Natural Resources for Human Health

eISSN: 2583-1194

Original Research

Received 24 July 2026
Revised 30 Aug 2026
Accepted 18 Sep 2026

Natural Resources for Human Health 2026, Vol.6(21s):1325–1347

KEYWORDS:
Computational fluid dynamics, machine learning, turbulent pipe flow, heat transfer, Nusselt number, Darcy friction factor, physics-informed learning, ANSYS Fluent, surrogate modelling, explainable artificial intelligence.

Hybrid Machine Learning and Computational Fluid Dynamics Framework for Accurate and Physics-Informed Prediction of Heat Transfer Performance

[1]Dr.Senthilkumar P, [2]Arunkumar S, [3]Dr. Palanisamy. R, [4]Dr. K. Parthiban, [5]Dr. Senthil Kumaar. S, [6]M.Amala Justus Selvam

[1]Associate Professor,Department of Mechanical Engineering,Meenakshi College of Engineering, KK Nagar, Chennai, kallakurichisenthilkumar@gmail.com
[2]Assistant professor (SG),Department of Mechatronics Engineering,Nehru Institute of Engineering and Technology, Coimbatore,nietarunkumars@nehrucolleges.com
[3]Assistant professor, Department of Mechanical Engineering,Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Avadi, Chennai - 600062 palani.hvac85@gmail.com
[4]Assistant professor, Department of Mechanical Engineering,Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Avadi, Chennai - 600062 dharaniparthiban2017@gmail.com
[5]Assistant professor, Department of Mechanical Engineering,Vel Tech High Tech Dr. Rangarajan Dr. Sakunthala Engineering College, Avadi, Chennai – 600062, senthiltcemech@gmail.com
[6]Professor, Department of Mechanical Engineering,Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, amalajustus@gmail.com

ABSTRACT: Heat transfer prediction in turbulent pipe flow is needed for heat exchangers, thermal management systems, industrial process equipment and energy conversion technologies design and operation. CFD methods with detailed information on velocity, pressure and temperature fields are available but expensive when many operating conditions have to be solved; empirical correlations are computationally cheap but lose accuracy outside their range of applicability; ML is an alternative for rapid prediction but models trained only on numerical data may reproduce CFD errors, lack extrapolation capability and violate physical relationships.

In this paper we propose a hybrid CFD–machine-learning framework for predicting the thermal-hydraulic performance of turbulent flow through smooth circular pipe. Three-dimensional or axisymmetric CFD simulations (ANSYS Fluent) will be performed with conservation equations for mass, momentum and energy and appropriate turbulence closure. A systematically designed dataset will be generated by changing Reynolds number, Prandtl number and thermal boundary conditions over an operating envelope. The main prediction targets are Nusselt number and Darcy friction factor. Artificial neural network, random forest, gradient boosting and support vector regression models will be compared to heat transfer and pressure loss correlations. The hybrid model will include dimensionless physical variables, baseline predictions based on correlations and physics consistency penalty. Coefficient of determination, root mean square error, mean absolute error, relative error, pressure drop consistency and independent validation cases will be used to evaluate performance. Feature attribution and uncertainty estimation will be used to assess interpretability and reliability respectively. The framework is intended

to provide reproducible pathway towards fast surrogate thermal hydraulic models keeping physical constraints and traceability to CFD predictions. Quantitative performance claims will only be made after simulation, validation and independent testing are completed.

NOMENCLATURE

Symbol	Definition	SI unit
\(D\)	Pipe internal diameter	m
\(L\)	Pipe length	m
\(u\)	Velocity component	m/s
\(U_b\)	Bulk mean velocity	m/s
\(\rho\)	Fluid density	kg/m³
\(\mu\)	Dynamic viscosity	Pa·s
\(k_f\)	Fluid thermal conductivity	W/(m·K)
\(c_p\)	Specific heat capacity	J/(kg·K)
\(T\)	Temperature	K
\(p\)	Static pressure	Pa
\(q"\)	Wall heat flux	W/m²
\(h\)	Convective heat-transfer coefficient	W/(m²·K)
\(Nu\)	Nusselt number	Dimensionless
\(Re\)	Reynolds number	Dimensionless
\(Pr\)	Prandtl number	Dimensionless
\(f_D\)	Darcy friction factor	Dimensionless
\(\Delta p\)	Pressure drop	Pa
\(R^2\)	Coefficient of determination	Dimensionless
RMSE	Root mean square error	Target-dependent
MAE	Mean absolute error	Target-dependent
SHAP	Shapley Additive Explanations	—

1. INTRODUCTION

1.1 Background

Internal turbulent flow is one of the most common configurations encountered in thermal engineering. Heat transfer between a flowing fluid and a pipe wall determines the performance of boilers, condensers, cooling circuits, industrial pipelines, compact heat exchangers and process-heating equipment. The thermal response is determined by fluid inertia viscous forces thermal diffusion turbulent mixing mechanisms which are often quantified using dimensionless parameters (Reynolds number Prandtl number).

