duval, V. Schmidt, S. Miret, Y. Bengio, A. Hernández-García, and D. Rolnick, *PhAST: Physics-Aware, Scalable, and Task-specific GNNs for Accelerated Catalyst Design*, arXiv:2211.12020 [physics], Mar. 2024.

[156] J. Tubiana, D. Schneidman-Duhovny, and H. J. Wolfson, "ScanNet: An interpretable geometric deep learning model for structure-based protein binding site prediction," en, *Nature Methods*, vol. 19, no. 6, pp. 730–739, Jun. 2022, Publisher: Nature Publishing Group, ISSN: 1548-7105.

[157] M. F. Kasim, D. Watson-Parris, L. Deaconu, *et al.*, "Building high accuracy emulators for scientific simulations with deep neural architecture search," en, *Machine Learning: Science and Technology*, vol. 3, no. 1, p. 015 013, Mar. 2022, arXiv:2001.08055 [physics, stat], ISSN: 2632-2153.

[158] Y. Lai, "A Comparison of Traditional Machine Learning and Deep Learning in Image Recognition," en, *Journal of Physics: Conference Series*, vol. 1314, no. 1, p. 012 148, Oct. 2019, ISSN: 1742-6588, 1742-6596.

[159] School of Computer Engineering, Suranaree University of Technology, Nakhon Ratchasima, 30000, Thailand, P. Poomka, N. Kerdprasop, and K. Kerdprasop, "Machine Learning Versus Deep Learning Performances on the Sentiment Analysis of Product Reviews,"

International Journal of Machine Learning and Computing, vol. 11, no. 2, pp. 103–109, Mar. 2021, ISSN: 20103700.

[160] L. Li, S. Pandey, T. Flynn, H. Liu, N. Wheeler, and A. Hoisie, "SimNet: Accurate and High-Performance Computer Architecture Simulation using Deep Learning," in *Abstract Proceedings of the 2022 ACM SIGMETRICS/IFIP PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems*, ser. SIGMETRICS/PERFORMANCE '22, New York, NY, USA: Association for Computing Machinery, Jun. 2022, pp. 67–68, ISBN: 978-1-4503-9141-2.

[161] C. van Leeuwen, D. Podareanu, V. Codreanu, *et al.*, "Deep-learning enhancement of large scale numerical simulations," en,

[162] S. Ghassemzadeh, M. Gonzalez Perdomo, M. Haghighi, and E. Abbasnejad, "A data-driven reservoir simulation for natural gas reservoirs," en, *Neural Computing and Applications*, vol. 33, no. 18, pp. 11 777–11 798, Sep. 2021, ISSN: 1433-3058.

[163] J. Chen, R. Du, P. Li, and L. Lyu, *Quasi-Monte Carlo sampling for machine-learning partial differential equations*, en, arXiv:1911.01612 [cs, math], Nov. 2019.

[164] W. Gao, X. Lu, Y. Peng, and L. Wu, "A Deep Learning Approach Replacing the Finite Difference Method for In Situ Stress Prediction," en, *IEEE Access*, vol. 8, pp. 44 063–44 074, 2020, ISSN: 2169-3536.

[165] Y. Zamora, "Machine Learning for Performance Acceleration and Prediction in Scientific Computing," en, Ph.D. dissertation, The University of Chicago.

[166] I. Nitze, U. Schulthess, and H. Asche, *COMPARISON OF MACHINE LEARNING ALGORITHMS RANDOM FOREST, ARTIFICIAL NEURAL NETWORK AND SUPPORT VECTOR MACHINE TO MAXIMUM LIKELIHOOD FOR SUPERVISED CROP TYPE CLASSIFICATION*. May 2012.

[167] H. Vardhan and J. Sztipanovits, "Deep Learning-based Finite Element Analysis (FEA) surrogate for sub-sea pressure vessel," Jun. 2022.

[168] J. Q. Toledo-Marín, G. Fox, J. P. Sluka, and J. A. Glazier, "Deep Learning Approaches to Surrogates for Solving the Diffusion Equation for Mechanistic Real-World Simulations," English, *Frontiers in Physiology*, vol. 12, Jun. 2021, Publisher: Frontiers, ISSN: 1664-042X.

[169] P. Yadav, N. Gupta, and P. K. Sharma, “A comprehensive study towards high-level approaches for weapon detection using classical machine learning and deep learning methods,” *Expert Systems with Applications*, vol. 212, p. 118 698, Feb. 2023, ISSN: 0957-4174.

[170] Z. Zheng, Y. Xu, and P. Wang, “Uncertainty Quantification Analysis on Mechanical Properties of the Structured Silicon Anode via Surrogate Models,” en, *Journal of The Electrochemical Society*, vol. 168, no. 4, p. 040 508, Apr. 2021, Publisher: IOP Publishing, ISSN: 1945-7111.

[171] Q. Xu, Z. Nie, H. Xu, H. Zhou, H. R. Attar, N. Li, F. Xie, and X.-J. Liu, “SuperMeshing: A New Deep Learning Architecture for Increasing the Mesh Density of Physical Fields in Metal Forming Numerical Simulation,” *Journal of Applied Mechanics*, vol. 89, no. 011002, Sep. 2021, ISSN: 0021-8936.

[172] A. G. Baydin, B. Poduval, and N. A. Schwadron, “A Surrogate Model for Studying Solar Energetic Particle Transport and the Seed Population,” en, *Space Weather*, vol. 21, no. 12, e2023SW003593, 2023, ISSN: 1542-7390.