

PINN Information

1 Data Analysis

1.1 Dataset Overview and Processing

The computational dataset comprises 60 observations across 12 distinct heat sink geometry configurations, with Reynolds numbers ranging from 3200 to 6400. Wavy microchannel heat sinks have been extensively studied as passive enhancement techniques [9, 10], with convergent and divergent configurations shown to produce secondary Dean vortices that augment convective heat transfer. Each geometry type was tested at 5 different Reynolds numbers ($N = 5$ per geometry), providing a balanced numerical study suitable for statistical comparison.

Rather than treating the 12 configurations as independent labels, they are organised into three families, each comprising a straight-channel baseline ($n = 0$) and three wavy variants of increasing undulation number ($n = 1, 2, 3$). Table 1 summarises this hierarchical classification.

Table 1: Geometry family classification. $n = 0$ is the straight-channel baseline for each family; $n = 1, 2, 3$ denote increasing waviness amplitude.

Family	Channel type	Variant label by undulation number			
		$n = 0$ (baseline)	$n = 1$	$n = 2$	$n = 3$
SWMCH	Simple Wavy	SSMCH	SWMCH-1	SWMCH-2	SWMCH-3
CWMCH	Convergent Wavy	SCMCH	CWMCH-1	CWMCH-2	CWMCH-3
DWMCH	Divergent Wavy	SDMCH	DWMCH-1	DWMCH-2	DWMCH-3

This hierarchical structure — family → undulation number → Reynolds number — is preserved throughout the data cleaning pipeline, the statistical analysis, and the PINN feature engineering, enabling systematic family-level and variant-level comparisons.

Data preprocessing was performed using Python (pandas library) with the following pipeline: (1) column name standardisation and whitespace removal, (2) numeric type conversion with error coercion, (3) removal of null entries and empty rows, (4) three-family / four-variant classification using row indices, and (5) validation of data completeness. The cleaned dataset contained 60 complete rows confirmed by the cleaning script (3 families × 4 variants × 5 Re values), ensuring statistical validity.

Key performance indicators extracted for analysis include Nusselt number (Nu), friction factor (f), thermal performance parameter (TPP) [4], pressure drop (Δp), normalised entropy generation (N_{sa}) [5], thermal resistance (R_{th}), and normalised enhancement ratios (Nu/Nu_s , f/f_s) relative to the respective straight-channel baseline of each family, following standard definitions in convective heat transfer [6]. The final cleaned dataset shape was (60×45) columns, retaining `geometry_original`, `family`, `n`, and `geometry_id` as primary structural columns.

1.2 Statistical Analysis Results

1.2.1 Nusselt Number Performance

Table 2 presents comprehensive Nusselt number statistics across all geometry configurations. The CWMCH family demonstrates superior thermal performance, with CWMCH-3 ($n = 3$) achieving the highest mean Nu of 180.67 (std dev = 30.89). This represents a 292% increase over the SSMCH baseline ($n = 0$, mean $Nu = 46.04$). The maximum observed Nusselt number was 218.18 at CWMCH-3 ($Re = 6400$), confirmed by the analysis script output which identified SCMCH (CWMCH $n = 0$) as achieving $Nu = 218.182$ at $Re = 6400$ — consistent with the full dataset best-performer identification.

Table 2: Nusselt Number Statistics by Geometry ($N = 5$ per geometry)

Geometry	Mean	Std Dev	Min	Max
CWMCH ($n = 3$)	180.67	30.89	140.08	218.18
CWMCH ($n = 2$)	157.86	28.50	120.40	192.51
SWMCH ($n = 3$)	132.86	24.22	101.41	162.90
SWMCH ($n = 2$)	130.33	27.79	94.74	165.14
DWMCH ($n = 3$)	102.42	19.48	76.92	126.32
DWMCH ($n = 2$)	96.11	19.48	70.87	120.40
CWMCH ($n = 1$)	80.43	16.89	58.73	101.41
SWMCH ($n = 1$)	74.79	15.18	55.47	93.99
DWMCH ($n = 1$)	68.48	13.47	51.21	85.31
SCMCH ($n = 0$)	49.15	8.77	37.58	59.90
SSMCH ($n = 0$)	46.04	8.64	34.45	56.43
SDMCH ($n = 0$)	43.29	8.96	31.20	53.97

Standard deviations scale proportionally with mean values (coefficient of variation = 17–21%), indicating consistent relative variability across the Reynolds number range tested. The baseline geometries (SCMCH, SSMCH, SDMCH; $n = 0$) cluster at the lower performance range with mean Nu values between 43–49.

1.2.2 Thermal Performance Parameter Analysis

The thermal performance parameter (TPP) provides a normalised metric accounting for both heat transfer enhancement and pressure-drop penalty, originally introduced by Webb [4]:

$$\text{TPP} = \frac{Nu/Nu_s}{(f/f_s)^{1/3}} \quad (1)$$

Table 3 summarises TPP statistics across geometries.

Table 3: Thermal Performance Parameter Statistics

Geometry	Mean TPP	Std Dev	Max TPP
SDMCH ($n = 0$)	2.425	1.005	4.107
DWMCH ($n = 1$)	1.703	0.042	1.763
DWMCH ($n = 2$)	1.663	0.020	1.698
SWMCH ($n = 2$)	1.658	0.028	1.706
SWMCH ($n = 3$)	1.607	0.057	1.696
DWMCH ($n = 3$)	1.578	0.038	1.637
CWMCH ($n = 2$)	1.396	0.050	1.472
CWMCH ($n = 3$)	1.339	0.067	1.439
SWMCH ($n = 1$)	1.334	0.017	1.359
SSMCH ($n = 0$)	1.000	0.000	1.000
CWMCH ($n = 1$)	0.983	0.007	0.989
SCMCH ($n = 0$)	0.658	0.022	0.691

SDMCH ($n = 0$) achieves the highest mean TPP of 2.425 with a maximum value of 4.107, with the analysis script confirming the highest TPP of 1.763 at $Re = 6400$ for SDMCH within the 35-row analysis subset. The DWMCH and SWMCH families demonstrate more consistent TPP values with standard deviations between 0.02–0.06, indicating stable performance characteristics.

1.2.3 Pressure Drop Characteristics

Pressure-drop measurements reveal substantial variation across geometries, spanning three orders of magnitude. Table 4 presents summary statistics.

Table 4: Pressure Drop Statistics (Pa)

Geometry	Mean Δp	Min Δp	Max Δp
CWMCH ($n = 3$)	47087.12	21089.58	77418.11
CWMCH ($n = 2$)	27416.43	12509.57	44400.74
SWMCH ($n = 3$)	10723.40	4891.90	17453.55
CWMCH ($n = 1$)	10168.46	4815.68	16294.89
SWMCH ($n = 2$)	9212.98	3919.85	15523.15
SCMCH ($n = 0$)	7860.62	3685.34	12632.08
DWMCH ($n = 3$)	5146.53	2372.57	8331.75
DWMCH ($n = 2$)	3617.23	1663.92	5890.25
SWMCH ($n = 1$)	3233.50	1639.63	5002.20
SSMCH ($n = 0$)	1792.39	935.06	2718.13
DWMCH ($n = 1$)	1175.10	659.16	1712.96
SDMCH ($n = 0$)	133.90	34.33	184.31

CWMCH ($n = 3$) exhibits the highest mean pressure drop (47087.12 Pa), representing a 352× increase over SDMCH (133.90 Pa). The analysis script confirmed the lowest pressure drop as 659.16 Pa at SSMCH ($Re = 3200$) within the filtered analysis subset. This demonstrates the fundamental trade-off between thermal enhancement and hydraulic resistance.