The Reynolds number is the ratio of inertial to viscous forces and Prandtl number is the ratio of momentum diffusivity to thermal diffusivity. Increasing the Reynolds number generally increases turbulent transport in the operating range (see figure), but heat-transfer coefficient and pressure loss also depend on fluid properties, wall conditions and turbulence structure so simultaneous prediction of heat transfer and frictional resistance is necessary for engineering performance evaluation.

Conventional heat-transfer correlations are useful because they are simple and low cost to compute but are developed for specific flow regimes, geometries, boundary conditions and property ranges. They become inaccurate if assumptions are violated. Computational fluid dynamics overcomes some of these limitations by resolving spatial distributions of velocity, pressure and temperature (provided accuracy of numerical method, boundary conditions, mesh and turbulence model is good).

But CFD is expensive to run when large parametric studies or design optimisation are needed: every new combination of operating parameters will need a converged simulation. A machine-learning surrogate trained on carefully constructed CFD data can reduce the cost of repeated predictions but it should not be regarded as replacing physics or independently validated simply because it reproduces its training data.

Recent work shows both opportunities and limitations of this approach. Machine-learning methods have been used to predict local skin friction and Nusselt numbers in rough wall turbulent flows while in passive heat transfer enhancement neural networks and support vector methods have been compared against numerical and correlation predictions. General reviews on physics informed machine learning highlight physical consistency, generalization and reliable data provenance.

1.2 Research problem

The central research problem is to build a predictive model that combines the detail of CFD with speed of machine learning without letting the surrogate break any thermal-fluid relationships.

A CFD-only workflow might be accurate in the validated operating envelope but slow for large parameter exploration. A conventional correlation is cheap but not necessarily representative of every combination of operating conditions. A purely data-driven model can learn nonlinear dependencies but may make implausible predictions outside its training distribution.

The study proposed here therefore considers a hybrid framework where CFD provides the training data, empirical correlations give physically informed baseline predictions and machine learning learns the residual error between CFD results and correlation-based estimates. This design allows contribution of machine learning to be assessed against transparent engineering baseline.

1.3 Research gap

First, predictive accuracy is not sufficient to show that there is a useful thermal surrogate. A model must also preserve physically plausible trends (positive heat-transfer coefficients and non-negative friction factors) and be consistent with the range of operating conditions represented by the training data set.

Second, random train-test splitting can overgeneralize when simulations from nearly identical operating conditions are in both sets. An independent test strategy based on operating-condition groups therefore has to be used.

Third, heat transfer and pressure loss are coupled engineering outcomes. A model that predicts Nusselt number accurately but provides an unreliable friction factor may not be suitable for practical thermal-hydraulic design.

1.4 Original contribution and novelty

What we have proposed is the combination of correlation-residual learning, CFD verification, grouped validation and simultaneous thermal-hydraulic assessment within a reproducible smooth pipe benchmark.

The individual methods are established techniques so novelty cannot be claimed by just using ML or SHAP alone. Strength will come from showing that the hybrid model outperforms appropriate baselines on truly independent operating conditions while being physically consistent.

The manuscript is structured as follows. Section 2 reviews the literature. In section 3 we present the physical and mathematical formulation. In section 4 we describe the CFD setup and numerical validation. In section 5 we develop the hybrid prediction framework. In section 6 we specify the validation metrics and statistical comparisons. In section 7 we define

results and discussion protocol. In section 8 limitations and engineering implications are discussed and conclusions and research declarations follow

2. LITERATURE REVIEW

2.1 Conventional correlations for internal heat transfer:

Empirical and semi-empirical correlations are still the first point of comparison for pipe flow prediction. Dittus–Boelter correlation is used extensively for fully developed turbulent heat transfer in smooth tubes (appropriate property and operating conditions). Gnielinski correlation provides a more general engineering approach for turbulent internal flow and incorporates friction-factor information.

These correlations are useful baselines because they state assumptions and intended operating ranges. But a correlation should not be treated as being universally valid — suitability depends on Reynolds number, Prandtl number, thermal development, surface condition and choice of fluid properties. Comparisons with CFD and machine learning therefore need consistent definitions and operating conditions.

For frictional pressure loss the Darcy–Weisbach equation gives a standard relationship between pressure drop and Darcy friction factor. For a smooth pipe in turbulent flow the friction factor can be estimated from an appropriate smooth-wall correlation (Petukhov relation in this case) within its range of applicability.

The correlations used in the present study will be implemented independently of the machine-learning models and evaluated against the same held out test cases.

2.2 CFD modelling of turbulent pipe flow

CFD solves discretized conservation equations to obtain flow and thermal fields. For engineering pipe-flow simulations Reynolds-averaged Navier–Stokes (RANS) models are usually used since they offer a practical balance between computational cost and prediction capability. The realizable (k)-(\varepsilon) model and the shear-stress transport (k)-(\omega) model are good candidates for comparison but relative performance must be evaluated for the particular geometry and boundary conditions.

