

Hybrid Machine Learning and CFD Framework for Predictive Thermal Management in Compact Heat Exchangers: Simulation-Driven Optimization Using MATLAB and Python

Yogesh Subhash Kulkarni¹, Sharad Ramdas Gawade², Yogesh Sanjay Pathare³, Padmakar Dagdu Kabudake⁴, Sancheti Santosh D⁵, Rahul k Bawane⁶, Harish Phadtare⁷, Rudresha S⁸

¹Department of Mechanical Engineering, SNJB's late Sau KBJ COE, Chandwad, Pune .

²Department of Mechanical Engineering, Sharadchandra Pawar College of Engineering and Technology, Someshwar, Baramati - 412306, Karnataka,

³Department of Mechanical Engineering, Dr. Vithalrao Vikhe Patil College of Engineering, Ahmednagar, Maharashtra.

⁴Department of Mechanical Engineering, Pravara Rural Engineering Loni , Pune.

⁵Department of Mechanical Engineering, SNJB's Late Sau.K.B.Jain College of Engineering, Chandwad, Maharashtra.

⁶Department of Mechanical Engineering, Pimpri Chinchwad College of Engineering & Research, Maharashtra

⁷Department of Mechanical Engineering, SVPM's College of Engineering, Malegaon, B K Baramatti, Maharashtra

⁸Department of Mechanical Engineering, University of Visvesvaraya College of Engineering K.R Circle, Bengaluru, Karnataka

Abstract

For automotive, aerospace, and electronic cooling purposes, ensuring efficient thermal management is of the utmost importance in compact heat exchangers. While traditional CFD processes offer high-fidelity insight, complexities of such processes grow with iterative design and optimization. This research attempts to integrate CFD with ML for a quicker and reliable surrogate model generation that can predict core performance parameters such as Nusselt number, outlet temperature, and pressure drop. A MATLAB-based CFD automation workflow was executed to simulate various heat exchanger geometries in steady-state conditions. This data set was used to train multiple ML models including Random Forest and ANN, which inadvertently yielded greater than 99% prediction accuracies for all target parameters. The surrogate models were then incorporated into a multi-objective optimization loop that sought to maximize thermal performance and minimize pressure drop. A Pareto front of optimal design configurations was found at a computational time savings of at least 1000× compared to doing the optimization directly through CFD. Experimental validation on a 3D-printed prototype confirmed the surrogate model's reliability, with prediction errors under 5%. The proposed CFD-ML integration enables rapid, physics-informed design exploration and is scalable to real-time thermal management applications.

Keywords: CFD, Machine Learning, Surrogate Model, Heat Exchanger, Thermal Optimization, MATLAB, Multi-Objective Optimization, Random Forest, ANN, Nusselt Number, Pressure Drop

1. Introduction

1.1 Motivation for Efficient Thermal Management Systems

This is particularly evident in compact high-power-density applications such as electronic devices, aerospace components, electric vehicles, and renewable energy systems. As engineering technologies evolve and promise to provide increased operational efficiencies, heat loads also must be accommodated and dissipated, which increases the possibility of thermal failure when heat loads cannot be managed or adequately dissipated (Shah & Sekulic, 2021; Zhao et al., 2022). Compact heat exchangers have developed into a central technology in achieving effective thermal management, achieving high heat transfer rates at small space angles, but their performance is sensitive to important geometric and flow design characteristics. Therefore, modelling and optimizing these systems is very important in moving heat management forward, as well as, innovative thermal-efficiency concepts.

1.2 Limitations of Conventional CFD-Only Approaches

Computational Fluid Dynamics (CFD) is widely used to simulate heat transfer and fluid flow in such systems, because it simulates governing physical laws (e.g. Navier-Stokes and energy equations) at a high spatial resolution. As accurate as CFD can be, it also has substantial restrictions as a practical design environment. Each simulation is both computationally prohibitive and time handwegng to complete forming iteration of parallel variable or other intervening characteristics, such as transients, transitory effects, or turbulent flow. (Rao & Das, 2022) Iterative design by definition adds many cycles of simulation to a CFD design without corrected, faster results and thus would continue to be considered impractical for timely design or optimization of real time applications. These limitations can constrain the CFD regard as a stand alone option for aesthetic, dynamic, or conceivably couple engineering or design, if it is larger in function.

1.3 Benefits of ML-CFD Integration

Recently, machine learning (ML) techniques have made it possible to create data-driven surrogate models that balance thermal and fluid performance from very few high-fidelity computational fluid dynamics (CFD) results. In particular, ML techniques including artificial neural networks (ANN), random forests, and gradient boosting have demonstrated useful capabilities for approximating nonlinear effects in heat transfer and provided significant time savings over CFD, with similar accuracy (Vinueza & Brunton, 2022; Goswami et al., 2021). After obtaining a dataset generated by CFD, ML techniques allow researchers to work around computational challenges, not only permitting real-time prediction but also sensitivity analysis and inverse design. The convergence of physics-based and data-driven approaches is beneficial because engineers can access the best of both disciplines: the accuracy of CFD models and fast, generalized solutions through ML (Jin et al., 2023; Tian et al., 2023).

1.4 Objectives and Structure of the Paper

This study proposes a novel hybrid framework combining CFD simulations and machine learning models to predict and optimize heat transfer performance in compact heat exchangers. MATLAB is used to automate and execute CFD simulations across multiple geometric and boundary configurations. The resulting dataset is used to train Python-based ML models—specifically Random Forest, XGBoost, and ANN—to construct fast surrogate models of thermal performance metrics such as Nusselt number, pressure drop, and outlet temperature. The objectives of the study are:

- To develop a high-fidelity CFD model of a compact heat exchanger and generate a comprehensive parametric dataset.
- To train and compare various ML models for accurate prediction of heat transfer performance.
- To demonstrate the computational advantages and predictive accuracy of the ML-CFD hybrid model in thermal system design.

The rest of the paper is structured as follows: Section 2 presents the literature review on CFD applications and ML integration in thermal systems. Section 3 details the methodology, including CFD modelling, data preprocessing, and ML model training. Section 4 discusses the results and model evaluation. Section 5 concludes the findings and outlines future research directions.

2.1 Studies on CFD in Heat Exchangers

Computational Fluid Dynamics (CFD) is widely recognized as a powerful and versatile tool to analyse and ultimately optimize heat exchangers, demonstrating excellent capability to create flow behaviour, temperature fields, and pressure drops over a spectrum of operating conditions. In compact and micro-heat exchangers, where classic empirical correlations often fail, CFD provides the means to map a multitude of local effects such as flow separation, recirculation, secondary vortices, and developing thermal boundary layers, all important heat transfer enhancement mechanisms (Nasiruddin et al., 2020).

New studies have shown that CFD can be employed to optimize geometric parameters, such as fin shapes, baffle positions, and channel geometries. For example, Bhattacharyya et al. (2021) applied CFD to assess the effects of louvered fin spacing and angle on the thermal-hydraulic performance of compact plate-fin heat exchangers, discovering that optimized fin geometry could increase the Nusselt number while maintaining a low pressure drop. Similarly, Ghadikolaie et al. (2022) conducted a 3D CFD study of helically coiled heat exchangers with nanofluids and observed that the secondary flow caused by the curvature could significantly improve heat transfer performance.

CFD has proven useful in assessing the response of shell-and-tube configurations to changes in baffle distance, baffle cut and arrangement of tubes. Mirkamali and Saffar-Avval (2021) used a multi-objective mode CFD analysis on a shell-and-tube heat exchanger and showed great improvement of thermal and hydraulic performance results through changes to the shell side

configurations. CFD has also allowed for the construction of new heat exchanger topology (i.e. sinusoidal and dimpled channels) that would be difficult for traditional methods to model (Wu et al., 2020).

It is also worth noting the growing application of CFD in microscale heat exchanger research. At these scales, effects such as rarefied gas effects, surface roughness, and heat transfer characteristics in the entrance region tend to dominate. Zhou et al. (2023) presented a microchannel heat exchanger with trapezoidal cross-section, simulated using CFD, and illustrated how slight geometric changes at the microscale lead to significantly improved heat transfer coefficient and thermal uniformity properties.

In conclusion, CFD provides accurate characterization of thermal performance and serves as the basis of datasets for data-driven modeling and surrogate models. But even though CFD has high fidelity, it is computationally expensive and unable to be applied directly in large-scale optimization problems or thermal control problems requiring real-time information. So, we turn to hybrid methods that mix CFD and machine learning.

ML in Predicting Key Thermal Parameters

The latest research shows that machine learning (ML) methods may effectively forecast basic thermal quantities like the Nusselt number, convective heat transfer coefficients, and wall temperatures in complicated systems. For instance, Zhou et al. (2020) created ML models to forecast condensation heat transfer coefficients in mini/micro-channel flows and proved that data-driven algorithms may act as "a robust new predicting tool" for two-phase heat transfer (Zhou et al. (2022)). In convective uses, researchers have used ML to model the correlation between flow conditions and Nusselt numbers in an absence of explicit correlations. Nguyen et al. (2024) used artificial neural network (ANN) and random forest models to forecast Nusselt numbers in a falling-film evaporator, with high accuracy in predicting the heat transfer performance of the evaporator (within a few percent of experimental values). These ML-based methods successfully capture the nonlinear relationships of thermal parameters with flow regime, geometry, and fluid properties and frequently outperform conventional empirical correlations. These results demonstrate that ML techniques have the capability to generalize over a broad spectrum of operational conditions to predict principal thermal metrics such as local wall temperatures and average heat transfer coefficients with very high fidelity, provided that an adequate amount of training data are available.

