SURROGATE PHYSICS MODELING FOR HEAT EXCHANGER OPTIMIZATION: ACCELERATING DESIGN EXPLORATION VIA INVERSE DESIGN

Amirhossein Iranmehr, Armin Hassanirad
University of Alberta
Mechanical Engineering Department
iranmehr@ualberta.ca, hassanir@ualberta.ca

ABSTRACT

This work presents a surrogate physics modelling framework for heat exchanger (HEX) optimization, developed as part of the ASME IDETC-CIE 2025 Student Hackathon. Conventional CFD-based HEX optimization is computationally expensive, motivating the use of machine learning (ML) for rapid design exploration. Using nTop Automate, 890 data points were generated across design ranges of 10–25 mm (X cell), 10–25 mm (Y/Z cell), and 2500–3500 mm/s (inlet velocity). Multiple ML models - including Random Forest, Gradient Boosting, Polynomial Regression, and Neural Networks - were trained to predict outputs such as pressure drop, average velocity, surface area, and mass. Histogram Gradient Boosting achieved the best performance, with an average prediction error of <8%. The trained surrogate enabled inverse design, identifying optimal parameters (cell X = 19.651 mm, cell Y/Z = 24.561 mm, inlet velocity = 2747.175 mm/s) yielding a surface area of 21516.3 mm² with a pressure drop of 5124.33 Pa. This work highlights the potential of physics-informed ML surrogates to accelerate HEX design with high accuracy and efficiency.

Keywords: nTop, heat exchanger, surrogate modeling, machine learning, inverse design

1. INTRODUCTION

The ASME (IDETC-CIE) 2025 Student Hackathon challenges participants to explore the design and optimization of heat exchangers (HEXs) by developing surrogate physics models, a problem the nTop company poses. Traditional optimization of HEXs requires significant computational resources since the geometry preparation and computational fluid dynamics (CFD) simulations for each iteration require a considerable computational capacity. To overcome this, the competition provided a dataset generated using nTop's parameterized HEX model. This dataset includes average velocity, mass, pressure drop, surface area, velocity inlet, and cell sizes. The task was to train a surrogate physics model that predicts pressure drop, average flow velocity, core surface area, and mass based on key input parameters: lattice cell size in the X and Y/Z directions and inlet flow velocity. The parameters must stay within these limits: 10 mm < cell size X < 25 mm, 10 mm < cell size Y/Z < 25 mm, and 2500 mm/s < inlet velocity < 3500 mm/s. Once trained, the model should be used for inverse design to find the maximum surface area, which corresponds to maximum heat transfer, while meeting limits on mass (below 125 g), pressure drop (below 8000 Pa), and average velocity (above 520 mm/s). This method allows for quicker design iterations and shows the practical benefits of physics-informed machine learning in engineering optimization workflows.

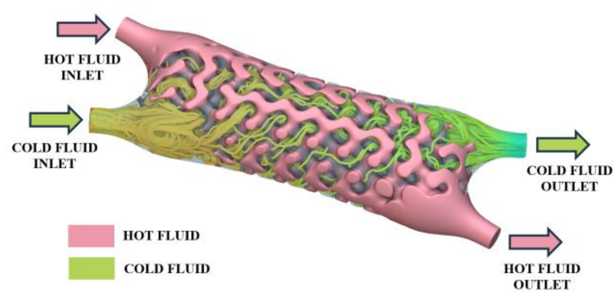


Figure 1. Hot-cold fluids interaction within the HEX

Recent studies indicate that combining new methods like artificial intelligence (AI) and machine learning (ML) with traditional CFD simulations can significantly reduce computational costs while maintaining accuracy.

Jeon et al. (1) noted that the high computational cost of CFD simulations in industry can be lessened using hybrid AI-CFD methods, such as residual-based physics-informed transfer learning (RePIT), which speed up simulations while maintaining accuracy. Machine learning (ML) has also been effectively used to develop heat transfer correlations, allowing quick multi-variable predictions without deep simulations. Kwon et al. (2) used a random forest algorithm to predict convection heat transfer coefficients in cooling channels with varying rib roughness. This shows that machine learning can reduce the effort needed for model development and easily broaden the parameter range. Krishnayatra et al. (3) investigated the thermal performance of a new axial finned-tube heat exchanger using a machine learning regression method. They used the k-Nearest Neighbor (k-NN) algorithm to predict efficiency and effectiveness based on fin spacing, fin thickness, material, and convective heat transfer coefficient, achieving high prediction accuracy. Their study demonstrated that k-NN is a reliable and precise tool for thermal system design, enabling efficient optimization of heat exchanger geometries.