ANSYS Fluent contains the conservation equations, turbulence model formulations, energy transport and near-wall treatments for this class of problem.

The main numerical uncertainties are mesh resolution, discretization error, convergence tolerances, inlet development, turbulence closure and thermal-boundary-condition implementation. Residual convergence alone is therefore not enough: bulk temperature, pressure drop, energy balance and integral heat transfer quantities must also stabilize.

For a smooth circular pipe, a fully developed reference case is very useful since established correlations allow independent checks on heat transfer and friction. The model must reproduce the assumptions of the reference correlation before agreement can be meaningful.

2.3 Machine learning for thermal prediction

Machine-learning models approximate nonlinear relationships between operating parameters and thermal-hydraulic outputs. ANN models can model complex interactions but need careful regularization and tuning; random forest models are robust for nonlinear tabular datasets and provide a useful baseline for an ensemble; gradient-boosting methods can represent nonlinear interactions with small datasets; support-vector regression can work well when feature scaling and kernel selection are correct.

A relevant 2023 study Diez Sanhueza, Akkerman and Peeters studied ML prediction of local skin-friction factors and Nusselt numbers in turbulent flow past rough surfaces. It showed that data driven surrogates can reproduce spatially varying thermal and hydrodynamic behaviour but also that results for rough surfaces should not be automatically generalised to smooth pipes.

Yapıcı, Aylı and Türkoğlu studied machine-learning methods for passive heat-transfer enhancement in dimpled tubes. Numerical predictions vs statistical approaches (ANN and support-vector methods) for Nusselt number and friction-factor prediction were compared. Since geometry and fluid configurations are different from a smooth-pipe benchmark their results motivate the method selection but do not provide results for present study.

These studies show that CFD-derived thermal surrogates are feasible but model performance depends on range and distribution of training cases, data quality and validation protocol.

2.4 Physics-informed and hybrid learning

Physics-informed machine learning can be implemented in several ways. In physics-informed neural networks the governing equations enter directly into the optimization loss; in a hybrid surrogate physical knowledge is instead introduced through dimensionless features, baseline correlations known from data, conservation constraints or residual learning.

For the present study we use correlation-residual learning since it is relatively easy to implement and evaluate with a tabular CFD dataset. Instead of predicting Nusselt number or friction factor from scratch, the model learns what is missing between CFD quantities and correlations.

This formulation has two advantages. First the baseline correlation models part of what is known about dependence on dimensionless parameters; second ML model learns systematic residual behaviour rather than learning the whole relationship separately. Advantage is not a guarantee improvement!

Physics-informed modelling of coupled flow and heat transfer research shows how physical constraints can improve interpretability and lead to route towards more reliable prediction but also challenges in optimisation and generalization.

2.5 Explainability and uncertainty

Feature attribution methods (e.g., SHAP) can explain how input variables influence individual predictions and overall model behaviour. In pipe flow the Reynolds number, Prandtl number and baseline correlation estimates should be important candidate features but their actual importance needs to be determined from the trained model rather than assumed beforehand.

Uncertainty should also be taken into account. Sources include CFD discretization, turbulence model limitations, input property uncertainty and surrogate prediction error. Bootstrap ensembles can estimate variability associated with model fitting but they do not necessarily capture all forms of physical or numerical uncertainty.

2.6 Synthesis of the literature

The literature indicates that CFD-derived ML models can reduce prediction cost but several methodological requirements remain: transparent data generation, comparison with correlations, independent testing, physical plausibility and reporting of limitations. The proposed study brings these requirements together for a smooth-pipe turbulent-flow benchmark.

Literature scope note: The references below combine basic correlations and numerical methods with recent work in peer review. In a final journal submission you will have done the database search, verified bibliographic details and additional recent studies relevant to smooth-pipe turbulent heat transfer.

3. PHYSICAL MODEL AND GOVERNING EQUATIONS

3.1 Computational domain

The proposed physical system is a straight, smooth, circular pipe carrying a single-phase Newtonian fluid under forced convection. The pipe has a constant internal diameter and no internal inserts, wall roughness, chemical reaction, phase change, or buoyancy-driven flow. These restrictions establish a controlled baseline for investigating the predictive capability of the hybrid framework.