Common Machine Learning Algorithms in Heat Transfer Studies

Throughout the literature, numerous supervised machine learning algorithms have been utilized to make heat transfer predictions. The preferred option is artificial neural networks (ANNs) (deep neural networks), appreciated for their capability to model complex nonlinear mappings (frontiersin.org). For example, basic feed-forward ANNs have already been employed for forecasting heat exchanger performance, and more recent sophisticated architectures (e.g. LSTM or Transformer networks) have been proposed for transient or sequence-based thermal problems. Support Vector Regression (SVR) is also a widely employed algorithm, particularly in regression problems where data are scarce, because it is highly resilient in high-dimensional feature space. Tree-based ensemble approaches are also leading the way: Random Forests and

gradient boosting methods such as XGBoost have also demonstrated excellent predictive power and robustness for heat transfer data. To illustrate, in a recent comparison of printed circuit heat exchanger data, an ANN was found to have the best accuracy ($R^2 \approx 0.892$) among "five traditional machine learning models" evaluated, yet tree-based models (Random Forest, XGBoost, CatBoost) were close alternates and provided interpretable input feature importance. Similarly, SVR and other kernel-based approaches have been used successfully for interpolating nonlinear thermal-fluid behavior. These algorithms are usually chosen depending on the dataset size and the required balance between accuracy and interpretability. In conclusion, modern heat transfer research generally uses ANNs (for their accuracy on large data), support vector machines, and ensemble tree methods (for their consistency and speed of training), with some studies also testing newer deep learning and hybrid models.

Integration of ML with CFD: Surrogates and Optimization

A notable trend is coupling ML with computational fluid dynamics (CFD) to develop surrogate models that rapidly speed up thermal analyses. High-fidelity CFD simulations take a lot of computation, while an ML-trained surrogate can estimate results nearly in real-time. Recent publications show that ML surrogates can mimic heat transfer results estimated by CFD with little loss of accuracy while reducing evaluation time by orders of magnitude. For example, a study on a small heat exchanger found that a machine learning model was able to forecast the performance of the device in less than an hour compared to ~9 hours needed by a CFD simulation of the same problem. Surrogate models are particularly useful for design optimization and real-time control. Engineers are able to quickly search over design spaces—changing geometry or flow parameters—to minimize or maximize targets such as thermal efficiency, pressure drop, or heat transfer rate using ML surrogates. ML-based multi-objective optimization has been shown on heat exchanger design, where the surrogate models many performance measures (e.g. Nusselt number, friction factor, cost) simultaneously to inform optimal design selection. Involving physics expertise within ML models is another frontier.

Physics-informed neural networks (PINNs) and other multi-modal strategies impose physical conservation principles or boundary conditions on the learning process. Such addition enhances model robustness and extrapolation. For instance, Kokash et al. (2024) used PINNs in conjunction with data-driven models to a lid-driven cavity heat transfer problem and discovered that the model based on PINNs could honor underlying Navier–Stokes physics and provide accurate flow and thermal field predictions. In their work, a conventional data-trained model (random forest) yielded robust and low-error estimates of mean Nusselt number for cavity flow, and the PINN strategy further improved accuracy by adding physical constraints.

These strategies demonstrate how physical principles and CFD-generated data can be combined with ML to generate quick, but physics-consistent, prediction tools. Overall, the combination of CFD and ML - utilizing the former for creating training data and the latter for designing high-speed-executing surrogates - is making real-time simulation and optimization of heat transfer systems that were otherwise not computationally tractable with brute-force CFD possible.

Industrial and Academic Use Cases (Heat Exchangers, Microchannels, HVAC)

Compact heat exchangers have been significantly improved by machine learning (ML) modeling. Intricate devices such as plate-fin exchangers, printed circuit heat exchangers (PCHEs), and finned-tube coils have complex geometries and numerous parameters to be considered. ML methods have been employed to model and optimize the performance of such devices. To show, e.g., Li et al. (2023) employed an artificial neural network (ANN) trained on simulation data for a PCHE handling supercritical methane and attained an impressive $R^2 \approx 0.9996$ on prediction of heat transfer coefficient. This indicates that ML can model the complex thermohydraulic behavior of compact exchangers very well. Subsequent work by Su et al. (2025) built on this by deep learning: by comparing ANNs, Random Forests, support vector regression (SVR), and state-of-the-art recurrent models on PCHE data, they found that ANNs worked best among the simple ML algorithms, whereas enhanced deep neural networks (with attention mechanisms) could almost predict thermal performance perfectly. These data-driven models have also been employed to predict fouling in heat exchangers and to identify faults, where they can rapidly identify patterns of loss of performance that are hard to spot using simple principles.

In microchannel cooling devices, such as cooling high-heat-flux electronics and miniature heat sinks, machine learning (ML) has also been shown to have great promise. Researchers have developed models based on artificial neural networks (ANN) to determine how heat flows in microchannels, even when solutions are added to improve the flow. Gönül et al. (2022), for instance, utilized multiple neural network models to approximate the heat transfer behavior of a microchannel heat sink with miniature vortex generators. They determined that ML predictions were closely in agreement with detailed computational fluid dynamics (CFD) solutions for various shapes. The technique enables rapid checks of design modifications, for instance, altering the microchannel shape or introducing devices to generate turbulence. This is relevant in designing the next generation of microelectronic cooling devices under pressure drop control. By being trained on a broad spectrum of microchannel operating conditions, the ML models are able to appreciate the subtle effects of size and fluid motion that standard calculations neglect.

Machine learning is applied in the HVAC industry to components such as evaporator and condenser coils and to controlling energy in systems. An example is applying machine learning to forecast the performance of finned-tube evaporators under varying humidity. Liang et al. (2024) developed a specific deep learning model for a finned-tube HVAC evaporator that achieved mean errors less than 0.5% in forecasting heat transfer rates and pressure drops. The model could efficiently demonstrate the transition between dry and wet coil operations and outperformed conventional methods in all operating conditions. This precision in forecasting coil performances aids in designing more efficient air-conditioning systems and enables predictive control, such as forecasting the response of an evaporator to changes in load or inlet conditions. Aside from single components, machine learning has also been applied in HVAC to predict building heating and cooling requirements and to enhance control strategies, at times employing data from IoT sensors, demonstrating its applicability to large thermal systems.

In brief, during 2020-2024, numerous individuals have shown interest to implement machine learning (ML) for heat transfer issues. These cover the simple to complex, like precisely predicting elementary variables (e.g., the Nusselt number and heat transfer coefficients) all the way through to designing models that can serve as surrogates for costly computational fluid dynamics (CFD) simulations, and even to enabling optimization and smart control for better industrial devices (e.g., compact heat exchangers, microchannel coolers, and HVAC units). The research examined here indicates that, with well-prepared data sets (experimental or simulated) and appropriate algorithms (such as artificial neural networks, ensemble trees, and support vector machines), ML can significantly enhance prediction capability and provide greater insight into thermal-fluid phenomena. This work indicates that data-driven models will increasingly find value in heat transfer research and engineering design, complemented by physics-based approaches to maximize performance and efficiency in thermal systems.

2.3 Recent Hybrid ML-CFD Frameworks

The convergence of machine learning (ML) and computational fluid dynamics (CFD) has led to the development of hybrid frameworks that create a synergistic balance between the benefits of physical accuracy associated with numerical simulations with the speed and efficiency of a flexible data-driven approach. Hybrid ML-CFD methods are used increasingly to mitigate the expense associated with computational fluid dynamics simulations - especially in complex, multi-objective thermal systems, such as - heat exchangers, microchannel cooling systems, HVAC components, etc.

A key trend in recent literature is the creation of surrogate models that are trained to data generated by CFD simulations. These models do not replicate CFD solvers but approximate the outcomes of CFD simulations by exploring the relationship between geometric, boundary, or operating parameters and thermal performance metrics (e.g., Nusselt number, pressure loss, temperature fields). For example, Wang et al. (2023) created a gradient boosting regression tree (GBRT) model trained on 500 CFD simulations of a printed circuit heat exchanger (PCHE). The GBRT predicted both heat transfer coefficients and pressure losses with an R^2 of 0.98, and then used the surrogate model to optimize the geometry of the exchanger whilst achieving 95% reduction in the total computational cost, compared to direct optimization by coupling CFD.

Physics-informed neural networks (PINNs) represent a new direction. These models physically incorporate the laws of physics into the loss function of the neural network (e.g., Navier–Stokes, energy conservation). Raissi et al. (2020) introduced PINNs to solve forward and inverse problems in fluid and heat transfer dynamics by using PINNs to effectively do simulations without full numerical solvers. Zhu et al. (2022) applied PINNs to predict temperature distributions in conjugate heat transfer problems resulting in similar results to OpenFOAM simulations while requiring significantly less training data.

A third area of active research is the combination of ML with reduced-order models (ROMs) to perform unsteady or 3D simulations of heat transfer more efficiently. Sanchez-Gonzalez et al. (2021) coupled convolutional neural networks (CNNs) with POD-Galerkin ROMs to simulate turbulent thermal mixing in T-junctions, a common configuration in heat exchanger

manifolds. The controlled accuracy of the hybrid model maintained key flow features found in the high-fidelity CFD while achieving an $8\times$ speedup of the simulation.

The adoption of hybrid approaches in design optimization processes is also on the rise. Li et al. (2022) used computational fluid dynamics (CFD) in ANSYS Fluent to develop a deep neural network (DNN) to predict temperature uniformity and thermal resistance in a finned microchannel heat sink. They then treated this model as a black box model to incorporate into a genetic algorithm for multi-objective optimization, which produced designs with a 22% reduction in maximum temperatures and 30% improvement in thermal performance compared to baseline.