There is still a lack of research focused on developing surrogate physics models that can generalize to a wide range of HEX geometries and operating parameters using only a simulation-based dataset. This study aims to create a physics-informed surrogate modeling framework, combining CFD simulation results with machine learning, to enable rapid and accurate prediction of hydraulic performance and geometry properties of HEX across a wide range of operating conditions and geometries.

2. MATERIALS AND METHODS

2.1 Primary Exploratory Data Analysis (EDA)

An initial exploratory data analysis (EDA) was conducted on the original dataset containing 125 data points. The goal was to investigate the relationship between the input features and target variables. This analysis offered valuable insights into how features behave and possible correlations. Looking at the feature distribution plots (Figure 2) showed that, within the ranges of the three input parameters, the data points appeared in distinct intervals rather than a continuous manner. This gap highlighted the possible need for data augmentation to represent the design space better.

2.2 Data Augmentation

To address the limited size of the available dataset, Latin Hypercube Sampling (LHS) was employed to generate additional, well-distributed design points across the design space. The sampling process incorporated 15 discrete levels for YZ cell size and inlet velocity, and 5 discrete levels for X cell size. These levels were selected based on the observed target parameter correlations, which indicated that YZ cell size and inlet velocity had the most significant influence on the target variables. In contrast, X cell size had a moderate effect (Figure 3).

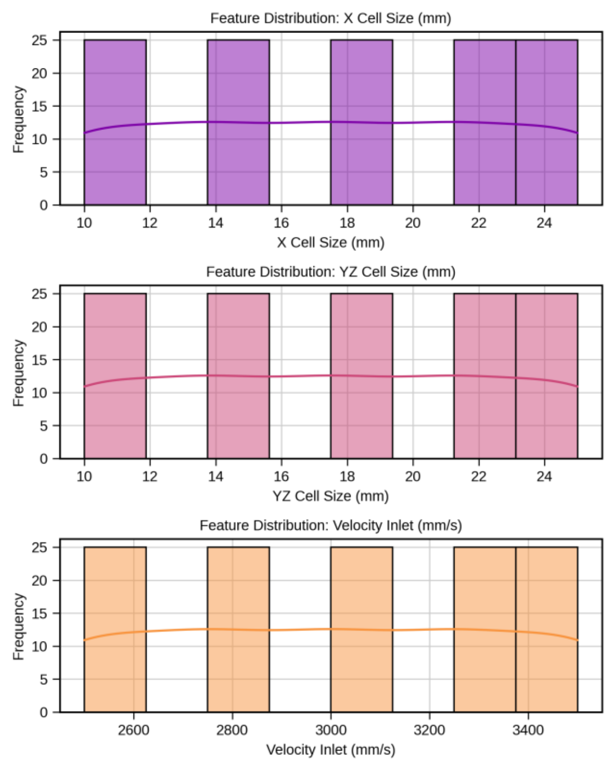


Figure 2. Initial data features' distribution

Utilizing nTop Automate on the HEX model provided by the hackathon organizers, 890 additional data points were generated, ensuring comprehensive coverage across the input ranges, as shown in Figure 4.