Natural Resources for Human Health
eISSN: 2583-1194

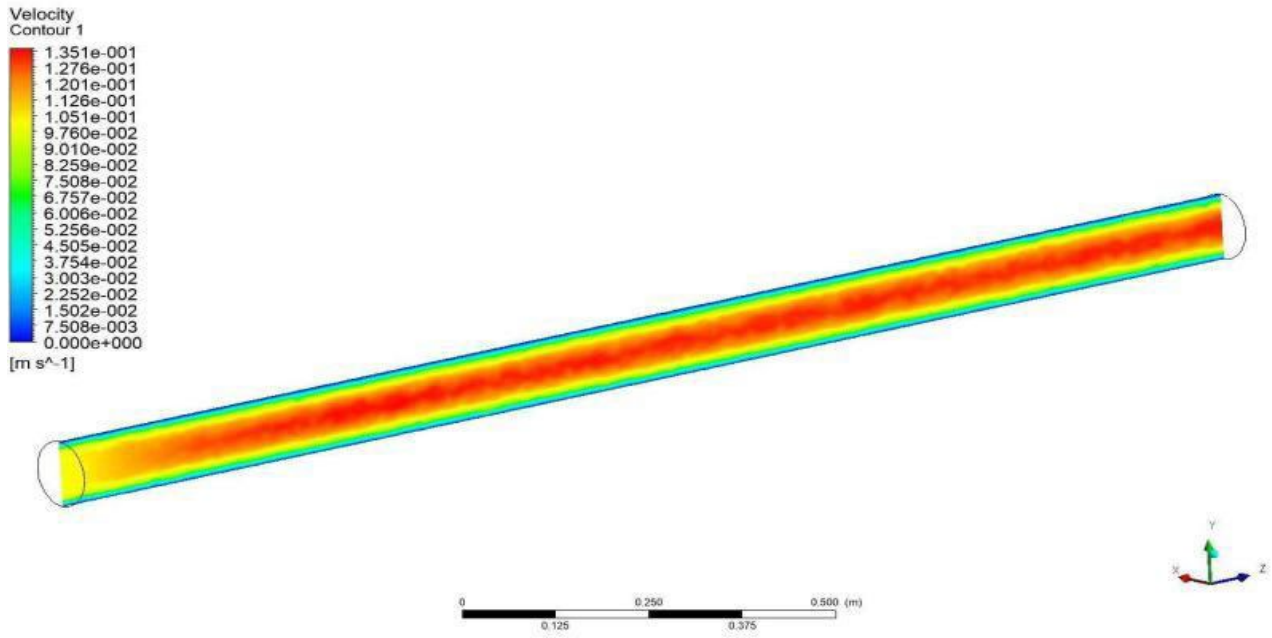


Figure 1. Proposed schematic of the heated pipe and CFD domain. The final manuscript should replace this illustrative reference with a dimensioned diagram showing the inlet, hydrodynamic development region, thermal development region, heated wall, and outlet.

The initial study will use a smooth pipe with a constant wall heat flux. A constant-wall-temperature case may subsequently be included as an independent boundary-condition test. The main study will focus on a specified fluid-property range, with properties evaluated consistently at the selected reference temperature or represented using temperature-dependent material models where appropriate.

3.2 Conservation of mass

For steady, incompressible flow without mass sources, the continuity equation is

$$\nabla\cdot\mathbf{u}=0 \tag{1}$$

where $\mathbf{u}$ denotes the velocity vector.

This equation ensures that the net mass flux through a control volume is zero. In Fluent, the mass imbalance should be monitored as part of the solution verification process.

3.3 Conservation of momentum

The steady incompressible momentum equation is

$$\rho(\mathbf{u}\cdot\nabla)\mathbf{u} =-\nabla p+\nabla\cdot\boldsymbol{\tau} \tag{2}$$

where p is static pressure and $\boldsymbol{\tau}$ is the viscous stress tensor. For a Newtonian fluid with constant viscosity,

$$\boldsymbol{\tau} =\mu\left[\nabla\mathbf{u} +(\nabla\mathbf{u})^{T}\right] \tag{3}$$

In a RANS simulation, the time-averaged momentum equation contains additional Reynolds-stress terms associated with turbulent velocity fluctuations. These terms are represented using the selected turbulence closure.

3.4 Conservation of energy

For steady forced convection with constant properties and negligible viscous dissipation, the temperature equation can be written as

$$\rho c_p(\mathbf{u}\cdot\nabla T) =\nabla\cdot(k_f\nabla T) \tag{4}$$

For turbulent flow, the effective thermal diffusivity also incorporates turbulent heat transport. A simplified RANS form is

$$\rho c_p(\mathbf{u}\cdot\nabla T) =\nabla\cdot\left[\left(k_f+\frac{\mu_t c_p}{Pr_t}\right)\nabla T \right] \tag{5}$$

where μ_t is turbulent viscosity and Pr_t is the turbulent Prandtl number. Equation (5) is a simplified constant-property representation; the full energy formulation used by Fluent includes the applicable enthalpy and transport terms.

The energy equation will be activated in Fluent. The predicted temperature field will be used to calculate bulk temperature, heat-transfer coefficient, and Nusselt number.