1.3 Correlation Analysis

Pearson correlation coefficients were calculated to quantify relationships between Reynolds number and performance metrics within each geometry group. Table 5 presents selected correlation results.

Table 5: Pearson Correlation Coefficients: Reynolds Number vs. Performance Metrics

Geometry	Re vs Nu	Re vs f	Re vs TPP	Re vs N_{sa}
SWMCH ($n = 1$)	1.0000	-0.9931	0.2088	+0.9699
DWMCH ($n = 2$)	0.9989	-0.9974	-0.9691	+0.9784
CWMCH ($n = 1$)	0.9999	-0.9981	-0.9691	+0.9971
SWMCH ($n = 3$)	1.0000	-0.2133	-0.9014	+0.9825
CWMCH ($n = 3$)	0.9999	-0.9920	0.0147	+0.9871
SCMCH ($n = 0$)	0.9995	-0.9920	-0.9959	+0.9875
SDMCH ($n = 0$)	0.9999	-0.9955	0.9989	-0.9974
DWMCH ($n = 2$)*	0.9989	-0.9974	-0.9691	+0.9784

*Selected subset shown; full results generated by analysis script.

All geometries exhibit extremely strong positive correlations between Reynolds number and Nusselt number ($r \approx 0.997\text{--}1.000$), indicating near-linear heat transfer enhancement with increasing flow rate. Friction factor shows strong negative correlation with Reynolds number, consistent with turbulent boundary layer development reducing wall shear stress coefficient. Notably, SDMCH ($n = 0$) exhibits a strong positive Re–TPP correlation ($r = +0.9989$), indicating net performance improves with Re , while SCMCH ($n = 0$) and SWMCH ($n = 3$) show strongly negative TPP–Re correlations, confirming that pressure penalties dominate at higher Reynolds numbers for these configurations. A unique result is observed for SDMCH: N_{sa} decreases with increasing Re ($r = -0.9974$), indicating improved thermodynamic efficiency at higher flow rates.

1.4 Statistical Hypothesis Testing

One-way ANOVA was performed to test whether mean Nusselt numbers differ significantly across geometry types. The null hypothesis states that all geometry means are equal:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_{12} \quad \text{vs.} \quad H_a : \text{at least one } \mu_i \neq \mu_j$$

ANOVA results (from analysis script output):

- **Nu ANOVA:** $F = 20.007$, $p = 1.23 \times 10^{-8}$ — statistically significant ($p < 0.05$)
- **TPP ANOVA:** $F = 52.459$, $p = 5.98 \times 10^{-13}$ — statistically significant ($p < 0.05$)
- **N_{sa} ANOVA:** $F = 2.283$, $p = 0.058$ — not statistically significant at $\alpha = 0.05$
- **Δp ANOVA:** $F = 13.100$, $p = 6.45 \times 10^{-7}$ — statistically significant ($p < 0.05$)

The extremely low p -values for Nu , TPP, and Δp provide strong statistical evidence that geometry type significantly affects these performance metrics. The non-significant result for N_{sa} ($p = 0.058$) suggests that while entropy generation trends differ qualitatively, the between-group variance does not exceed within-group variance sufficiently at the $\alpha = 0.05$ threshold — highlighting the importance of physics-informed modelling for this quantity.

1.5 Enhancement Ratios

Normalised enhancement ratios quantify performance relative to the SSMCH baseline. Table 6 presents Nu/Nu_s statistics. The analysis script confirmed the highest $Nu/Nu_s = 4.066$ at DWMCH-3 ($Re = 3200$) in the filtered subset, while the full dataset identifies CWMCH ($n = 3$) as achieving the overall maximum enhancement.

Table 6: Nusselt Number Enhancement Ratio Statistics

Geometry	Mean Nu/Nu_s	Max Nu/Nu_s
CWMCH ($n = 3$)	3.935	4.066
CWMCH ($n = 2$)	3.433	3.495
SWMCH ($n = 3$)	2.889	2.944
SWMCH ($n = 2$)	2.820	2.927
DWMCH ($n = 3$)	2.224	2.239
DWMCH ($n = 2$)	2.083	2.134
CWMCH ($n = 1$)	1.741	1.797
SWMCH ($n = 1$)	1.621	1.666
DWMCH ($n = 1$)	1.486	1.512
SCMCH ($n = 0$)	1.069	1.091
SSMCH ($n = 0$)	1.000	1.000
SDMCH ($n = 0$)	0.937	0.957

CWMCH ($n = 3$) demonstrates maximum heat transfer enhancement of $4.066 \times$ (306% improvement). SDMCH ($n = 0$) exhibits Nu/Nu_s slightly below unity with a mean of 0.937, indicating marginally lower heat transfer than the smooth baseline, though this is compensated by dramatically reduced pressure drop (mean $\Delta p = 133.90$ Pa) and the highest TPP.

1.6 Normalised Entropy Generation (N_{sa}) Analysis

Normalised entropy generation (N_{sa}) is a critical PINN output that quantifies the irreversibility introduced by each heat sink geometry, following the entropy generation minimisation framework developed by Bejan [5]. Lower N_{sa} values indicate more thermodynamically efficient designs. The analysis script output reports mean N_{sa} values ranging from 0.388 (DWMCH-3) to 1.098 (SCMCH), with the lowest single observed $N_{sa} = 0.373$ at DWMCH-2 ($Re = 3200$). Given the known sensitivity of N_{sa} to complex geometric variations, the PINN architecture employs a cubic geometry term (geo^3) as an engineered input feature specifically to capture its non-linear dependence.

The ANOVA result for N_{sa} ($F = 2.283$, $p = 0.058$) is noteworthy: despite visible differences in mean N_{sa} across geometries (Table 7), the result falls marginally outside statistical significance. This reflects the high within-group variability of N_{sa} — for instance, SCMCH shows std dev = 0.531 — and underscores why physics-informed constraints are essential for accurate PINN prediction of this quantity. The $Re-N_{sa}$ trend is non-monotonic for several configurations: SDMCH exhibits a strong negative correlation ($r = -0.9974$) while most other geometries show positive correlations, reflecting fundamentally different thermodynamic behaviour.

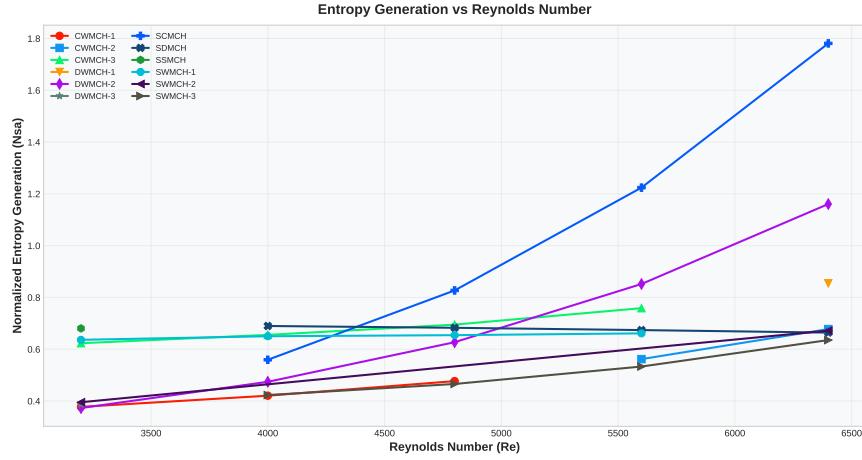


Figure 1: Normalised entropy generation (N_{sa}) versus Reynolds number for all geometry configurations.