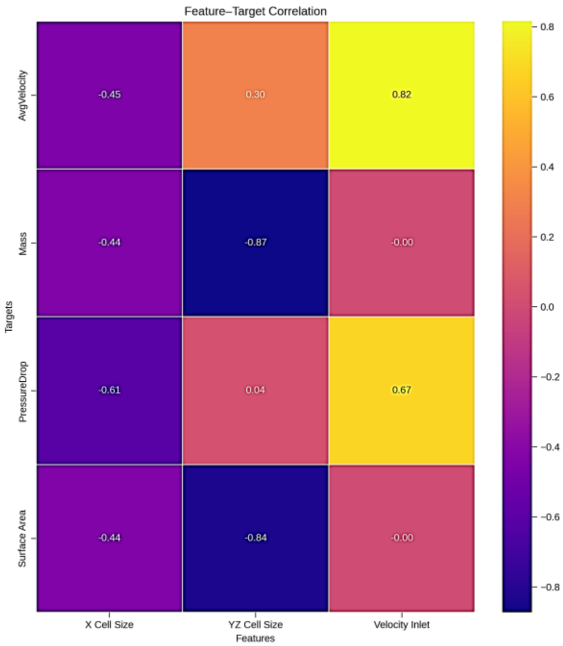


Figure 3. Feature-target correlation

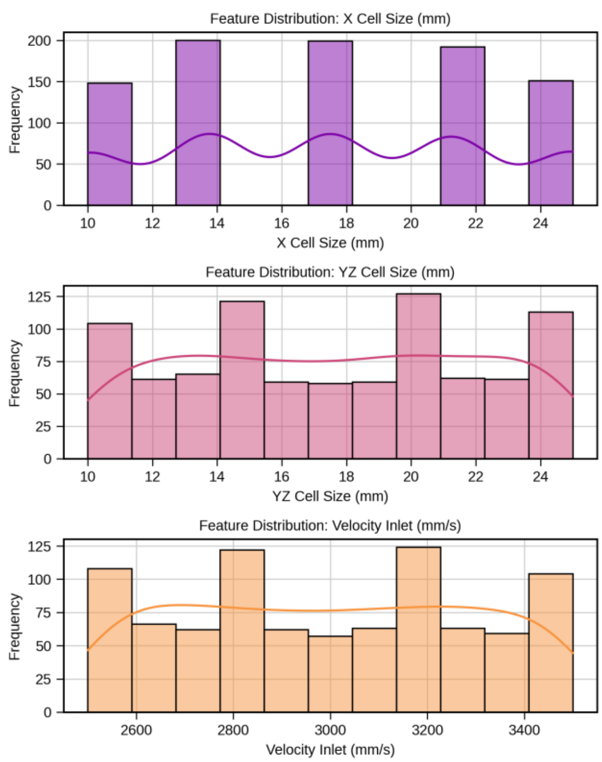


Figure 4. Augmented data features' distribution

2.3 Secondary Exploratory Data Analysis (EDA)

After the data augmentation, secondary EDA is needed to more accurately determine the patterns and relationships between features and targets. This would further help us improve our physics-informed analysis and choose better models for our problem. Figure 5 illustrates the feasible design according to the problem statements, revealing the trade-off between different features to reach the possible design. To better understand the dataset, the cold channel void volume (V_S) of the HEX was extracted from nTop and used to calculate the hydraulic diameter (D_h). Equation 1 was used to calculate the hydraulic diameter, while Equation 2 was applied to determine the Reynolds number (Re), a key parameter for analyzing the hydraulic behavior of the HEX. In this context, V_S represents the cold fluid volume, A_S is the surface area, ρ is the water density, U is the inlet velocity, and μ denotes the dynamic viscosity of water.

Depicting from Figure 6, Reynolds number shows a clear linear trend with inlet velocity, as expected. The scattered distribution of Re at a given inlet velocity corresponds to variations in hydraulic diameter, which are influenced by changes in the X and Y/Z cell sizes. Overall, the range of Re confirms that the cold fluid flow is fully turbulent.

$$\mathrm{D}_h = \frac{4V_S}{A_S} \quad (1)$$

$$Re = \frac{\rho U \mathrm{D}_h}{\mu} \quad (2)$$

2.4 Machine Learning Models Setup, Training, and Evaluation

A systematic pipeline was implemented for the training process. Data was preprocessed for the ML model's training and followed by cross-validation (CV) to obtain reliable performance estimates and avoid overfitting. Different machine learning and deep learning algorithms were modeled, including Random Forest, Gradient Boosting, Polynomial, and Neural Networks (MLP), with each model trained on the expanded dataset. Their predictive accuracy was assessed using RMSE and R^2 score, consistent with this framework. Accordingly, the computational time of each model was investigated further to