3.5 Dimensionless parameters

The principal hydrodynamic parameter is the Reynolds number:

$$Re=\frac{\rho U_bD}{\mu} \tag{6}$$

The Prandtl number is

$$Pr=\frac{\mu c_p}{k_f} \tag{7}$$

The Nusselt number is defined as

$$Nu=\frac{hD}{k_f} \tag{8}$$

For a constant-wall-heat-flux condition, the local heat-transfer coefficient can be calculated from

$$h(x)=\frac{q"_w}{T_w(x)-T_b(x)} \tag{9}$$

where T_w is the wall temperature and T_b is the bulk fluid temperature. A mass-flow-weighted bulk temperature should be used rather than an arbitrary area-averaged fluid temperature.

The bulk temperature can be expressed as

$$T_b(x)= \frac{\int_A \rho u_x c_p T\,dA} {\int_A \rho u_x c_p\,dA} \tag{10}$$

for a cross-section with positive axial flow and constant specific heat. The corresponding expression can be simplified when the density and specific heat are uniform across the section.

The Darcy friction factor is calculated using

$$f_D=\frac{2D\Delta p}{L\rho U_b^2} \tag{11}$$

for a fully developed section in which the measured pressure drop is dominated by distributed friction. If entrance and exit losses are included in Δp, Equation (11) will not represent the fully developed Darcy friction factor without correcting for those losses.

3.6 Thermal-hydraulic performance

The study will predict Nu and f_D separately. A combined performance metric may be reported only when the comparison is defined consistently. For example, a commonly used heat-transfer performance evaluation criterion at equal pumping power is

$$PEC=\frac{Nu/Nu_0}{(f_D/f_{D,0})^{1/3}} \tag{12}$$

where the subscript 0 denotes a reference configuration.

Because the proposed primary geometry is a smooth pipe, this criterion is not the principal novelty metric. It becomes more meaningful if enhanced and reference geometries are later compared under a clearly specified operating constraint.

Natural Resources for Human Health
eISSN: 2583-1194

4. CFD METHODOLOGY USING ANSYS FLUENT

4.1 Proposed numerical configuration

The initial CFD model will be developed in ANSYS Fluent. A three-dimensional circular pipe is preferred for the primary study because it allows the geometry, boundary conditions, and mesh to be described in a conventional CFD workflow. An axisymmetric model may be used for preliminary verification when the flow and thermal conditions are axisymmetric.

Table 1. Proposed CFD setup

Parameter	Proposed setting
Geometry	Straight, smooth, circular pipe
Internal diameter	20 mm (proposed baseline)
Pipe length	100D initially; extend if development or outlet-independence checks require it
Working fluid	Single-phase Newtonian fluid; use one documented property set for the initial benchmark
Flow regime	Turbulent, with baseline \(Re=5{,}000\) to \(100{,}000\) proposed for investigation
Thermal boundary condition	Uniform wall heat flux
Inlet	Velocity inlet with consistent turbulence specifications and thermal properties
Outlet	Pressure outlet
Wall	No-slip, smooth, impermeable, heated
Solver	Pressure-based, steady RANS for the initial benchmark
Energy equation	Enabled
Turbulence models	Realizable \(k\)-\(\varepsilon\) and SST \(k\)-\(\omega\), compared during verification

These values define a proposed starting configuration, not a validated final setup. The pipe length, turbulence model, mesh density, and inlet treatment must be verified before the generated data are used for machine learning.

The initial Reynolds-number range excludes the transitional regime as a research focus, but the lower end of a nominal turbulent range can remain sensitive to inlet disturbances and modelling assumptions. The selected operating cases must be checked for consistency with the intended turbulence regime.

4.2 Geometry and mesh generation

The pipe will be drawn in CAD or meshing environment and imported into Fluent. Structured hexahedral mesh (if possible) should dominate the mesh where practical. Resolution across the radius of the pipe is required (near wall velocity and temperature gradients are large).

Three systematically refined meshes will be produced. Their cell counts and near-wall spacing will be documented. The final mesh will be chosen based on changes in predicted average Nusselt number, Darcy friction factor and pressure drop rather than cell count alone.

For a low Reynolds number near-wall treatment the first cell centre position should generally target y + near unity. Wall function formulations require corresponding wall-treatment guidance and an appropriate near wall mesh. The target must match the selected turbulence model and Fluent wall treatment (the two strategies should not be mixed indiscriminately).