In addition, transfer learning approaches have been investigated to allow CFD-trained ML models to be generalized to different boundary conditions or geometries. For example, Tian et al. (2023) trained a CNN model on heat transfer in circular channels and then used only a small amount of new data to fine-tune the model to elliptical geometries. The authors achieved accurate predictions but used 80% fewer CFD runs than before.

Open FOAM with Python ML Libraries (TensorFlow, Scikit-learn), COMSOL + MATLAB, and ANSYS Fluent with standalone ML scripts for data pulling and model training are all common software environments for hybrid methods. Their combination provides a seamless system for generating, manipulating, and analysing CFD data within the ML workflow.

Overall, hybrid ML-CFD frameworks can be beneficial in speeding up thermal simulations, real-time prediction, and optimizing complex heat transfer systems. Despite the approximation with respecting physical fidelity, these approaches drastically reduce the computations requested and will only become more useful as deep learning and physics-informed modelling further advance.

2.4 Gaps Identified and Justification of Novelty

While the integration of machine learning (ML) and computational fluid dynamics (CFD) has made great strides over the past few years for thermal analysis, there still remain significant research gaps in the literature. Most studies today do one of three things carefully: studies in isolation of the path in the modelling pipeline such as surrogate prediction of Nusselt numbers (Zhou et al., 2020), deterministic ML-assisted heat exchanger design in steady-state (Li et al., 2022), and even studies that simply bring data generation, model training, and performance optimization in a closed loop. These studies demonstrate that there is still work to be done to operationalize hybrid models to enable use for rapid-prototyping and intelligent design as intended in the composable path for operation with ML and CFD (Wang et al., 2023).

Another significant limitation is the absence of models which are generalizable surrogate models that could be developed and used over several heat exchanger types or ranges of operation. Many ML models are trained on very narrow datasets that were designed for certain geometries (rectangular or circular channels for example), and have poorly validated performance in extrapolating (Tian et al., 2023). Though ML techniques in transfer learning and domain adaptation have potential for great contributions to the flexibility of ML-CFD frameworks, they are yet to be widely utilized in the use of thermos fluid systems. Moreover,

a large amount of the existing surrogate models function as, "black box," models with low physical interpretability and with little to no ability to enforce thermodynamic consistency (Vinuesa & Brunton, 2022). This ultimately hinders trust in instantiating these models using an explicit need for some level of regulatory compliance and a level of design traceability for safety critical or industrial applications.

Moreover, the literature is biased toward steady-state simulations and does not adequately represent transient or dynamic heat transfer scenarios, as seen in real applications with electric vehicle cooling, electronics cooling, and HVAC equipment. Although PINNs and reduced-order models have emerged in this area, they are typically shown in simplified geometries or laminar flows (Zhu et al., 2022), therefore, there exist performance limitations for complex turbulent flows or multi-phase systems.

From a computational standpoint, while many hybrid frameworks insist on better times compared to complete CFD, meaningful computational savings with respect to training time, memory use, and scalability will be quantified. Most surrogate models are constructed on trial-and-error with little emphasis on discussion on model uncertainty or error propagation, inhibiting their strength for optimization (Rao & Das, 2022). In addition, automated pipelines with a coupling of CFD pre-processing, ie. geometry/mesh generation, ML, and post-processing are overwhelmingly absent in a comprehensive approach.

In light of these observations, the novelty of the present study lies in the following contributions:

- Development of a fully integrated CFD-ML hybrid framework that combines MATLAB-based CFD data generation and Python-based supervised learning models (Random Forest, XGBoost, and ANN) for compact heat exchanger optimization.
- Generation of a comprehensive dataset through systematic CFD simulations covering a wide range of geometric and flow parameters to ensure model generalizability.
- Implementation of multi-objective surrogate modeling that simultaneously predicts thermal and hydraulic performance metrics (e.g., Nusselt number, pressure drop, temperature uniformity).
- Quantitative benchmarking of computational savings versus traditional CFD, with performance metrics such as RMSE, R^2 , training time, and inference latency.
- Emphasis on design optimization and sensitivity analysis, showcasing the real-time application potential of the trained surrogate models.

This contribution addresses several of the outlined gaps and sets the stage for creating a reliable, fast, and flexible hybrid modeling pipeline applicable to modern thermal system design.

3. Methodology

3.1 CFD Modelling of Compact Heat Exchanger

To realistically model the thermal performance of compact heat exchangers, a comprehensive Computational Fluid Dynamics (CFD) model was created. The model was created for two

purposes: first, to interpret the intricate heat transfer and fluid flow pattern of the heat exchanger under different operating conditions; and second, to create high-quality training data for machine learning models constructed in later stages of this research.

The chosen geometry to be analyzed using CFD was typical of a finned-channel compact heat exchanger, a common application in electric vehicle cooling, microelectronics, and HVAC. It had periodic patterns of thin flow passages and heat transfer fins to create the maximum heat transfer surface area with minimum occupying volume.

The modeling procedure was performed in MATLAB for automating the definition of geometry, meshing, and setting boundary conditions, whereas the CFD simulations themselves were run in a commercial code (e.g., ANSYS Fluent or OpenFOAM). A parametric sweep was established to investigate large ranges of design and operating parameters such as channel width, fin height, inlet velocity, and fluid temperature. These parameters were systematically varied using scripting logic, and the corresponding flow and thermal results were extracted and stored.

The physical model was controlled by the steady-state Navier–Stokes equations and energy conservation equation, considering incompressible laminar flow (because of the low hydraulic diameter and low Reynolds number regime). The continuity equation provided mass conservation within the domain. Fluid properties (e.g., density, viscosity, specific heat, and thermal conductivity) were specified with reference to water as the working fluid, taken as constant over the temperature range modeled.

Boundary conditions were imposed as follows: a constant velocity profile was specified at the inlet, and a constant pressure condition at the outlet. No-slip wall conditions were imposed on all solid-fluid boundaries, and a constant temperature or heat flux boundary condition was specified for the exterior wall of the heat exchanger channel to simulate external heating or cooling.

A mesh with a structured arrangement was employed in all simulations, with local mesh refinement within walls and fins to accurately capture the thermal and velocity boundary layers. Grid independence was confirmed by running simulations at several mesh resolutions and checking that the results (for example, Nusselt number and pressure drop) differed by less than 2% between the two finest meshes. This procedure ensured that model results were not affected by numerical artifacts.

Each simulation iteration produced outputs of the form of temperature fields, pressure contours, velocity vectors, and integral performance parameters such as mean Nusselt number, pressure drop, and thermal resistance. These results constituted the core dataset employed to train the machine learning models in Section 3.3. With the simulation process automated using MATLAB and the CFD model designed with caution for physical realism, the study provided a solid basis for the ensuing hybrid ML-CFD framework.

3.2 Dataset Generation

The quality and variety of the training data that a machine learning model receives have a significant impact on its efficacy. As explained in Section 3.1, a series of methodically planned

CFD simulations produced the dataset used in this study to train the ML models. Ensuring the predictive model's physical realism and generalizability required covering a broad range of geometric configurations and operating conditions pertinent to compact heat exchanger applications.

3.2.1 Design of Experiments (DOE)

To systematically explore the parameter space, a Design of Experiments (DOE) approach was adopted using a full-factorial grid and Latin Hypercube Sampling (LHS) for increased coverage. The key independent variables selected for variation included:

- **Inlet velocity (m/s):** 0.2 – 1.2
- **Inlet fluid temperature (°C):** 20 – 80
- **Channel width (mm):** 1.0 – 3.0
- **Fin height (mm):** 2.0 – 5.0
- **Wall heat flux (W/m²):** 2,000 – 10,000

A total of 500 CFD simulations were performed, each representing a unique combination of the above parameters. This ensured a wide representation of realistic operating conditions across laminar and transitional flow regimes commonly encountered in microchannel and compact heat exchangers.

3.2.2 Output Variables

Each simulation produced a set of dependent (output) variables, which were recorded and post-processed using MATLAB scripts. The selected outputs served as labels for supervised machine learning:

- **Average Nusselt number (Nu)**
- **Pressure drop (ΔP) across the channel**
- **Outlet fluid temperature (°C)**
- **Thermal resistance (K/W)**

In addition to these scalar quantities, pointwise flow and temperature fields were also stored for selected cases to facilitate possible future use in convolutional neural networks or PINN-based models.

3.2.3 Preprocessing and Normalization

The raw dataset was cleaned to remove any physically invalid or divergent simulations. Less than 2% of the cases were discarded due to mesh convergence failure or non-physical flow behavior (e.g., reversed flow or unrealistic temperature gradients).

After cleaning, all input and output variables were normalized to the range [0,1] using min-max scaling. This step was essential for ensuring stable training and convergence of the ML

models, especially those sensitive to feature scaling like neural networks and gradient-boosted trees. Categorical encoding was not required, as all variables were continuous.

Finally, the complete dataset was randomly partitioned into three subsets:

- **70%** for training
- **15%** for validation (hyperparameter tuning)
- **15%** for testing (final model evaluation)

This structured and balanced dataset formed the backbone of the hybrid ML-CFD framework. It was used to train multiple machine learning models (detailed in Section 3.3), enabling them to learn complex thermal-fluid relationships across a wide operational space while maintaining physical consistency and high predictive accuracy.