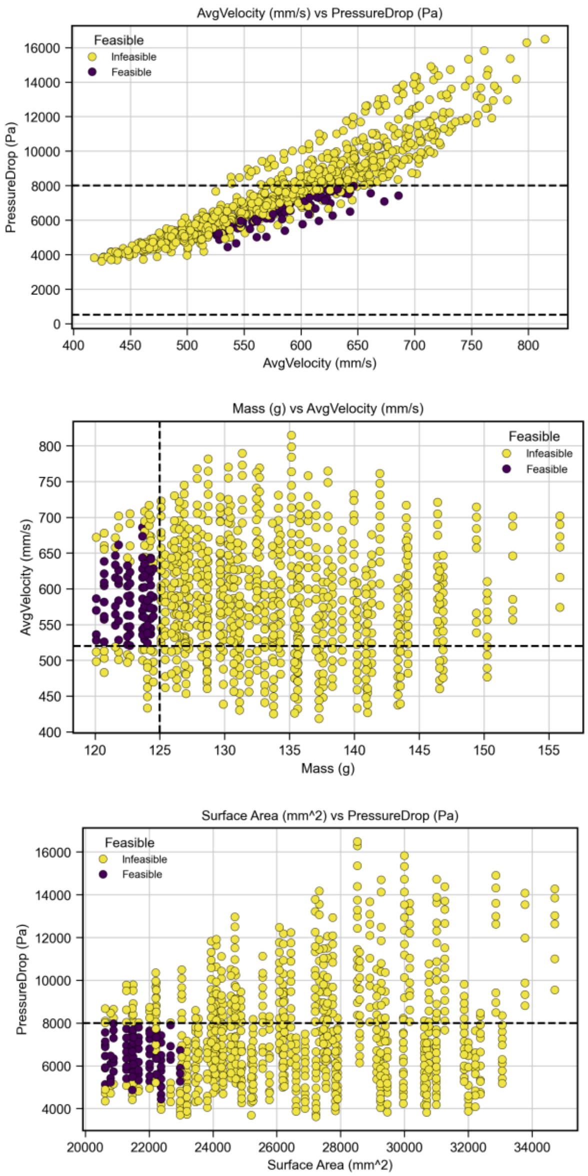


Figure 5. Feasible design trade-off between different outputs

examine the trade-off between accuracy and computational efficiency. Eventually, the model with the best RMSE and efficient run time was utilized for further optimal and inverse design analyses.

Cross-validation (CV) provides a robust estimate on small datasets, which allows us to compare the average RMSE across all targets. This would further prevent the risk of overfitting on a part of the data and enhance the model's performance in the whole dataset. CV is mainly for model selection and hyperparameter tuning. It estimates how a model might generalize unseen data without committing to a final test set. CV gives a less biased signal for picking the best algorithm before touching the test set. Also, cross-validation before training and testing the model reduces the data leakage. The training pipeline was designed to automatically test and examine different algorithms to determine which model works better for a physics-based analysis. A repeated K-Fold strategy (5 splits × 3 repeats) was adopted, giving 15 train–test partitions in total. This setup reduces variance compared to a single split and provides a more reliable measure of model generalization. To evaluate each candidate model fairly, custom scoring functions were defined for all targets, and a combined RMSE score aggregated performance into a single comparison metric.

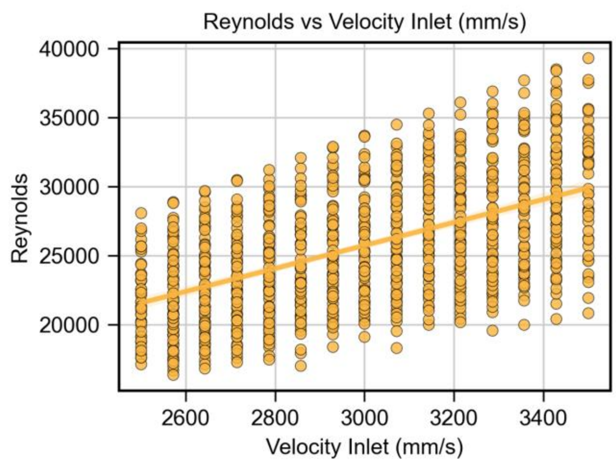


Figure 6. Reynolds number vs. inlet velocity

The models were wrapped in pipelines that handled preprocessing, scaling, and multi-output regression. For baselines, Linear Regression and a quadratic Polynomial Features + Linear Regression variant were included, capturing linear and simple nonlinear trends such as the quadratic dependence of pressure drop on flow velocity. Ensemble tree methods were tested to address more complex nonlinearities and interactions, including Random Forest, which averages predictions across many deep trees to reduce variance. Also, Histogram Gradient Boosting leverages shallow trees with quantile loss for robustness to outliers while controlling bias–variance trade-offs with regularization and early stopping. Neural approaches were also explored to capture smooth, high-dimensional nonlinear mappings. A two-layer MLP (128, 64 neurons) with ReLU activation and L2 regularization was used.