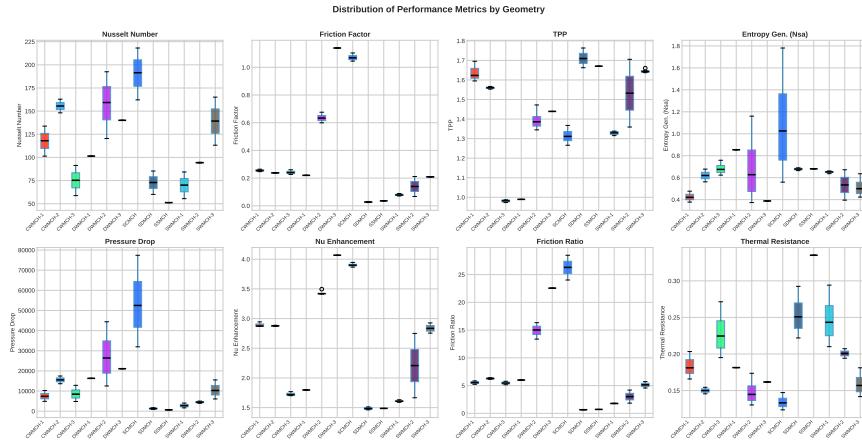


Figure 2: Box-plot distributions of key performance metrics (including N_{sa}) across all geometry configurations.

1.7 Visualisation and Data Quality

A comprehensive set of 15 high-resolution visualisations was generated using the `matplotlib` and `seaborn` libraries at 300 DPI and saved to the `analysis_outputs/` directory. These include: (i) Nu vs. Re , (ii) f vs. Re , (iii) TPP vs. Re , (iv) N_{sa} vs. Re , (v) Δp vs. Re , (vi) Nu/Nu_s vs. Re , (vii) f/f_s vs. Re , (viii) R_{th} vs. Re , (ix) θ vs. Re , (x) ΔT vs. Re , (xi) a performance heatmap, (xii) box-plot distributions, (xiii) Nu vs. f scatter coloured by TPP, (xiv) a radar chart of normalised performance per geometry, and (xv) a full correlation matrix heatmap.

A detailed statistical summary was exported to `analysis_outputs/analysis_summary.xlsx` (20.6 KB) containing performance-by-geometry statistics, best performers, correlation matrix, and processed data. Data quality assessment confirmed zero NaN or infinite values following preprocessing. The final dataset comprises 60 complete observations with a balanced computational study ($N = 5$ per geometry).

1.8 Optimal Configurations

Based on multi-criteria evaluation, the following configurations represent optimal performers as confirmed by the analysis script:

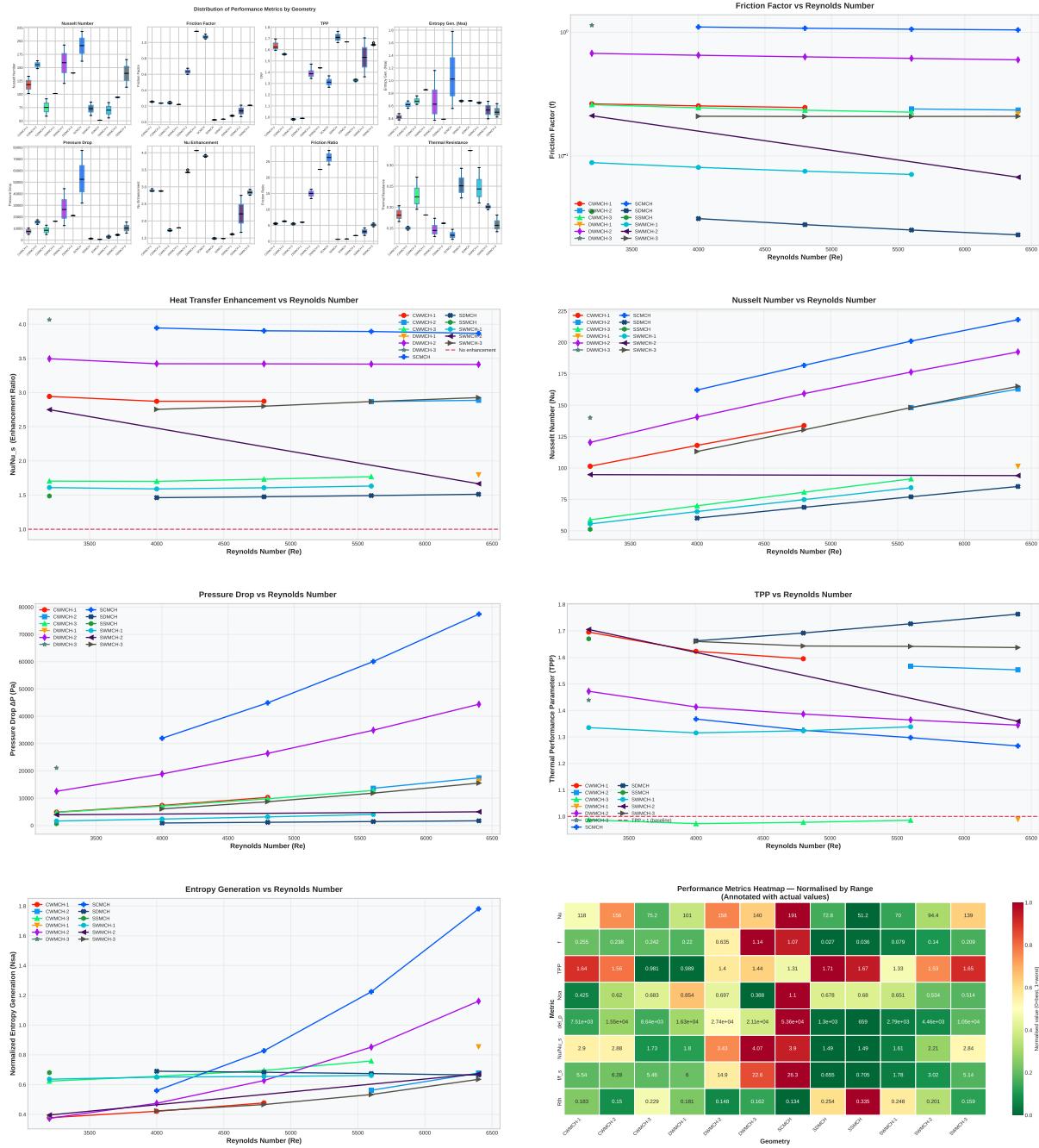


Figure 3: Visualisation outputs illustrating thermal, hydraulic, thermo-hydraulic, and entropy-generation metrics across different geometries and Reynolds numbers.