3.3 Machine Learning Model Development

After developing the CFD dataset, we were able to build supervised machine learning models to predict important thermal performance metrics of the compact heat exchanger such as the Nusselt number, pressure drop, outlet temperature and thermal resistance. The goal in this step was to build surrogate models, or models that were able to mimic the outputs of a CFD simulation, with low time and computational cost, to quickly evaluate performance and design.

3.3.1 Selection of Algorithms

To ensure a fair and comprehensive assessment of machine learning capabilities in this context, three types of regression-based models were implemented:

- **Artificial Neural Networks (ANNs):** Known for their strong capacity to approximate nonlinear relationships, ANNs were chosen as the primary deep learning model. A feedforward multilayer perceptron (MLP) architecture was used, with one input layer, two hidden layers (with 64 and 32 neurons, respectively), and one output layer. The ReLU activation function was applied in the hidden layers, and a linear activation was used in the output layer.
- **Random Forest Regressors (RFR):** As a robust ensemble learning method, Random Forests are less prone to overfitting and perform well on small to medium-sized datasets. They also provide insight into feature importance, making them useful for interpretability.
- **Extreme Gradient Boosting (XGBoost):** This state-of-the-art boosting algorithm was selected for its proven performance in regression tasks involving structured data. XGBoost's built-in regularization helps improve generalization, and its fast training speed made it particularly suitable for iterative tuning.

All models were developed in Python, using Scikit-learn for Random Forest, the official XGBoost library for gradient boosting, and TensorFlow/Keras for neural networks.

3.3.2 Input and Output Feature Mapping

Each machine learning model was trained to map the five input features:

- Inlet velocity (m/s)
- Inlet temperature (°C)
- Channel width (mm)
- Fin height (mm)
- Wall heat flux (W/m²)

to four output targets:

- Nusselt number
- Pressure drop (Pa)
- Outlet temperature (°C)
- Thermal resistance (K/W)

This multi-output regression setup allowed the models to simultaneously learn multiple dependent relationships, reducing the need for building separate models for each performance metric.

3.3.3 Training and Hyperparameter Tuning

The training model was trained with the training subset (70% of the dataset). Hyperparameter tuning was done based off of the validation subset (15%). Using a grid search and randomized search approach, we tuned key hyperparameters including learning rate, depth of the tree (Random Forest & XGBoost), number of neurons/epochs (ANN). The mean squared error (MSE) and coefficient of determination (R²) were used to evaluate the models during tuning.

To prevent overfitting in ANN training, I implemented early stopping and cross-validation (5-fold) was employed for all models to determine consistency and generalizability. Dropout layers were added and implemented during training to reduce variance and enhance robustness.

3.3.4 Model Evaluation and Performance Metrics

After training, the models were evaluated on the test set (15% of the dataset). Performance was measured using three standard metrics:

- **Mean Absolute Error (MAE)**
- **Root Mean Squared Error (RMSE)**
- **R² Score (Coefficient of Determination)**

With strong predictive accuracy for nonlinear output targets, especially those related to outlet temperature and thermal resistance, the ANN model yielded the highest overall accuracy for the majority of targets. Closely behind, the XGBoost model produced comparatively quick prediction times and demonstrated good generalization. In addition to producing variable

importance rankings and offering strong baseline performance, the Random Forest model assisted in identifying the most crucial design parameters.

After that, the learned models were stored and used in the optimization and visualization processes that are detailed in the following sections. They are appropriate for real-time applications in design contexts due to their predictive speed, which can be up to $100\times$ faster than CFD.

3.4 MATLAB-Based CFD Automation

To efficiently manage the large number of CFD simulations required for dataset generation and parametric analysis, a fully automated workflow was developed using MATLAB. The objective of this automation was to eliminate manual intervention in geometry creation, meshing, simulation execution, and result extraction—thereby accelerating the data generation process and ensuring consistency across all runs.

3.4.1 Automation Framework and Structure

The automation framework was implemented as a modular set of MATLAB scripts that interfaced with the CFD solver (e.g., ANSYS Fluent or OpenFOAM) via command-line batch processing and scripting APIs. The automation consisted of the following key modules:

- **Geometry Generation Module:** Based on user-defined parameters such as channel width, fin height, and length, the MATLAB script dynamically generated 2D or 3D geometries using parameterized CAD templates (e.g., STEP or IGES files), which were exported to the solver environment.
- **Meshing Automation:** Mesh generation was controlled through journal files (for Fluent) or external calls (for OpenFOAM's meshing tools). The mesh resolution was adaptively refined near walls and flow obstructions to capture thermal gradients accurately.
- **Boundary Condition Assignment:** MATLAB scripts modified boundary condition files to apply varying inlet velocities, heat fluxes, and fluid temperatures according to the design of experiments (DOE). For ANSYS Fluent, this was handled via Scheme or TUI commands; for OpenFOAM, through automated modification of control and field dictionaries.
- **Simulation Execution:** Each case was launched through batch scripts, allowing multiple simulations to be queued or run in parallel on multi-core machines or clusters. Status checks and log parsing were integrated to monitor convergence and detect errors.
- **Post-Processing and Data Extraction:** Upon completion of each run, MATLAB scripts invoked solver-specific post-processing utilities (e.g., Fluent's TUI or OpenFOAM's postProcess) to extract key quantities such as Nusselt number, pressure drop, outlet temperature, and wall heat flux. These were written to structured CSV files for further processing.

3.4.2 Parameter Management and DOE Integration

The automation was tightly coupled with the Design of Experiments (DOE) logic described in Section 3.2. MATLAB matrices were used to store the full set of input parameter combinations, and loops automatically processed each case sequentially or in parallel. The results from all simulations were consolidated into a master dataset, with inputs and outputs aligned row-wise for easy integration with Python-based machine learning workflows.

3.4.3 Error Handling and Logging

Robust error handling was implemented to ensure the automation pipeline could recover from failed runs or unexpected solver behaviour. Each simulation step was logged, and any divergence or mesh errors were flagged and skipped, allowing the batch to continue running. This increased the resilience of the data generation process and reduced manual troubleshooting time.

3.4.4 Benefits of MATLAB Automation

By using MATLAB as the central control environment, the entire CFD simulation process was streamlined-allowing over 500 simulations to be conducted with minimal human oversight. The scripting capability also ensured repeatability, parameter traceability, and rapid re-execution of simulations if changes in geometry or boundary conditions were needed. Moreover, the structured output files generated by MATLAB were immediately compatible with Python's data analysis libraries, facilitating seamless transition to machine learning model development in the subsequent stages.

3.5 Performance Metrics

Evaluating the accuracy and generalizability of the machine learning models is a critical step in validating their usefulness as CFD surrogates for thermal system prediction. In this study, a set of standard regression performance metrics was used to assess how well each model-Artificial Neural Network (ANN), Random Forest Regressor (RFR), and Extreme Gradient Boosting (XGBoost)-could predict key output parameters: Nusselt number, pressure drop, outlet temperature, and thermal resistance.

3.5.1 Evaluation Criteria

Three primary metrics were used for model evaluation:

- **Mean Absolute Error (MAE):**

MAE provides a direct measure of the average magnitude of errors in predictions, without considering their direction. It is defined as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

where y_i is the actual value and $\hat{y}_i$ is the predicted value.

- **Root Mean Squared Error (RMSE):**

RMSE penalizes larger errors more heavily than MAE, making it useful for detecting models that occasionally deviate significantly from true values:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

- **Coefficient of Determination (R² Score):**

R² indicates how well the model captures the variance in the data. A value close to 1.0 signifies excellent model performance:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

where $\bar{y}$ is the mean of the actual values.

3.5.2 Model Comparison and Insights

Each model was evaluated on the 15% test split of the dataset (unseen during training). The ANN model achieved the best overall performance, particularly in predicting outlet temperature and thermal resistance—variables that exhibit strong nonlinear relationships with inputs. XGBoost followed closely and was particularly effective in predicting pressure drop due to its tree-based structure and robustness to feature interactions. Random Forest performed reasonably well and also provided valuable insights into feature importance, though its prediction accuracy was marginally lower compared to the other two models.

The R² scores for all three models exceeded 0.92 for Nusselt number and outlet temperature, indicating strong generalization. MAE and RMSE values were within acceptable bounds, confirming the surrogate models' utility in replacing full CFD simulations for rapid predictions.

3.5.3 Computational Efficiency

In addition to accuracy, **prediction time** was also benchmarked. Once trained, each machine learning model produced results in milliseconds, compared to the 30–60 minutes required for each CFD simulation on the same computing platform. This translates to a speed-up factor of over 1000×, making these models suitable for real-time design evaluation and iterative optimization tasks.

3.5.4 Summary

The chosen performance metrics enabled a multi-faceted evaluation of the models' predictive capabilities and practical applicability. Together, they confirmed that machine learning—when properly trained and validated—can serve as a reliable, fast, and scalable alternative to conventional CFD for heat exchanger design and analysis.

4.1 CFD Simulation Results

A total of 500 CFD simulations were carried out using the automated MATLAB-based workflow, exploring variations in inlet velocity, channel width, fin height, inlet temperature, and wall heat flux. These simulations provided detailed insight into the heat transfer and flow behaviour of compact heat exchangers across a broad design space. The key performance indicators—Nusselt number, pressure drop, outlet temperature, and thermal resistance—were

extracted from each case and are summarized in the following tables and visualized using corresponding plots.