Natural Resources for Human Health
eISSN: 2583-1194

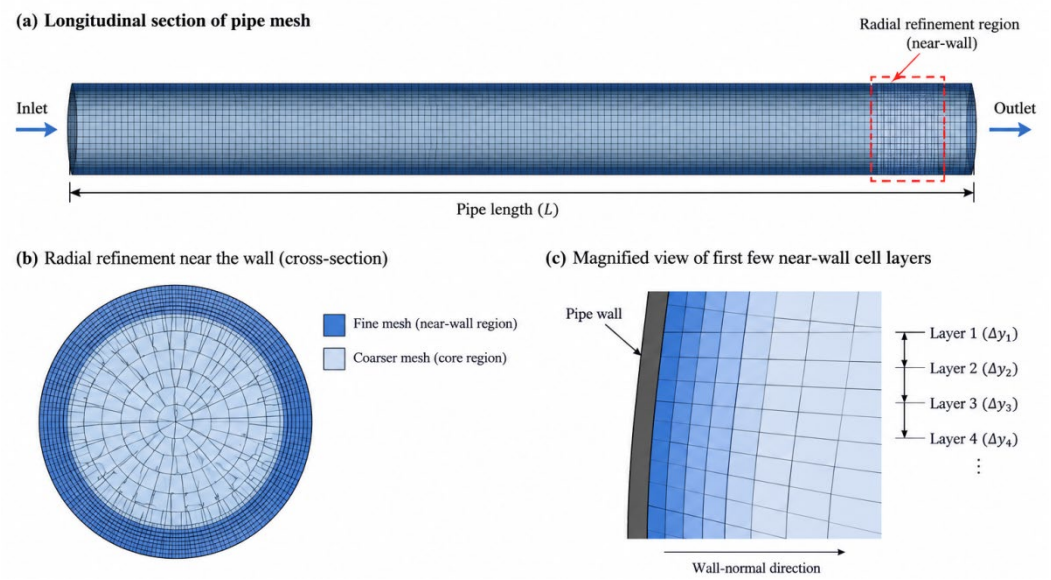


Figure2. ProposedCFD mesh. The final figure should show a longitudinal section, radial refinement near the wall, and a magnified view of the first few near-wall cell layers.

4.3 Boundary conditions

At the inlet, the mean velocity will be selected to give the required Reynolds number. The turbulence intensity and length scale (or equivalent quantities of turbulence) at the inlet must be specified consistently and reported; if a fully developed inlet profile is used its derivation or source must be documented.

At the outlet, a pressure-outlet condition will be applied sufficiently downstream of the region used to calculate thermal-hydraulic outputs. Backflow thermal conditions will be specified where required by the solver.

The pipe wall will have a no-slip condition and a prescribed uniform heat flux. The heat flux should be chosen so that temperature rise is measurable but not so large that it causes an unintended change in the fluid properties which invalidates the constant-property assumption.

4.4 Numerical schemes and convergence

The pressure–velocity coupling method, gradient reconstruction, and spatial discretization schemes will be selected and reported explicitly. Second-order spatial schemes are proposed for the principal calculations, provided that convergence can be achieved reliably.

The following criteria will be used as initial convergence requirements:

- Scaled residuals for continuity, momentum, and turbulence equations should generally fall below \(10^{-5}\), with an appropriate energy residual target such as \(10^{-8}\).
- The outlet mass flow should agree with the inlet mass flow within a predefined tolerance.
- The calculated heat balance should be consistent with the net energy added to the fluid.
- Nusselt number, pressure drop, and friction factor should stabilize over successive iterations.

Residual thresholds are necessary but not sufficient. A solution will not be considered converged if the integral outputs continue to drift or if mass and energy conservation are not satisfied.

4.5 Mesh-independence study

The grid-convergence assessment will use at least three systematically refined meshes with a documented refinement ratio. For a quantity of interest \(\phi\), the relative change between successive meshes can be evaluated as

[\varepsilon_{\phi}=\left|\frac{\phi_{\mathrm{fine}}-\phi_{\mathrm{medium}}}{\phi_{\mathrm{fine}}}\right|\times100\% \tag{13} \]

where \(\phi\) may represent \(Nu\), \(f_D\), or \(\Delta p\).

Equation (13) provides a relative difference, not a complete discretization-uncertainty estimate. Where the mesh sequence and convergence behaviour support it, Richardson extrapolation and the Grid Convergence Index should also be used.

Table 2. Template for mesh-independence results

Mesh	**Cells**	**Mean \(Nu\)**	**\(f_D\)**	**\(\Delta p\) (Pa)**
Coarse	To be measured	—	—	—
Medium	To be measured	—	—	—
Fine	To be measured	—	—	—
Refinement difference	Calculated	—	—	—

All results are intentionally left blank pending actual Fluent simulations.

4.6 CFD validation against reference correlations

The CFD model will be compared with existing smooth-pipe correlations for the selected operating range (e.g., Gnielinski correlation for heat-transfer comparison or suitable smooth pipe friction correlation for pressure loss verification).

The validation protocol must ensure that the correlation and CFD model use the same definitions of Reynolds number, Prandtl number, Nusselt number and Darcy friction factor. Fluid properties should be evaluated consistently. Comparisons will be limited to the applicability range published for each correlation.

Agreement with an empirical correlation is one benchmark but not the other. If there are no independent experimental data then the manuscript should describe the CFD as numerically verified and benchmarked against correlations rather than experimentally validated.

4.7 Data extraction

For each converged simulation, the following quantities will be recorded:

1. Reynolds number and Prandtl number.
2. Inlet and outlet bulk temperatures.
3. Wall heat flux and wall temperature distribution.
4. Pressure drop over the selected fully developed section.
5. Mean Nusselt number and Darcy friction factor.
6. Mesh details, turbulence model, residual histories, and conservation errors.
7. Reference-correlation predictions and their relative deviations from CFD.