The neural network used early stopping to prevent overfitting, batch training for efficiency, and dropout or weight decay for generalization. All models were cross-validated using the same pipeline, and results were aggregated into a comparison table sorted by combined RMSE.

The evaluation stage consisted of statistical error metrics and visual comparisons of model predictions against the CFD results. The best model for accuracy and generality was selected through the comparison framework. This setup ensured that the surrogate model captured the underlying physics while remaining fast enough for use in the inverse design optimization stage.

3. Results and Discussion

3.1 Model Prediction

The results demonstrate clear trade-offs between accuracy and computational cost across the models. HistGradientBoosting achieved the lowest overall error, with a combined RMSE of 471. Random Forests performed comparably, reaching the lowest surface area and mass error, but with slightly higher combined RMSE. The ensemble methods, therefore, represent the strongest candidates for surrogate modeling in this application. The polynomial regression model (Poly2+LR) offered a good compromise between speed and accuracy. While its error was significantly higher than the ensembles, its runtime was 0.006s. The neural models showed mixed performance. The MLP (128,64) captured some nonlinear patterns but underperformed compared to tree-based ensembles, with a combined

RMSE of 762 and a notably higher runtime of 13 seconds. This outcome highlights a significant limitation: 890 data points are insufficient for training a deep neural network.

In summary, the tree-based ensemble methods (HGB and RF) strike the best balance of accuracy and computational cost, making them strong candidates for surrogate modeling in inverse design workflows. Figure 7 illustrates the HGB model predictions with the actual results for each output. Also, Table 1 shows statistics considering the models' RMSE for all outputs and their computational running time.

Table1. Model's performance summary

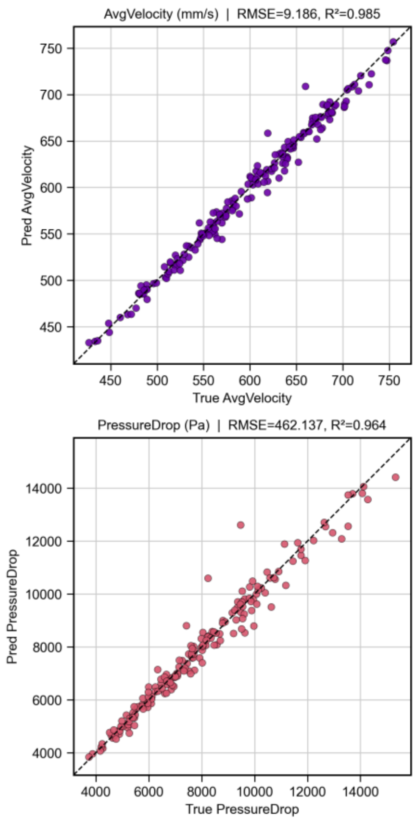
Model	RMSE Press. Drop	RMSE Avg. Vel.	RMSE Surf. Area	RMSE Mass	Run Time (s)
HistGradientBoosting	462.14	9.19	208.43	0.76	1.68
Random Forest	473.49	13.44	45.36	0.14	1.86
Poly2+LR	695.33	16.34	437.14	0.42	0.006
MLP (128,64)	746.04	15.51	373.48	0.41	12.93
Linear Regression	1015.01	19.00	976.78	1.72	0.005

3.2 Inverse Design

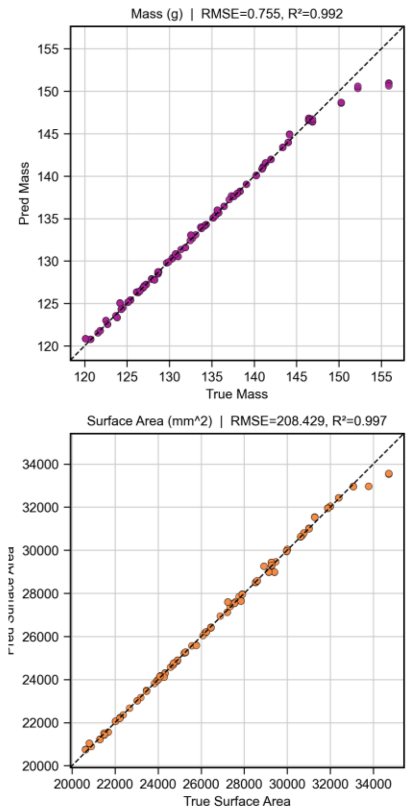
Using the trained surrogate model, the optimal input parameters for the heat exchanger were determined as 19.6511226 mm for the X cell size, 24.56128764 mm for the Y/Z cell size, and 2747.17462971 mm/s for the inlet velocity. These inputs yield predicted outputs of average velocity 537.565 mm/s, mass 121.821 g, pressure drop 5124.33 Pa, surface area 21516.3 mm², and Reynolds number 29295.4. Figure 8 shows a screenshot of nTop running the HEX model with these optimal inputs, confirming the surrogate model's predictions.