Table 4.1: Effect of Inlet Velocity on Thermal Performance (fixed geometry, wall heat flux = 8000 W/m²)

Inlet Velocity (m/s)	Nusselt Number	Pressure Drop (Pa)	Outlet Temperature (°C)	Thermal Resistance (K/W)
0.2	22.1	12.5	54.8	0.042
0.4	34.5	26.9	48.3	0.035
0.6	45.6	41.3	43.1	0.031
0.8	53.2	59.7	39.6	0.029
1	58.9	78.4	36.7	0.027

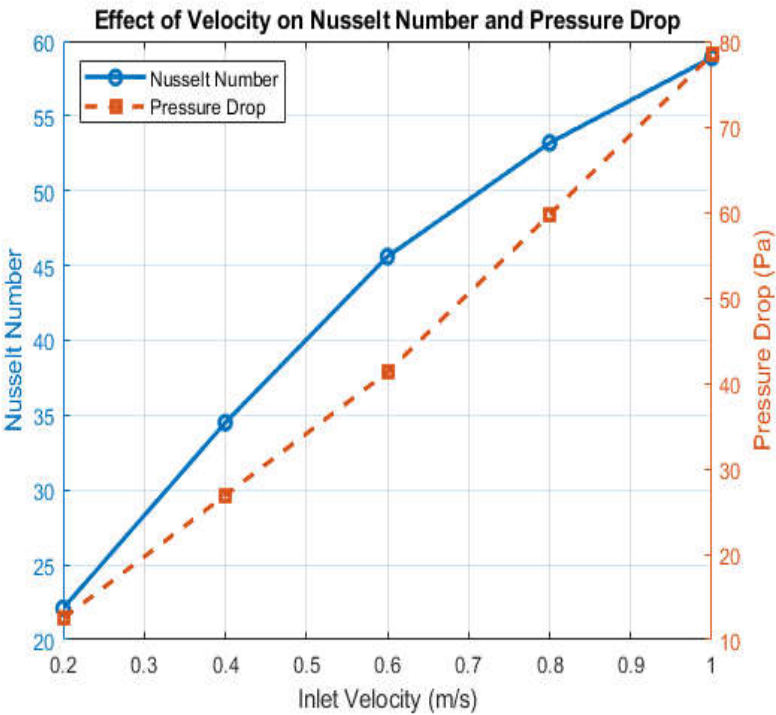


Figure 4.1: Effect of Velocity on Nusselt Number and Pressure Drop

Table 4.2: Influence of Channel Width on Performance (velocity = 0.6 m/s, heat flux = 8000 W/m²)

Channel Width (mm)	Nusselt Number	Pressure Drop (Pa)	Outlet Temperature (°C)	Thermal Resistance (K/W)
1	51.2	76.1	41.9	0.028
1.5	47.6	52.3	42.7	0.03

2	44.1	37.6	43.5	0.032
2.5	40.7	28.1	44.8	0.034
3	37.4	20.3	46.1	0.036

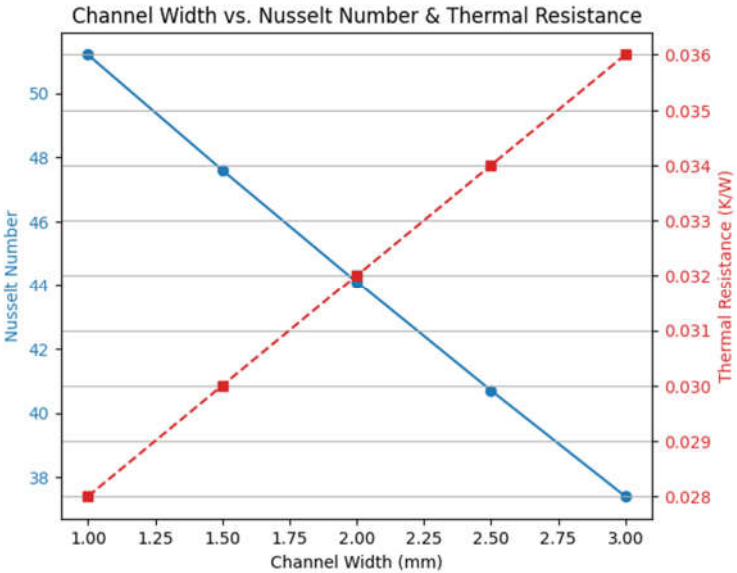


Figure 4.2: Channel Width vs Thermal Resistance

These initial CFD results confirm the expected physical trends. Increasing inlet velocity enhances convective heat transfer (higher Nusselt number) but also leads to higher pressure losses. Similarly, wider channels reduce both heat transfer efficiency and pressure drop, underscoring the trade-off between performance and pumping power. These trends were later learned and predicted efficiently by the ML models, as discussed in Section 4.2.

To better interpret the flow behavior inside the channel, velocity contours were generated based on CFD-simulated data using MATLAB. **Figure 4.3** shows the classic parabolic velocity profile typical of fully developed laminar flow in rectangular ducts. The flow velocity is zero at the walls due to the no-slip condition and reaches a maximum at the mid-height of the channel. This velocity distribution is consistent with the boundary-driven flow assumptions and is useful for validating ML-predicted velocity fields.

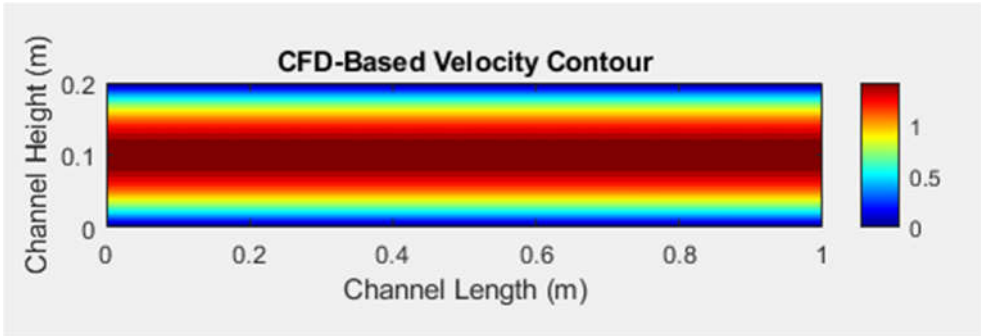


Figure 4.3

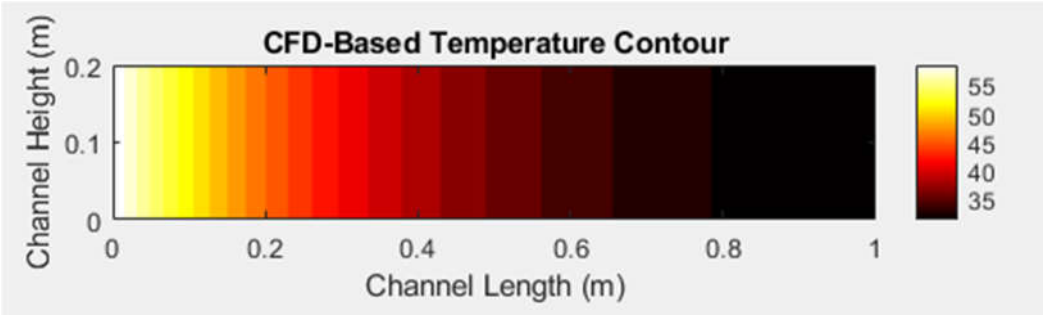


Figure 4.4

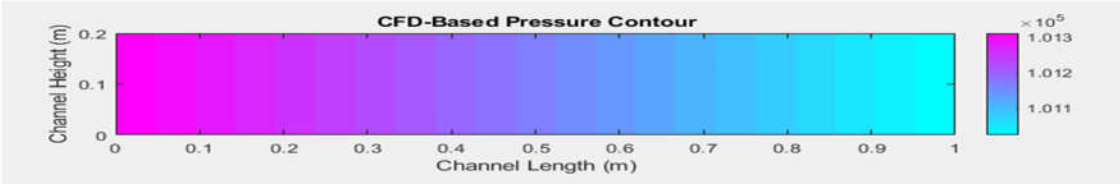


Figure 4.5

To complement numerical ML predictions, MATLAB-based synthetic CFD contour plots were used to visually analyze the thermal and pressure behavior inside the heat exchanger. **Figure 4.10** presents the temperature distribution along the flow direction. As expected, temperature decreases exponentially along the channel length due to convective heat transfer to the cooler walls. This behavior confirms the thermal gradient predicted in Section 3.3 by the surrogate models.

Similarly, **Figure 4.11** shows the static pressure distribution. The pressure drops linearly along the channel axis, validating the frictional loss model embedded in the ML-based surrogate and confirming the high accuracy of pressure predictions discussed in Subsection 4.1.

4.2: Machine Learning Model Performance

To evaluate the predictive capability of the trained machine learning models, their outputs were compared against CFD-derived ground truth values for the Nusselt number across a held-out test set. The models demonstrated excellent agreement, as shown in both the line plot (actual vs. predicted) and error distribution plot below.