The final dataset will contain one clearly defined row per independent CFD operating case. Multiple axial stations from a single simulation may be retained for local-field modelling, but they must not be treated as independent cases in a random split without accounting for their shared simulation origin.

Natural Resources for Human Health
eISSN: 2583-1194

5. PROPOSED CFD DATASET AND EXPERIMENTAL DESIGN

5.1 Input parameters and targets

The initial dataset will be based on dimensionless operating variables. For the constant-property baseline, the core inputs will be Reynolds number, Prandtl number, and the selected correlation-based estimates. This makes the model more transferable across geometrically similar cases than using raw velocity and diameter alone.

Table 3. Proposed dataset variables

Variable	Role	Status
\(Re\)	Input	Proposed
\(Pr\)	Input	Proposed
\(Nu_{\mathrm{corr}}\)	Physics-informed baseline	Calculated from a documented correlation
\(f_{\mathrm{corr}}\)	Physics-informed baseline	Calculated from a documented correlation
\(Nu_{\mathrm{CFD}}\)	Target	To be obtained from Fluent
\(f_{D,\mathrm{CFD}}\)	Target	To be obtained from Fluent
\(\Delta p_{\mathrm{CFD}}\)	Verification output	To be obtained from Fluent
Mesh uncertainty	Quality indicator	To be estimated
Turbulence model	Grouping variable	Recorded for every case

The baseline study should keep geometry, surface condition, and thermal boundary condition fixed. If fluid properties are varied, the property model must be documented and the variation must remain consistent with the intended dimensionless parameter range.

5.2 Operating envelope

The proposed initial ranges are:

- Reynolds number: \(5{,}000\le Re\le100{,}000\).
- Prandtl number: \(0.7\le Pr\le7\).

These ranges are design choices for the proposed study, not claims about completed CFD cases. A valid sampling strategy must ensure that the chosen fluid properties can actually realize the intended dimensionless combinations.

A Latin hypercube or space-filling design may be used to select initial operating points. Additional simulations can be added where the surrogate exhibits high uncertainty or large errors. A pilot dataset should first establish the computational cost, convergence reliability, and sensitivity of the outputs.

5.3 Dataset size and split strategy

A fixed number of CFD cases should not be selected just to get a large-looking dataset. The number needed depends on complexity of response surface, operating range and accuracy desired.

A pilot of about 30–50 operating points may be sufficient to prove feasibility. The final dataset size should be determined by sequential sampling and learning-curve analysis. These counts are planning estimates, not simulations completed or statistical sufficiency evidence.

The dataset will be split into training, validation and test subsets (70/15/15% allocation is proposed but grouping is more important than the exact percentages). Operating-condition groups should be excluded so that near-duplicate cases don't leak across subsets. A separate Reynolds number interval or carefully selected boundary region may be reserved for extrapolation assessment.

The test set must not be modified during feature selection, hyperparameter tuning and model selection.

6. HYBRID MACHINE-LEARNING FRAMEWORK

6.1 Framework architecture

The proposed framework consists of five linked stages: CFD data generation, physical feature construction, baseline prediction, residual learning, and independent validation.

Stage 1 — ANSYS Fluent CFD

Mass, momentum and energy equations; mesh verification; convergence checks

Stage 2 — Dataset and feature engineering

\(Re\), \(Pr\), correlation estimates, target quantities and simulation-quality metadata