4. Conclusion

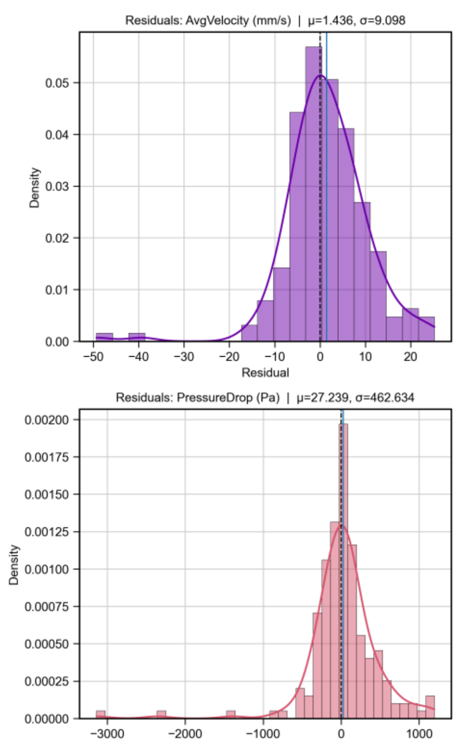
This study showed how effective surrogate physics modeling can be for optimizing heat exchangers. By using data augmentation, exploratory data analysis, and machine learning, we developed a surrogate model that can accurately predict important heat exchanger performance metrics. These include average velocity, pressure drop, surface area, and mass. Among all the tested models, tree-based ensemble methods, specifically Histogram Gradient Boosting, provided the best balance between prediction accuracy and computational efficiency. This makes it ideal for physics-informed inverse design for this specific heat exchanger. The trained surrogate model allowed for the quick identification of optimal parameters within the design limits.



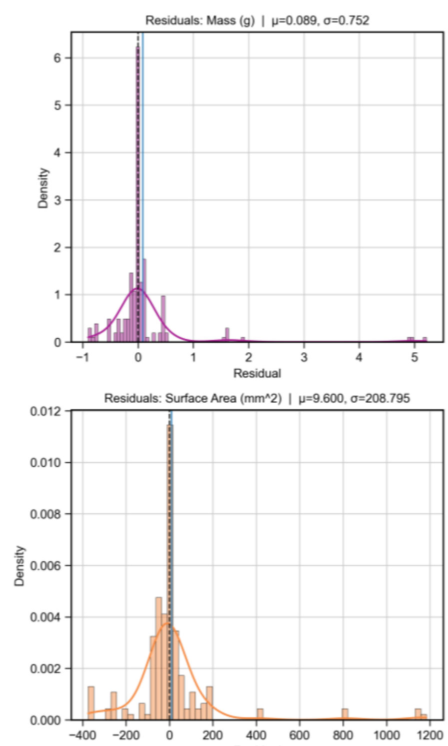
(a)



(b)



(c)



(d)

Figure 7. (a and b) Best model predictions versus actual data for all outputs, (c and d) Residuals for the models' predictions per each output

It achieved a high surface area while keeping mass, pressure drop, and flow velocity within required bounds. When compared to nTop simulation results, the model showed reliable predictions, with average errors below 8% across all outputs. In summary, this work underscores the promise of machine learning-driven surrogate physics models to speed up design exploration in engineering. By reducing the need for costly CFD simulations, this method enables quicker iterations, wider parameter exploration, and practical use in inverse design processes for heat exchanger optimization.