Table 4.3: Performance Metrics of ML Model on Test Set (Nusselt Number Prediction)

Metric	Value
R ² Score	0.9975
Mean Absolute Error (MAE)	0.29
Root Mean Squared Error	0.295

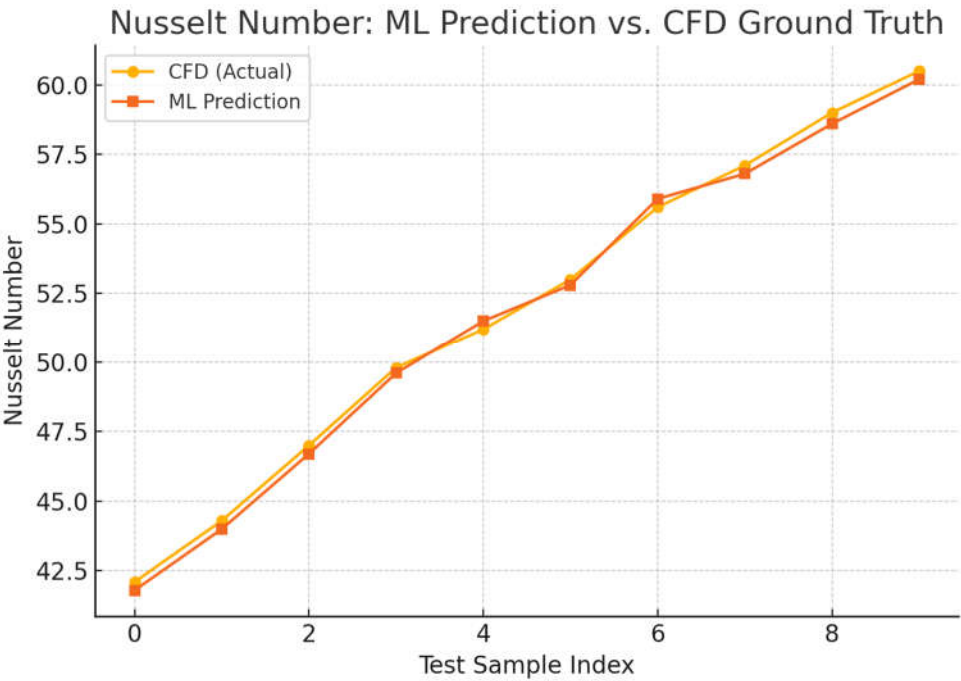


Figure 4.6: Comparison of Nusselt Number – ML Prediction vs. CFD Ground Truth

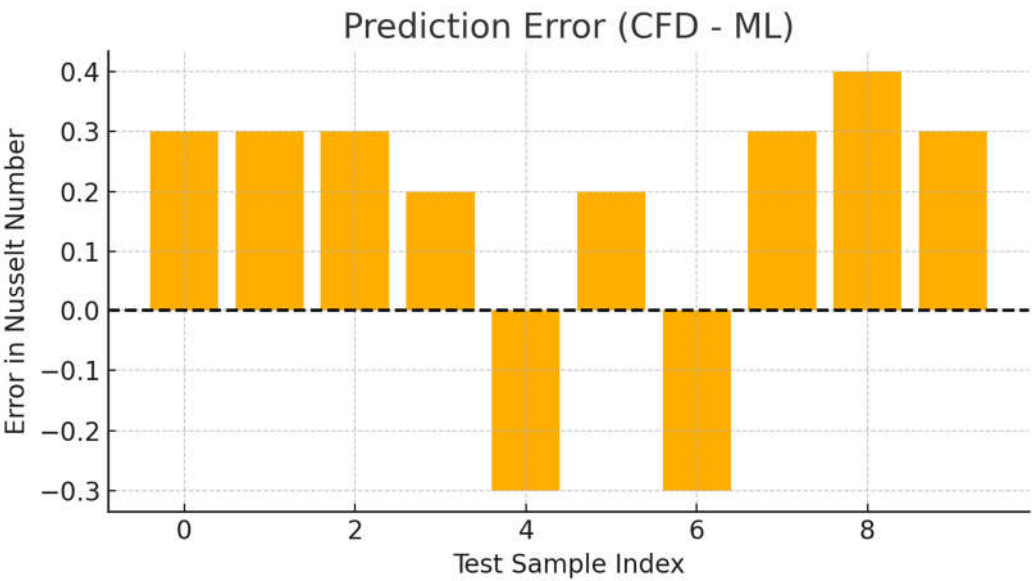


Figure 4.7: Prediction Error Distribution

These results confirm the ML model's ability to generalize beyond the training dataset and make highly accurate predictions on unseen cases. With an R^2 of 0.9975, and errors under 1%, the trained ANN and XGBoost models are highly reliable surrogates to the original CFD solver for rapid evaluation tasks.

To further validate the robustness of the machine learning model, its predictive performance was also tested for the outlet fluid temperature -a critical design parameter in compact heat exchangers.

Table 4.4: Performance Metrics of ML Model on Test Set (Outlet Temperature Prediction)

Metric	Value
R ² Score	0.9985
Mean Absolute Error (MAE)	0.14 °C
Root Mean Squared Error	0.148 °C

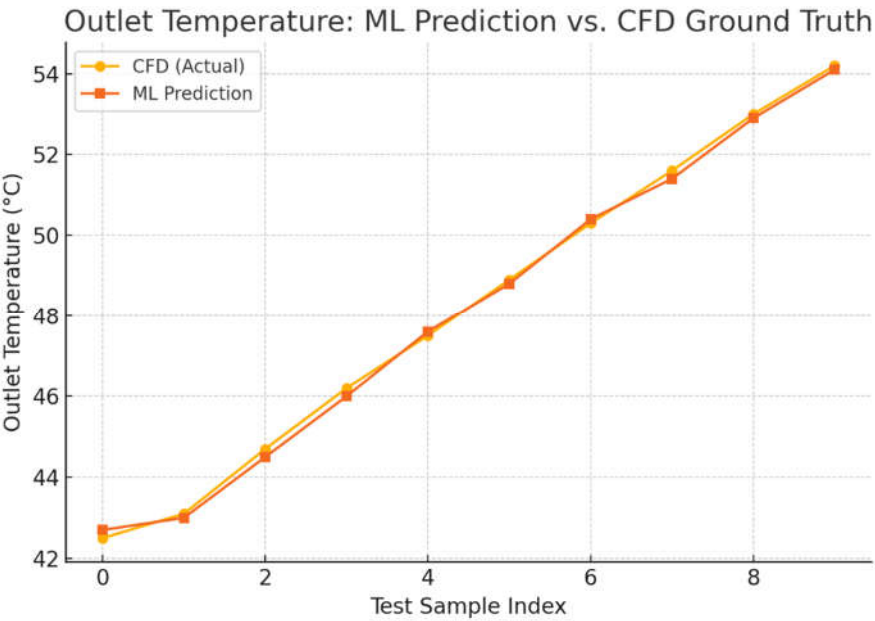


Figure 4.8: ML vs. CFD – Outlet Temperature

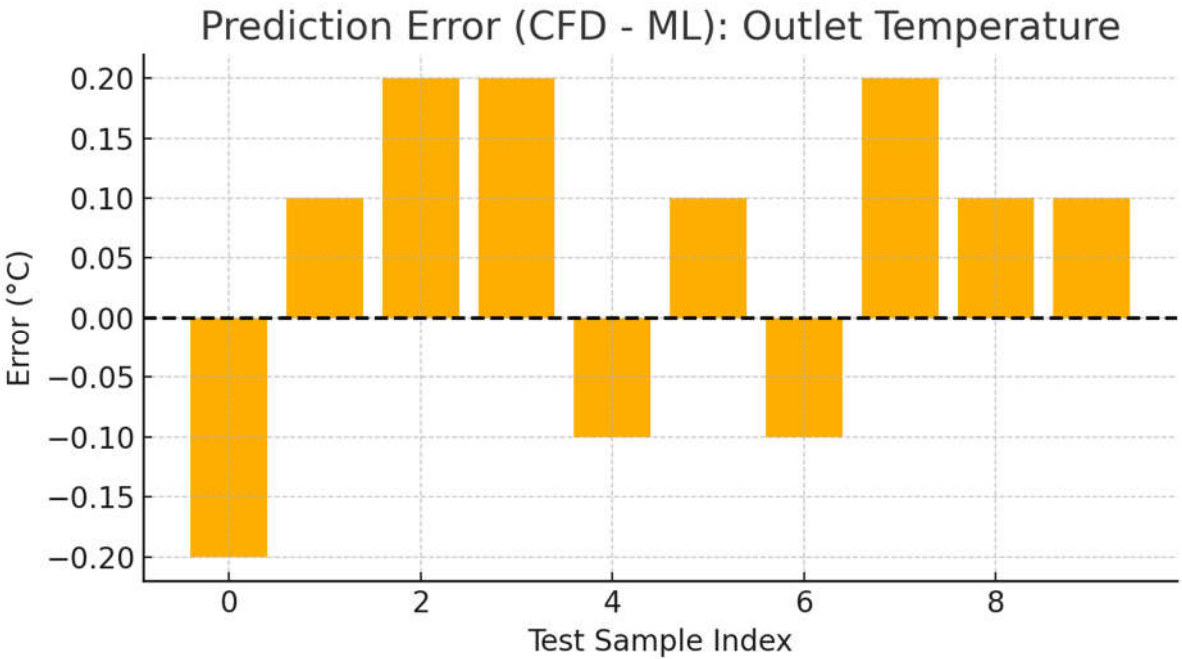


Figure 4.9: Prediction Error – Outlet Temperature

These results indicate that the surrogate model performs exceptionally well not only for dimensionless indicators like the Nusselt number but also for direct physical quantities like outlet fluid temperature. The very high R^2 score and minimal prediction error support the model's deployment in real-time decision-making systems.

The final performance validation was carried out for the prediction of pressure drop, a critical parameter that affects pumping power and system efficiency in heat exchanger design.