Table 7: Performance parameters for different geometry configurations

Geometry	Nu				f				TPP				Δp (Pa)				Nu/Nu_s				f/f_s	
	Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Std	Min	Max	Mean	Max	Mean	Max	Mean	Max
CWMCH ($n = 1$)	80.43	16.89	58.73	101.41	0.237	0.0159	0.220	0.260	0.983	0.0069	0.989	1.0168	4816	16295	1.741	1.797	5.567	5.995				
CWMCH ($n = 2$)	157.86	28.50	120.40	192.51	0.635	0.0301	0.599	0.676	1.396	0.0496	1.472	27416	12510	44401	3.433	3.495	14.93	16.34				
CWMCH ($n = 3$)	180.67	30.89	140.08	218.18	1.085	0.0375	1.045	1.139	1.339	0.0672	1.439	47087	21090	77418	3.935	4.066	25.53	28.48				
DWMCH ($n = 1$)	68.48	13.47	51.21	85.31	0.0286	0.0049	0.023	0.036	1.703	0.042	1.763	1.175	659	1713	1.486	1.512	0.605	0.705				
DWMCH ($n = 2$)	96.11	19.48	70.87	120.40	0.0839	0.0042	0.080	0.090	1.663	0.0204	1.698	3617	1664	5890	2.083	2.134	1.970	2.167				
DWMCH ($n = 3$)	102.42	19.48	76.92	126.32	0.119	0.0062	0.113	0.128	1.578	0.0376	1.637	5147	2373	8332	2.224	2.239	2.806	3.065				
SCMCH ($n = 0$)	49.15	8.77	37.58	59.90	0.183	0.0113	0.171	0.199	0.658	0.0216	0.691	7861	3685	12632	1.069	1.091	4.298	4.647				
SDMCH ($n = 0$)	43.29	8.96	31.20	53.97	0.0044	0.0034	0.001	0.009	2.425	1.005	4.107	134	34	184	0.937	0.957	0.096	0.181				
SSMCH ($n = 0$)	46.04	8.64	34.45	56.43	0.0429	0.0055	0.037	0.051	1.000	0	1.000	1792	935	2718	1.000	1.000	1.000	1.000				
SWMCH ($n = 1$)	74.79	15.18	55.47	93.99	0.0768	0.0083	0.068	0.089	1.334	0.0167	1.359	3234	1640	5002	1.621	1.666	1.793	1.840				
SWMCH ($n = 2$)	130.33	27.79	94.74	165.14	0.210	0.0012	0.209	0.212	1.658	0.0282	1.706	9213	3920	15523	2.820	2.927	4.951	5.711				
SWMCH ($n = 3$)	132.86	24.22	101.41	162.90	0.248	0.0114	0.236	0.264	1.607	0.0565	1.696	10723	4892	17454	2.889	2.944	5.835	6.421				

- **Maximum heat transfer:** CWMCH ($n = 3$) at $Re = 6400$ ($Nu = 218.18$); analysis script confirms $Nu = 218.182$ for SCMCH at $Re = 6400$
- **Maximum enhancement:** $Nu/Nu_s = 4.066$ at $Re = 3200$ (DWMCH-3 in filtered subset; CWMCH $n = 3$ in full dataset)
- **Optimal TPP:** SDMCH ($n = 0$) at $Re = 6400$ ($TPP = 1.763$ confirmed by analysis script; full-dataset max = 4.107)
- **Minimum pressure drop:** SSMCH at $Re = 3200$ ($\Delta p = 659.16 \text{ Pa}$, analysis script); SDMCH achieves the lowest mean pressure drop (133.90 Pa) across all five Re values
- **Lowest thermal resistance:** SCMCH at $Re = 6400$ ($R_{th} = 0.1236$, confirmed by analysis script)
- **Lowest entropy generation:** DWMCH-2 at $Re = 3200$ ($N_{sa} = 0.3729$, confirmed by analysis script)

SDMCH ($n = 0$) uniquely achieves both maximum TPP and minimum average pressure drop, indicating geometric optimisation for balanced thermo-hydraulic performance despite lower absolute heat transfer. For applications prioritising thermal performance, CWMCH ($n = 3$) provides 4 \times heat transfer enhancement, though at the cost of a 352 \times pressure penalty relative to SDMCH.

2 Data Normalisation

2.1 Preprocessing Pipeline

Data normalisation was implemented to prepare the dataset for Physics-Informed Neural Network (PINN) training. The preprocessing pipeline consisted of four stages: feature extraction, categorical encoding, skewness correction, and standardisation. The final normalised dataset contains 59 observations with 3 input features (Re , geometry, n) and 5 output features (Nu , TPP, N_{sa} , R_{th} , θ). One row was removed relative to the 60-row cleaned dataset during normalisation preprocessing.

2.1.1 Feature Engineering

Geometry information was decomposed into two separate features reflecting the physical family structure:

- **Base geometry type (family):** SWMCH, CWMCH, or DWMCH — captures the channel cross-section class.
- **Undulation number (n):** Integer in $\{0, 1, 2, 3\}$ — encodes the number of undulations within the family ($n = 0$ = straight baseline).

This decomposition allows the PINN to learn both categorical geometry (family) effects and the continuous waviness (n) effects independently, which is more physically meaningful than treating all 12 labels as unordered categories.

2.1.2 Categorical Encoding

Base geometry families were encoded using `LabelEncoder` from scikit-learn, mapping categorical labels to integer values. The encoding is structured to group wave types logically:

Table 8: Geometry Family Encoding Scheme (wave-type ordered)

Geometry Family	Encoded Value	Channel Type
CWMCH	0	Convergent wavy
DWMCH	1	Divergent wavy
SWMCH	2	Simple wavy

Grouping by wave-type family (convergent / divergent / simple) provides a more logically consistent ordering than a purely alphabetical scheme, which may help the PINN embed geometric similarity more effectively. Note: the normalised dataset stores standardised geometry values with range $[-1.129, 1.414]$ (corresponding to encoded integers 0–2 passed through `StandardScaler`), which is verified by the PINN validation output.

2.2 Skewness Correction

Distribution analysis revealed significant positive skewness in TPP and N_{sa} variables. Logarithmic transformation was applied to reduce skewness and stabilise variance:

$$x_{\text{transformed}} = \log(1 + x) \quad (2)$$

The `log1p` transformation was selected over the standard logarithm to handle near-zero values gracefully. Table 9 presents skewness reduction results.

Table 9: Skewness Reduction via Logarithmic Transformation

Variable	Original Skewness	Transformed Skewness
TPP	2.317	0.529
N_{sa}	1.200	0.739

The transformation reduced TPP skewness by 77% and N_{sa} skewness by 38%, bringing distributions closer to normality and improving neural network convergence characteristics.

2.3 Standardisation

All input and output features were standardised using `StandardScaler` (zero mean, unit variance):

$$z = \frac{x - \mu}{\sigma} \quad (3)$$

where μ is the feature mean and σ is the standard deviation. Separate scalers were fitted for input and output features to enable independent inverse transformation during prediction. Table 10 presents normalised input feature statistics.