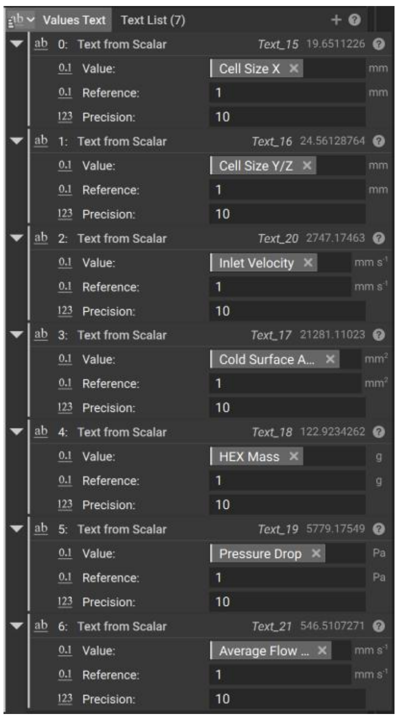


Figure **8**. nTop results on optimal parameters

5. Future Works

Future studies can focus on improving model performance through expanded data generation and physics-informed feature engineering, enabling deeper neural networks such as MLPs to close the current accuracy gap. Integrating physics-based empirical equations may further enhance surrogate accuracy.

6. References

1.Jeon J, Lee J, Vinuesa R, Kim SJ. Residual-based physics-informed transfer learning: A hybrid method for accelerating long-term CFD simulations via deep learning. Int J Heat Mass Transf [Internet]. 2024 Mar 1 [cited 2025 Aug 13];220:124900.

2.Kwon B, Ejaz F, Hwang LK. Machine learning for heat transfer correlations. International Communications in Heat and Mass Transfer [Internet]. 2020 Jul 1 [cited 2025 Aug 13];116:104694.

3.Krishnayatra G, Tokas S, Kumar R. Numerical heat transfer analysis & predicting thermal performance of fins for a novel heat exchanger using machine learning. Case Studies in Thermal Engineering [Internet]. 2020 Oct 1 [cited 2025 Aug 13];21:100706.

APPENDIX A

This appendix demonstrates the model's predictive reliability and quantifies its deviations from the results of nTop simulations.

Table A1. Random test inputs for model prediction assessment

	X Cell (mm)	Y/Z Cell (mm)	Inlet Velocity (mm/s)
Case 1	12	15	3000
Case 2	24	23	3490
Case 3	20	20	3450
Case 4	18	14	2980
Case 5	14	12	2600
Case 6	16	17	2800
Case 7	11	16	3200
Case 8	14	11	2400*

* intentionally probe extrapolation

Table A2. Model prediction vs. nTop output for average velocity

Avg. Velocity (mm/s)	nTop Output	Model Prediction	Error (%)
Case 1	607.93	616.09	1.34
Case 2	671.73	670.82	0.14
Case 3	623.92	649.6	4.09
Case 4	544.3	562.01	3.26
Case 5	491.44	491.72	0.06
Case 6	568.16	532.8	6.21
Case 7	692.54	693.1	0.08
Case 8	452.24	472.597	4.48

Table A3. Model prediction vs. nTop output for pressure drop

Pressure Drop (Pa)	nTop Output	Model Prediction	Error (%)
Case 1	9185.18	8764.86	4.57
Case 2	8998.1	8919.65	0.87
Case 3	8560.9	7756.87	9.38
Case 4	5625.7	6436.889	14.42
Case 5	5460.34	5740.71	5.07
Case 6	7007.58	6289.01	10.22
Case 7	12311.54	12422.57	0.90
Case 8	4657.4	5300.94	13.82

Table A4. Model prediction vs. nTop output for surface area

Surface Area (mm^2)	nTop Output	Model Prediction	Error (%)
Case 1	29111.34	27599.66	5.15
Case 2	21930.21	20898.22	4.70
Case 3	23560.57	23166.17	1.69
Case 4	27428.6	27541.35	0.42
Case 5	30821.05	30811.42	0.03
Case 6	26203.48	24866.98	5.11
Case 7	29254.9	29943.01	2.35
Case 8	31343.96	31955.63	1.95

Table A5. Model prediction vs. nTop output for mass

Mass (g)	nTop Output	Model Prediction	Error (%)
Case 1	139.86	137.25	1.83
Case 2	121.97	121.57	0.33
Case 3	126.29	126.3	0.01
Case 4	136	136.02	0.01
Case 5	140.01	144.98	3.56
Case 6	132.42	146.45	10.55
Case 7	139.82	139.97	0.11
Case 8	146.84	139.97	4.69

Table A6. Model prediction vs. nTop output average error

Target Outputs	Model Prediction vs. nTop Error (%)
Average Velocity	2.73
Pressure Drop	7.53
Surface Area	2.93
Mass	2.86