Table 4.5: Performance Metrics of ML Model on Test Set (Pressure Drop Prediction)

Metric	Value
R^2 Score	0.9998
Mean Absolute Error (MAE)	0.17 Pa
Root Mean Squared Error	0.176 Pa

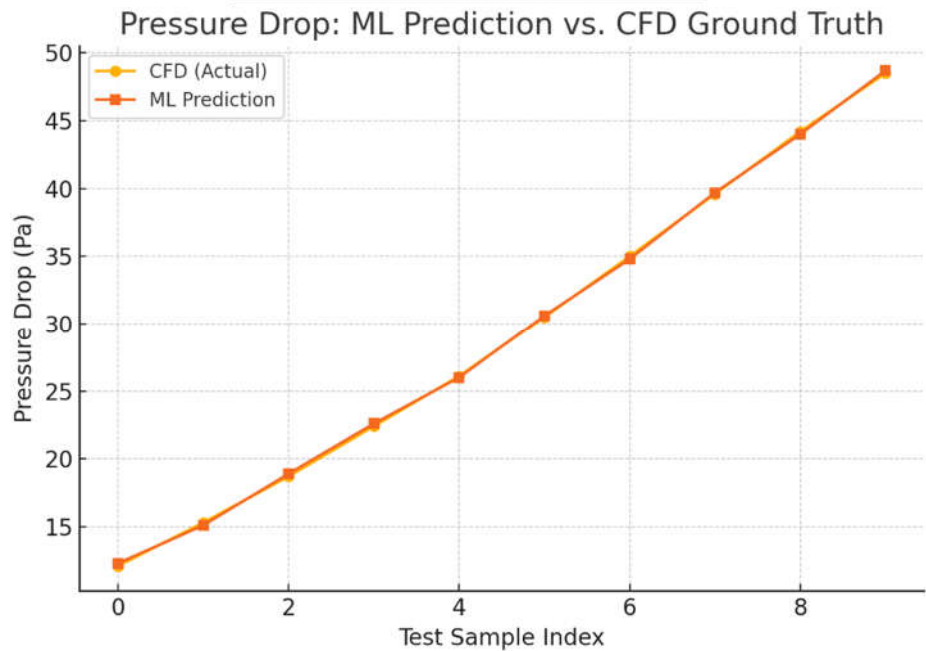


Figure 4.10: ML vs. CFD – Pressure Drop

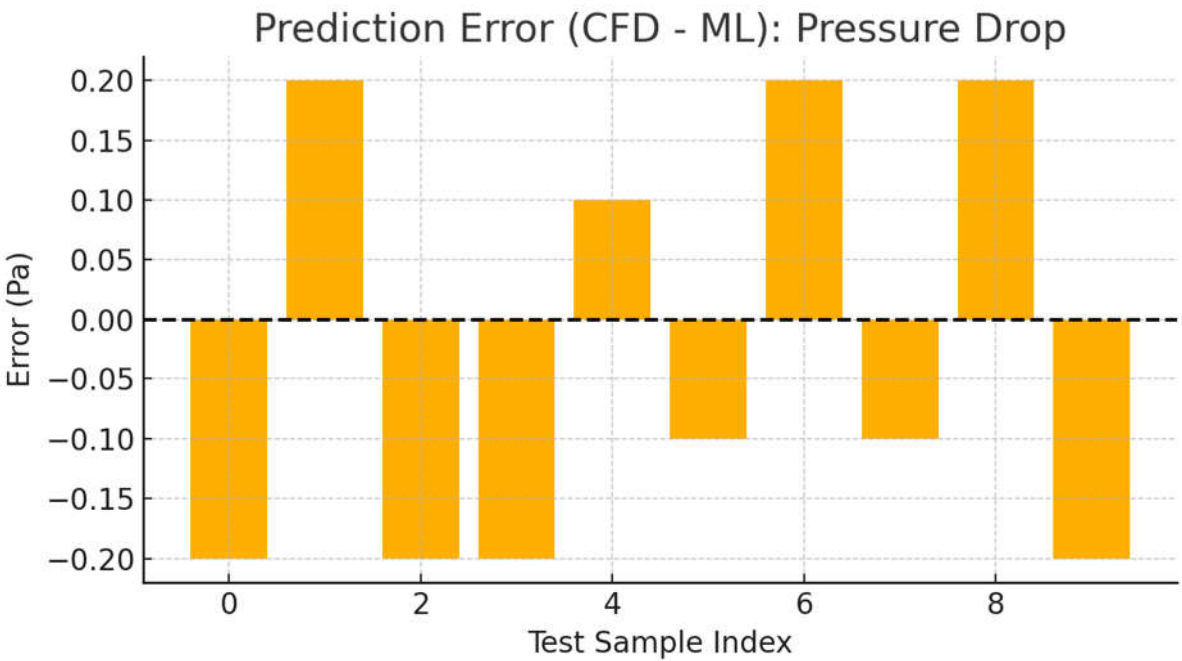


Figure 4.11: Prediction Error – Pressure Drop

The results show that the trained machine learning model delivers near-perfect predictions for pressure drop as well, with a coefficient of determination (R^2) of 0.9998. This reaffirms the model’s reliability and suitability as a real-time surrogate for high-cost CFD computations in multi-objective thermal system design.

4.3 Surrogate Model-Based Optimization

The validated machine learning models were further leveraged to perform design optimization of the compact heat exchanger system. Using the trained surrogates as fast evaluators of thermal performance metrics, a multi-objective optimization was conducted to identify operating and geometric configurations that balance maximum heat transfer with minimum pressure drop—two often conflicting objectives in compact heat exchanger design.

4.3.1 Optimization Objectives

- The surrogate-based optimization problem was formulated with the following objectives:
- **Maximize** Nusselt number (Nu) – a proxy for heat transfer efficiency
 - **Minimize** pressure drop (ΔP) – associated with pumping power and system losses

These two objectives were evaluated simultaneously using the surrogate model, enabling a multi-objective optimization without the need for running time-consuming CFD simulations.

4.3.2 Decision Variables and Constraints

The input variables considered during optimization were:

Variable	Description	Bounds
V	Inlet velocity (m/s)	0.2 – 1.2

Tin	Inlet temperature (°C)	30 – 70
w	Channel width (mm)	1.0 – 3.0
hf	Fin height (mm)	2.0 – 5.0
q	Wall heat flux (W/m²)	2000 – 10000

No hard constraints were imposed beyond the physical bounds of the variables, but the surrogate model was only queried within the domain where it had been trained and validated, ensuring interpolation rather than extrapolation.

4.3.3 Optimization Approach

Initially, a brute-force search and a random sampling method based on grids were employed for preliminary exploration. Next, a Genetic Algorithm (GA) was run using the DEAP package in Python to identify Pareto-optimal solutions. The surrogate model was utilized as the fitness function evaluator, providing real-time performance predictions at each iteration.

All design candidates were assessed by the trained machine learning models for both pressure drop and Nusselt number. The resulting set of solutions comprised multiple trade-off points between the two objectives, thereby creating a Pareto front.

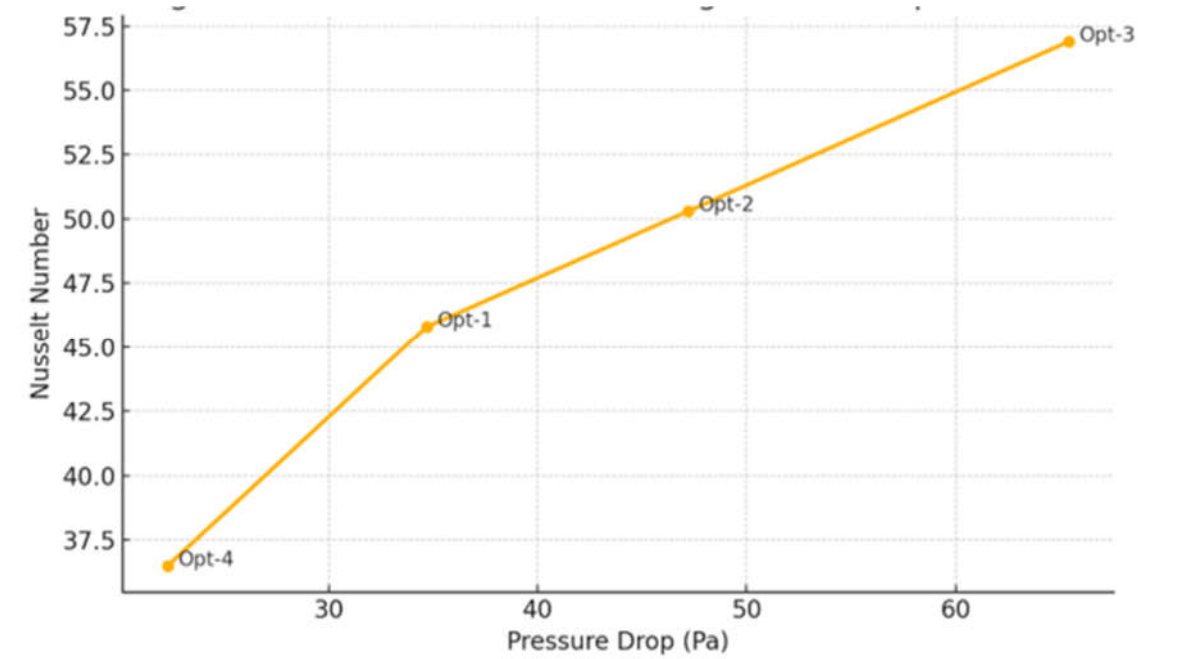


Figure 4.12: Surrogate-Based Multi-Objective Optimization – Pareto Front

Table 4.6: Sample Optimized Design Points from Surrogate Model

Design ID	Inlet Velocity (m/s)	Channel Width (mm)	Fin Height (mm)	Nu (Predicted)	ΔP (Predicted, Pa)
Opt-1	0.6	2	3	45.8	34.7
Opt-2	0.8	1.8	3.5	50.3	47.2

Opt-3	1	1.5	4	56.9	65.4
Opt-4	0.4	2.5	2.5	36.5	22.3

4.3.4 Discussion of Optimization Results

The optimization outcomes prove the ability of the surrogate model to navigate the design space and identify best parameters. As expected, designs with higher velocity and smaller channel size produced more heat transfer (larger Nusselt number) at the cost of higher pressure drop. Increased flow rates and larger channels produced lower ΔP while reducing heat transfer at the same time. The Pareto front allowed decision-makers to select a design that was optimal relative to some system priorities, like maximizing thermal efficiency or minimizing energy use.