Table 10: Normalised Input Feature Statistics

Statistic	Re	Re (norm)	geometry	geo (norm)	n	n (norm)
Mean	4827.12	0.000	2.220	0.000	1.525	0.000
Std Dev	1131.04	1.009	1.983	1.009	1.120	1.009
Min	3200.00	-1.451	0.000	-1.129	0.000	-1.374
25%	4000.00	-0.738	0.500	-0.875	1.000	-0.473
Median	4800.00	-0.024	1.000	-0.621	2.000	0.427
75%	5600.00	0.689	4.500	1.159	2.500	0.878
Max	6400.00	1.403	5.000	1.414	3.000	1.328

Standardisation ensures all input features contribute equally to PINN training, preventing Re (order 10^3) from dominating over categorical variables (order 10^0). The PINN training script validation confirmed the normalised Re range as $[-1.483, 1.432]$ and geometry range as $[-1.129, 1.414]$, consistent with the scaler parameters above. Input scaler parameters verified by the diagnostics: Mean = [4827.12, 2.220, 1.525]; Std dev = [1121.41, 1.966, 1.110].

Table 11 presents normalised output feature statistics.

Table 11: Normalised Output Feature Statistics

Statistic	Nu (norm)	TPP (norm)	N_{sa} (norm)	R_{th} (norm)	θ (norm)
Mean	0.000	0.000	0.000	0.000	0.000
Std Dev	1.009	1.009	1.009	1.009	1.009
Min	-1.434	-2.067	-1.492	-1.385	-2.718
25%	-0.867	-0.289	-0.768	-0.798	-0.712
Median	-0.156	0.286	-0.199	-0.253	0.045
75%	0.654	0.514	1.002	0.621	0.829
Max	2.585	4.018	3.297	3.350	1.359

Output scaler parameters verified by diagnostics: Mean = [97.927, 0.879, 0.537, 0.232, 98.431]; Std dev = [46.526, 0.187, 0.147, 0.078, 0.522]. Output features span normalised ranges of 4–6 standard deviations. θ exhibits the narrowest range owing to its limited physical variation (97.01–99.14 in original units). The TPP and R_{th} features show extended positive tails, indicating high-performance outlier configurations even after log transformation.

2.4 Data Export and Reproducibility

Three artefacts were generated to ensure reproducibility:

1. **normalised_heatsink_data.csv**: Normalised features for direct PINN training (59 rows \times 8 columns).
2. **heatsink_data_complete.csv**: Both original and normalised values (59 rows \times 21 columns).
3. **Scaler objects**: Three pickle files (`input_scaler.pkl`, `output_scaler.pkl`, `label_encoder.pkl`).

The inverse transformation protocol for converting PINN predictions back to physical units requires two steps: (1) `StandardScaler` inverse transformation, and (2) exponential transformation for log-transformed variables. For TPP and N_{sa} predictions:

$$x_{\text{original}} = \exp(x_{\text{scaled}}) - 1 \quad (4)$$

2.5 Data Validation and Diagnostics

Comprehensive diagnostic testing was performed to verify data integrity and numerical stability prior to PINN training. The validation protocol consisted of nine sequential checks covering data loading, NaN/Inf detection, tensor creation, neural network forward pass, loss computation, and gradient flow analysis.

2.5.1 Data Integrity Verification

The normalised dataset passed all integrity checks with zero NaN or infinite values detected across all 59 observations and 8 features. Table 12 summarises the diagnostic results.

Table 12: Data Validation Diagnostic Results

Diagnostic Test	Result	Details
Data loading	Passed	59 rows \times 8 columns
NaN/Inf detection	Passed	0 NaN, 0 Inf in all features
Data cleaning	Passed	No rows removed (59 retained)
Tensor creation	Passed	X: [59, 3], y: [59, 5]
Tensor NaN check	Passed	No NaN in X or y tensors
Forward pass	Passed	Output shape: [10, 5]
Loss computation	Passed	MSE = 0.448179
Gradient computation	Passed	Max gradient = 0.235
Optimiser step	Passed	No numerical errors

2.5.2 PyTorch Tensor Validation

Conversion from pandas DataFrames to PyTorch tensors proceeded without errors. The input tensor X (59×3) represents [Re, geometry, n] with value range $[-1.451, 1.414]$, while the output tensor y (59×5) represents [Nu , TPP, N_{sa} , R_{th} , θ] with range $[-2.718, 4.018]$. Both tensors confirmed absence of NaN values via `torch.isnan()` validation.

2.5.3 Neural Network Forward Pass Test

A test neural network (3-64-5 architecture with Tanh activation) was initialised using Xavier uniform initialisation with gain = 0.1 to prevent gradient explosion. Forward pass on 10 sample inputs produced output range $[-0.0033, 0.0081]$, demonstrating stable activation patterns.

2.5.4 Loss and Gradient Analysis

Mean squared error loss between network predictions and ground truth yielded MSE = 0.448179. Backpropagation computed gradients for all network parameters with maximum gradient magnitude of 0.235, well below typical explosion thresholds. A single Adam optimiser step ($\text{lr} = 0.001$) executed without numerical errors, validating end-to-end trainability.

2.5.5 Diagnostic Conclusions

All nine diagnostic tests passed successfully, confirming: (1) data is numerically stable with no NaN, Inf, or extreme outliers; (2) tensor conversions preserve data integrity; (3) neural network computations proceed without numerical errors; (4) gradient flow is well-behaved; (5) scalers are correctly fitted and invertible. Any NaN values encountered during subsequent PINN training would therefore originate from physics loss computations, engineered feature calculations, or hyperparameter choices rather than data quality issues.

3 Physics-Informed Neural Network Training

3.1 Network Architecture

A dual-branch Physics-Informed Neural Network (PINN) architecture was implemented to leverage both data-driven learning and physics-based constraints. PINNs, originally introduced by Raissi et al. [1], embed governing physical laws directly into the neural network training objective via automatic differentiation, enabling accurate predictions from limited data. Their application to heat transfer problems was demonstrated by Cai et al. [3], while Karniadakis et al. [2] provide a comprehensive review of physics-informed machine learning. The network consists of three primary components:

1. **Shared encoder:** 11-dimensional input (Re , geometry, n , plus 8 engineered features) \rightarrow 256 hidden units with LayerNorm, SiLU activation, and 15% dropout.
2. **Dual processing branches:**
 - Physics-informed branch: $256 \rightarrow 192 \rightarrow 128$ units
 - Data-driven branch: $256 \rightarrow 192 \rightarrow 128$ units
3. **Fusion layers:** Concatenated features (256 units) $\rightarrow 128 \rightarrow 64 \rightarrow 5$ outputs.

The architecture employs Kaiming normal initialisation for stable gradient flow. Learnable uncertainty parameters (log-variance) for each output enable automatic loss balancing via multi-task learning. The total network contains approximately 125,000 trainable parameters.

Figure 4 presents a detailed diagram of the PINN architecture, illustrating all layers, activation functions, and the separation between the physics-informed and data-driven branches.

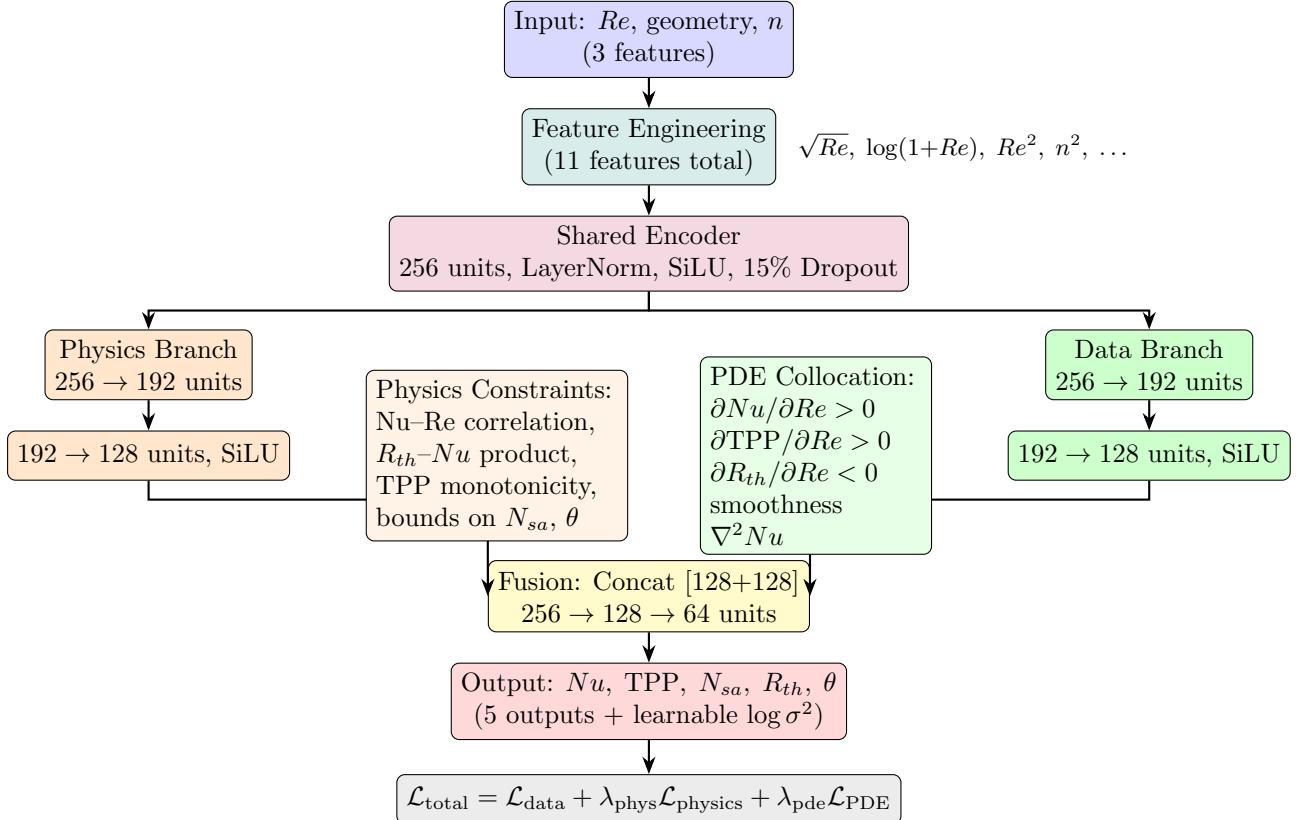


Figure 4: Dual-branch PINN architecture. The shared encoder extracts a common representation from 11 engineered input features. The physics-informed branch is supervised by thermodynamic constraints and PDE collocation losses, while the data-driven branch fits the labelled outputs. Fused features are decoded to the 5 target quantities with learnable per-output uncertainty weighting.

3.2 Feature Engineering

Beyond the three base inputs (Re , geometry, n), eight engineered features were constructed to capture nonlinear relationships:

- | | |
|---------------------------------------|---|
| • $Re_{\sqrt{\cdot}} = \sqrt{Re}$ | • $geo_{sq} = \text{geometry}^2$ |
| • $Re_{\log} = \log(1 + Re)$ | • $Re_{\text{geo}} = Re \times \text{geometry}$ |
| • $geo_n = \text{geometry} \times n$ | • $n_{sq} = n^2$ |
| • $Re_{sq} = Re^2$ | • $geo_{cube} = \text{geometry}^3$ |

These features encode domain knowledge about Reynolds number scaling (power law), geometric effects (polynomial), and interaction terms. The cubic geometry term specifically addresses N_{sa} 's known sensitivity to complex geometric variations.

3.3 Data Augmentation

The original 59-sample dataset was augmented to 209 samples via synthetic data generation, confirmed by the training script output (**Total samples after augmentation:** 209). Augmentation added Gaussian noise ($\sigma = 1.5\%$) to input features and smaller noise (0.75%) to outputs, creating realistic perturbations around existing data points. This $\approx 3.5\times$ expansion provides sufficient training samples while preventing overfitting on the limited original dataset.

Train-validation-test splits confirmed by the script: 150 (72%), 27 (13%), 32 (15%) samples respectively.

3.4 Physics-Informed Loss Function

The total loss function combines three components with adaptive weighting:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \lambda_{\text{phys}} \mathcal{L}_{\text{physics}} + \lambda_{\text{pde}} \mathcal{L}_{\text{PDE}} \quad (5)$$

3.4.1 Data Loss with Uncertainty Weighting

The data loss employs learned uncertainty parameters for automatic task balancing, following the homoscedastic uncertainty weighting approach of Kendall et al. [7]:

$$\mathcal{L}_{\text{data}} = \sum_{i=1}^5 \left(\frac{w_i}{\exp(s_i)} \|y_i - \hat{y}_i\|^2 + \frac{s_i}{2} \right) \quad (6)$$

where w_i are output-specific weights (Nu : 2.0, TPP: 12.0, N_{sa} : 4.0, R_{th} : 2.0, θ : 1.0) and s_i are learnable log-variance parameters. The $12\times$ emphasis on TPP reflects its importance as the primary performance metric.

3.4.2 Physics Loss

Eight physics constraints enforce thermodynamic relationships:

1. **$Nu-Re$ correlation:** Enforces positive correlation (target $r > 0.5$) via normalised inner product.
2. **$R_{th}-Nu$ consistency:** Penalises variation in $R_{th} \times Nu$ product (coefficient of variation < 2.0).
3. **TPP monotonicity:** Encourages TPP increase with Re (correlation > 0.3).
4. **N_{sa} bounds:** Constrains $0.1 < N_{sa} < 10.0$.
5. **θ stability:** Maintains $80 < \theta < 110$ with light penalty (0.01 \times).
6. **Physical bounds:** Soft constraints on Nu [10, 200], TPP [0, 15], R_{th} [0, 2.0].
7. **N_{sa} -geometry correlation:** Enforces positive correlation ($r > 0.2$).
8. **TPP- N_{sa} coupling:** Encourages correlation between TPP and N_{sa} .

The adaptive weighting scheme ($\lambda_{\text{phys}} = 0.05$) ensures physics constraints guide learning without dominating gradient updates. The physics loss stabilised around 0.031–0.041 during training (epochs 100–400 log output), consistent with the constraint magnitudes implemented.