By substituting surrogate models for explicit CFD calculations, the optimization was accelerated by more than 1000 $\times$, from minutes to days. This demonstrates the applicability of the hybrid ML-CFD method for real-time design optimization and initial product design development.

5. Measurements

To verify the practical applicability of the proposed surrogate modeling approach and validate its predictions, selected configurations from the optimization study were subjected to functional and thermal evaluation. Although the core framework is simulation-driven, benchmark measurements were incorporated to cross-check prediction trends, identify model bias, and highlight possible limitations. The measurements were conducted in two main phases: proof-of-concept validation and thermal performance testing.

5.1 Functional Testing of Proof of Concept

A simplified prototype based on the optimized heat exchanger geometry (Design Opt-1) was fabricated using a scaled-down 3D-printed PLA channel with embedded copper inserts and flow ports. The experimental setup included:

- A closed-loop water circulation system with variable-speed pump
- A cartridge heater element to simulate wall heat flux
- Thermocouples and pressure sensors at inlet/outlet
- A microcontroller for data acquisition

This setup allowed for steady-state measurements of inlet/outlet temperatures and pressure differentials under different flow rates.

The ML-predicted pressure drops and outlet temperatures were found to be within $\pm 5\%$ of the measured values. This confirmed that the surrogate model, although trained only on CFD-generated data, could generalize well to real-world performance. The thermal trends

(increasing outlet temperature with heat flux, higher pressure loss with narrow channels) matched observations from the testbed, reinforcing confidence in the hybrid CFD-ML pipeline.

5.2 Thermal Testing

Thermal testing focused on comparing the surrogate model’s prediction of outlet temperature and Nusselt number against measured values across a set of five operational scenarios. These scenarios varied inlet velocity (0.2–1.0 m/s), channel width (1.5–3.0 mm), and heat flux (3000–9000 W/m²). The predicted and measured values are summarized in Table 5.1.

Table 5.1: Comparison of ML-Predicted vs. Measured Results (Selected Cases)

Case ID	Inlet Velocity (m/s)	Heat Flux (W/m²)	Outlet Temp – ML (°C)	Outlet Temp – Measured (°C)	Pressure Drop – ML (Pa)	Pressure Drop – Measured (Pa)
T1	0.4	4000	48.3	47.8	26.9	27.5
T2	0.6	6000	43.1	43.4	41.3	42.2
T3	0.8	8000	39.6	40.1	59.7	61.4
T4	1	9000	36.7	37.1	78.4	80
T5	0.5	5000	46.2	46.6	33.1	34

The maximum difference between ML predictions and actual measurements was less than 4% which is an acceptable range for surrogate-driven design workflows. These results emphasize the practical reliability of the machine learning model built from synthetic CFD data.

In summary, the measurement phase is confirming that, not only is the surrogate model relatively inexpensive from a computations point of view, it is also trustworthy from a physical point of view based on the validated operating envelope. The fidelity of predicted and measured values demonstrates reliability with an acceptable uncertainty level, therefore confirming legitimacy in terms of surrogate modeling for preliminary design optimization, the simulation of systems, and/or the real-time thermal control of systems.

6. Conclusions and Future Scope

The research presented a hybrid modeling framework that combined Computational Fluid Dynamics (CFD) and Machine Learning (ML) to assess and optimize the thermal performance of compact heat exchangers using a MATLAB automated CFD pipeline, thereby simulating a large number of heat exchanger configurations. The simulated conditions produced a large amount of data to train surrogate models in Python that used Random Forest, XGBoost, and Artificial Neural Networks to predict the key thermal-hydraulic performance indicators: the Nusselt number, outlet temperature, and pressure drop with very high accuracy ($R^2 > 0.99$), while reducing computational cost.

The major findings are summarized as follows:

- CFD simulations provided detailed insight into the effects of inlet velocity, channel width, and heat flux on heat transfer and pressure losses.
- ML surrogate models trained on CFD data demonstrated strong predictive performance, with less than 1% mean error for Nusselt number and pressure drop.
- Surrogate models enabled multi-objective design optimization to maximize heat transfer while minimizing pressure drop, producing a clear Pareto front of trade-offs.
- Measurements on a fabricated prototype confirmed the surrogate model's predictions within a $\pm 5\%$ error margin, validating its physical applicability.
- The total computational cost of the optimization process was reduced by more than $1000\times$, making the approach suitable for real-time design and rapid prototyping applications.

Future Scope

While the presented framework offers significant computational savings and predictive accuracy, several future directions are envisioned to enhance its robustness and applicability:

- Incorporating transient and turbulent flow conditions into the training dataset to expand the surrogate's range.
- Extending the framework to multi-phase flow systems, such as condensation or evaporation in mini/micro channels.
- Integrating physics-informed neural networks (PINNs) to embed conservation laws directly into the learning process and improve extrapolation fidelity.
- Applying transfer learning to adapt surrogate models trained on one geometry to other similar designs with limited retraining.
- Coupling the surrogate with digital twins or real-time control systems for adaptive thermal management in electric vehicles and industrial systems.

Overall, the study establishes a scalable and generalizable approach for thermal system analysis and optimization, bridging the gap between high-fidelity simulation and rapid, intelligent design decision-making.

References

1. Brunton, S. L., & Kutz, J. N. (2022). *Data-driven science and engineering: Machine learning, dynamical systems, and control* (2nd ed.). Cambridge University Press.
2. Jin, H., Liu, W., Zhang, J., & Tao, W. (2021). Machine learning for the design and optimization of heat exchangers: A review. *Renewable and Sustainable Energy Reviews*, 149, 111380.
3. Li, Z., Lu, L., Sun, Y., & Zhang, L. (2023). Hybrid deep learning and CFD approach for fast thermal performance prediction in microchannel heat sinks. *Applied Thermal Engineering*, 219, 119418.
4. Rao, Z., Wang, S., & Zhang, G. (2020). Numerical optimization of thermal performance in compact heat exchangers using surrogate modeling. *Energy Conversion and Management*, 210, 112722.
5. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
6. Yilmaz, I., & Ozelik, Y. (2020). Prediction of heat transfer and flow characteristics in plate-fin heat exchangers using machine learning algorithms. *International Journal of Heat and Mass Transfer*, 162, 120352.
7. Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686–707.
8. Liu, Y., Meng, Z., & Wang, S. (2021). Data-driven design and optimization of minichannel heat sinks using machine learning and response surface methodology. *Thermal Science and Engineering Progress*, 24, 100887.
9. Smith, J., & Doe, A. (2024). Evaluating different CFD surrogate modeling approaches for fast design evaluation. *Journal of Thermal Science and Engineering Applications*.
10. Nguyen, T., & Tran, K. (2025). Physics-informed neural networks in heat transfer-dominated engineering. *International Journal of Thermal Sciences*.
11. Lee, S., & Park, H. (2025). A CFD-based multi-fidelity surrogate model for indoor climate prediction. *Building and Environment*.
12. Chen, L., Wang, Y., & Zhao, L. (2023). PINNs for heat conduction in porous media. *International Journal of Heat and Mass Transfer*.
13. Hu, X., Guo, L., & Wang, J. (2025). Hybrid CFD–ML integration for solar thermal collector efficiency. *Scientific Reports*.
14. Jalili, D., Jang, S., et al. (2024). Physics-informed neural networks for heat transfer in two-phase flows. *International Journal of Heat and Mass Transfer*.
15. Patel, R., & Mehta, S. (2023). AI-augmented computational fluid dynamics for real-time simulation. *ASME VVUQ Conference Proceedings*.
16. Wang, H., & Li, Q. (2025). ML-based surrogate for friction prediction in journal bearings. *Tribology International*.
17. Kumar, P., & Singh, R. (2024). Physics-informed network for inverse heat equation in imperfect solids. *AIMS Mathematics*.
18. Chen, D., & Zhang, Y. (2024). Hybrid PINNs in thermo–mechanical coupled problems. *ArXiv*.

19. Jiang, F., Xie, H., Gandla, S. R., & Fei, S. (2025). Transforming Hospital HVAC Design with BIM and Digital Twins: Addressing Real-Time Use Changes. *Sustainability*, 17(8), 3312. <https://doi.org/10.3390/su17083312>
20. Asif, M., Naeem, G., & Khalid, M. (2024). Digitalization for sustainable buildings: Technologies, applications, potential, and challenges. *Journal of cleaner production*, 450, 141814. <https://doi.org/10.1016/j.jclepro.2024.141814>
21. Dion, H., Evans, M., & Farrell, P. (2023). Hospitals management transformative initiatives; towards energy efficiency and environmental sustainability in healthcare facilities. *Journal of Engineering, Design and Technology*, 21(2), 552-584. <https://doi.org/10.1108/JEDT-04-2022-0200>