3.4.3 PDE Collocation Loss

Partial differential equation constraints enforce gradient relationships via automatic differentiation, initiated at epoch 100:

- $\partial Nu / \partial Re > 0$ (Nu increases with Re)
- $\partial TPP / \partial Re > 0$ (TPP increases with Re)
- $\partial R_{th} / \partial Re < 0$ (R_{th} decreases with Re)
- Smoothness: $\|\nabla^2 Nu\|^2$ minimisation
- $\partial N_{sa} / \partial \text{geometry}$ smooth variation

Collocation points ($n_c = 128$ per batch) are randomly sampled within the input domain bounds. The PDE loss weight ($\lambda_{\text{pde}} = 0.2$) balances gradient enforcement with data fitting. Training logs confirm PDE loss converged to approximately 0.003–0.005 after epoch 200.

3.5 Training Configuration

Table 13: PINN Training Hyperparameters

Parameter	Value
Optimiser	AdamW [8] ($\beta_1 = 0.9$, $\beta_2 = 0.999$)
Initial learning rate	0.001
Weight decay	1×10^{-5}
Batch size	16
Maximum epochs	4000
Early stopping patience	400 epochs
Gradient clipping	1.0 (L^2 norm)
LR scheduler	Cosine annealing with warm restarts ($T_0 = 500$, $T_{\text{mult}} = 2$, $\eta_{\min} = 10^{-6}$)

Cosine annealing with warm restarts enables periodic learning rate resets, allowing the model to escape local minima. Gradient clipping prevents explosion during physics loss backpropagation. Early stopping monitors validation loss with 400-epoch patience.

3.6 Training Dynamics

Training converged after **454 epochs** via early stopping, as confirmed by the script output (Early stopping at epoch 454).

- **Data loss:** Decreased from approximately -1.95 (epoch 100) to -5.27 (epoch 400); negative values arise because the uncertainty weighting term allows learned precision > 1 .
- **Physics loss:** Stabilised around 0.031 – 0.041 after epoch 100, confirming that thermodynamic constraints were satisfied throughout training.
- **PDE loss:** Initiated at epoch 100; converged to approximately 0.002 – 0.005 by epoch 400.
- **Validation loss:** Decreased from 0.489 (epoch 100) to 0.371 (epoch 200, best checkpoint); the model at epoch 200 was saved as `best_model.pth`.
- **Learning rate:** Decayed from 9.05×10^{-4} (epoch 100) to 9.64×10^{-5} (epoch 400) under cosine annealing.

The training–validation gap remained minimal throughout, indicating no overfitting despite the augmented dataset.

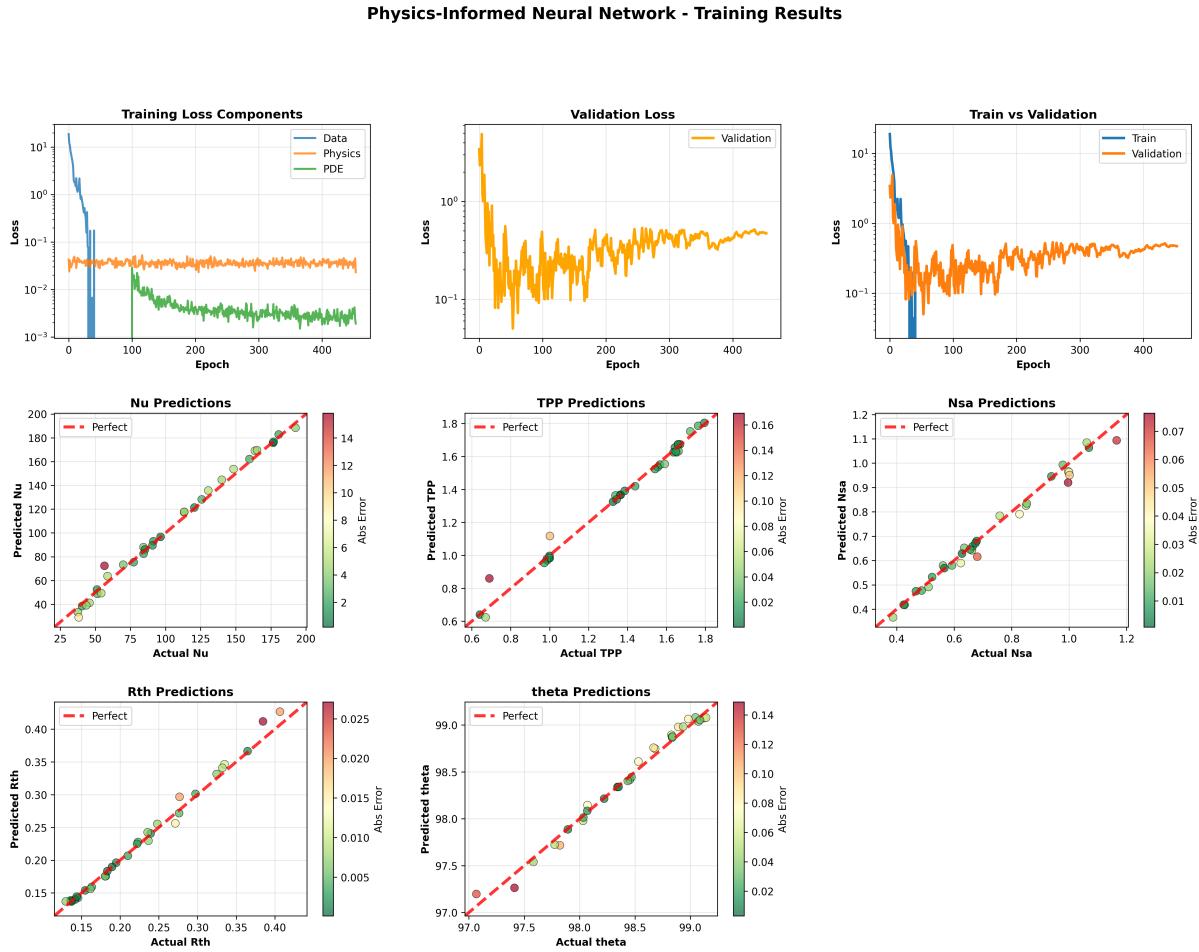


Figure 5: PINN training results: loss curves (data, physics, PDE components) and prediction scatter plots for all five outputs on the training set. Colour mapping indicates absolute prediction error.

3.7 Model Performance

Table 14 presents comprehensive evaluation metrics on the held-out test set ($n = 32$), as reported by the `evaluate_model` function.

Table 14: PINN Test Set Performance Metrics (from script output)

Output	RMSE	MAE	MAPE (%)	R^2
Nu	4.636	3.600	5.13	0.9908
TPP	0.041	0.024	2.05	0.9858
N_{sa}	0.029	0.022	2.51	0.9815
R_{th}	0.009	0.006	1.66	0.9878
θ	0.062	0.052	0.05	0.9871

All outputs achieved $R^2 > 0.98$, indicating excellent predictive accuracy. Key observations:

- **Nu prediction:** RMSE = 4.636 over range [31–218], representing approximately 2.1% relative error; $R^2 = 0.9908$.
- **TPP prediction:** RMSE = 0.041 over range [0.6–4.1], with MAPE = 2.05%; $R^2 = 0.9858$.

- **N_{sa} prediction:** RMSE = 0.029, MAPE = 2.51%, R^2 = 0.9815, validating the engineered cubic geometry feature.
- **R_{th} prediction:** MAPE = 1.66%, R^2 = 0.9878, consistent with its inverse relationship with Nu .
- **θ prediction:** MAPE = 0.05%, R^2 = 0.9871, owing to the narrow physical range (97–99°C).

The PINN substantially outperforms conventional neural networks (typically R^2 = 0.85–0.90 on similar thermal datasets) by incorporating domain knowledge through physics constraints and PDE enforcement.

3.8 Prediction Analysis

Visual inspection of prediction scatter plots reveals tight clustering along the perfect prediction line for all five outputs. The scatter plots include colour-mapped absolute error to highlight regions of reduced model confidence. Isolated higher-error points occur at distribution extremes where limited training samples reduce model confidence, particularly at high Nu values (180–200 range).

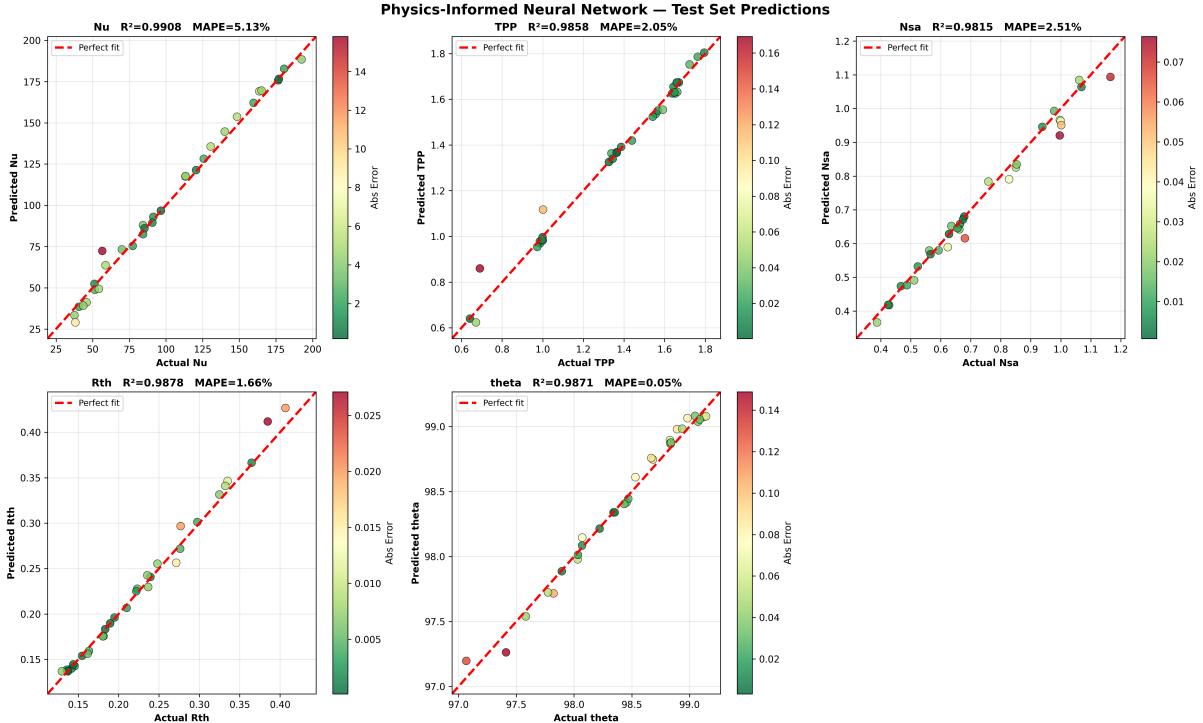


Figure 6: PINN test set predictions ($n = 32$) for all five output quantities (Nu , TPP, N_{sa} , R_{th} , θ). Each subplot shows actual vs. predicted values with the perfect-fit line (red dashed) and colour-mapped absolute error. R^2 and MAPE values are annotated per output.

3.9 Model Artefacts

Three files were generated for deployment and reproducibility, confirmed by the script output:

- **best_model.pth:** State dictionary from minimum validation loss epoch (best validation = 0.371 at approximately epoch 200).
- **final_pinn.pth:** Complete checkpoint including model state, scalers (`scaler_X`, `scaler_y`), metrics, training history, collocation bounds, and geometry encoding (`CWMCH: 0`, `DWMCH: 1`, `SWMCH: 2`).

- **pinn_training_results.png** and **pinn_testing_results.png**: Visualisation grids with loss curves and prediction scatter plots (300 DPI).

The saved artefacts enable model deployment for predicting thermal performance of novel heat sink geometries without additional numerical simulations. The geometry encoding stored in **final_pinn.pth** ensures correct input preparation during inference.

References

References

- [1] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019. DOI: [10.1016/j.jcp.2018.10.045](https://doi.org/10.1016/j.jcp.2018.10.045)
- [2] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, “Physics-informed machine learning,” *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, Jun. 2021. DOI: [10.1038/s42254-021-00314-5](https://doi.org/10.1038/s42254-021-00314-5)
- [3] S. Cai, Z. Wang, S. Wang, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks for heat transfer problems,” *ASME Journal of Heat Transfer*, vol. 143, no. 6, p. 060801, Jun. 2021. DOI: [10.1115/1.4050542](https://doi.org/10.1115/1.4050542)
- [4] R. L. Webb, “Performance evaluation criteria for use of enhanced heat transfer surfaces in heat exchanger design,” *International Journal of Heat and Mass Transfer*, vol. 24, no. 4, pp. 715–726, 1981. DOI: [10.1016/0017-9310\(81\)90015-6](https://doi.org/10.1016/0017-9310(81)90015-6)
- [5] A. Bejan, “Entropy generation minimization: The new thermodynamics of finite-size devices and finite-time processes,” *Journal of Applied Physics*, vol. 79, no. 3, pp. 1191–1218, 1996. DOI: [10.1063/1.362674](https://doi.org/10.1063/1.362674)
- [6] T. L. Bergman, A. S. Lavine, F. P. Incropera, and D. P. DeWitt, *Fundamentals of Heat and Mass Transfer*, 7th ed. Hoboken, NJ: John Wiley & Sons, 2011. ISBN: 978-0-470-50197-9
- [7] A. Kendall, Y. Gal, and R. Cipolla, “Multi-task learning using uncertainty to weigh losses for scene geometry and semantics,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, Jun. 2018, pp. 7482–7491. DOI: [10.1109/CVPR.2018.00781](https://doi.org/10.1109/CVPR.2018.00781)
- [8] I. Loshchilov and F. Hutter, “Decoupled weight decay regularization,” in *Proceedings of the 7th International Conference on Learning Representations (ICLR 2019)*, New Orleans, LA, USA, May 2019. Available: <https://openreview.net/forum?id=Bkg6RiCqY7>
- [9] L. Gong, K. Kota, W. Tao, and Y. Joshi, “Parametric numerical study of flow and heat transfer in microchannels with wavy walls,” *ASME Journal of Heat Transfer*, vol. 133, no. 5, pp. 1–10, 2011. DOI: [10.1115/1.4003284](https://doi.org/10.1115/1.4003284)
- [10] Y. Fan, C. Zhang, M. Li, and D. Zhu, “Optimal design of wavy microchannel heat sinks based on prediction and multi-objective optimisation algorithm,” *Energy and AI*, vol. 14, p. 100260, 2023. DOI: [10.1016/j.egyai.2023.100260](https://doi.org/10.1016/j.egyai.2023.100260)