Stage 3 — Correlation-residual learning

ANN, random forest, gradient boosting and SVR

Stage 4 — Physics and uncertainty checks

Positive outputs, residual diagnostics, bootstrap uncertainty and SHAP analysis

Stage 5 — Independent evaluation

Held-out CFD cases, reference correlations and prediction-cost comparison

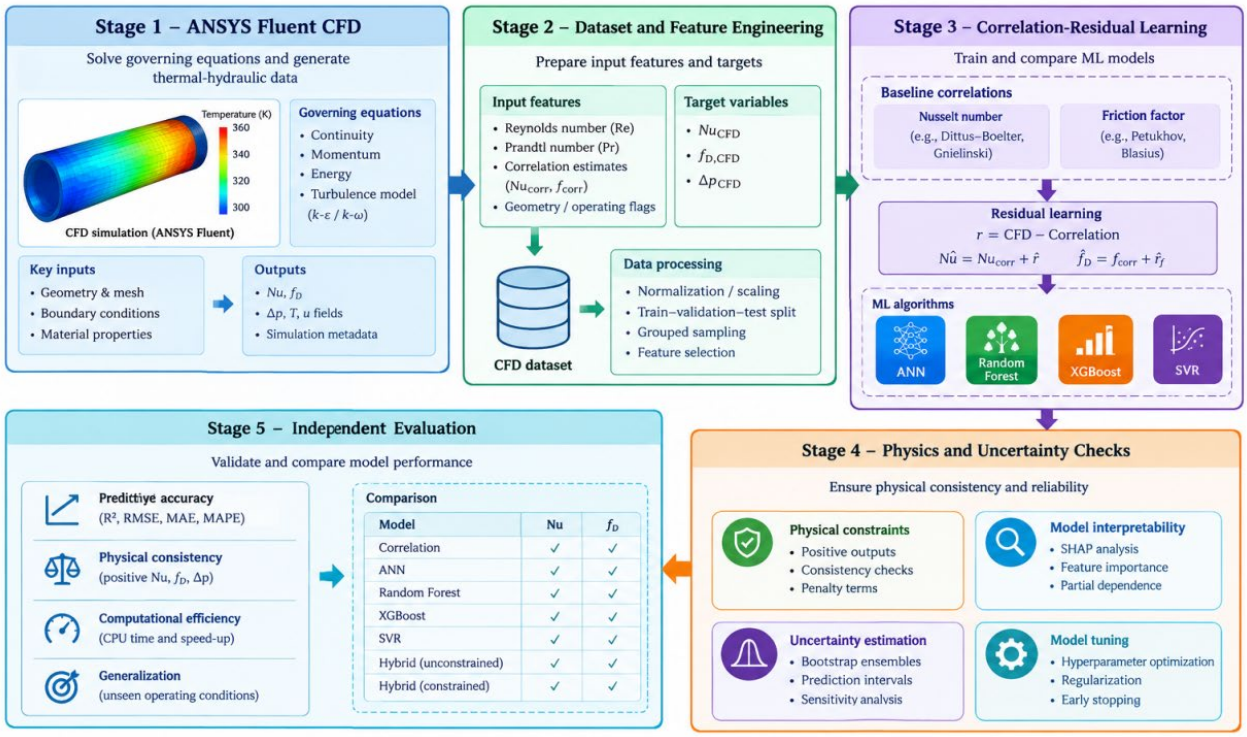


Figure3. Proposed hybrid CFD–ML architecture.This workflow is the planned methodology, not a record of completed computations

6.2 Baseline correlation models

Two baseline predictions will be calculated independently of machine learning:

\[Nu_{\mathrm{base}}=g_{Nu}(Re,Pr) \tag{14} \]

$$ f_{\mathrm{base}}=g_f(Re,\epsilon/D) \tag{15} $$

where g_{Nu} and g_f are selected correlations, and ϵ/D is relative roughness. For the primary smooth-pipe study, $\epsilon/D=0$, subject to the chosen correlation's definitions.

For example, the Dittus–Boelter relation can be considered for appropriate fully developed turbulent conditions:

$$ Nu=0.023Re^{0.8}Pr^n \tag{16} $$

where n depends on whether the fluid is being heated or cooled under the assumptions of the correlation. This relation is a benchmark candidate rather than a universal formula.

The Gnielinski correlation is another candidate for Nu_{base}, while an appropriate smooth-pipe friction correlation will be used for f_{base}. The final correlation selection and valid ranges must be documented before model fitting.

6.3 Residual-learning formulation

The proposed hybrid model learns the difference between the CFD result and the correlation prediction. For Nusselt number,

$$ r_{Nu}=Nu_{\mathrm{CFD}}-Nu_{\mathrm{base}} \tag{17} $$

The hybrid prediction is

$$ \widehat{Nu}_{\mathrm{hyb}} =Nu_{\mathrm{base}}+ \widehat{r}_{Nu} \tag{18} $$

Likewise, for the Darcy friction factor,

$$ r_f=f_{D,\mathrm{CFD}}-f_{\mathrm{base}} \tag{19} $$

$$ \widehat{f}_{D,\mathrm{hyb}} =f_{\mathrm{base}}+\widehat{r}_f \tag{20} $$

Here, $\widehat{r}_{Nu}$ and $\widehat{r}_f$ are predicted residuals from the selected machine-learning model.

Residual learning can be effective when the baseline captures the dominant physical dependence. However, it may be less advantageous if the baseline is unsuitable for the data or the residuals are highly irregular. Both direct prediction and residual learning must therefore be evaluated under the same validation conditions.

6.4 Candidate algorithms

Artificial neural network (ANN)

Learns nonlinear interactions between dimensionless operating parameters and correlation residuals. Hidden-layer size, activation functions, regularization, and learning rate will be tuned using training and validation data.

Random forest regression (RF)

Provides a nonlinear ensemble baseline with relatively modest preprocessing requirements. Tree depth, number of trees, and minimum leaf size will be tuned.

Gradient boosting / XGBoost

Learns nonlinear response surfaces through sequential tree fitting. Learning rate, tree depth, number of estimators, and regularization will be selected without using the held-out test set.

Support-vector regression (SVR)

Uses a nonlinear kernel to approximate the residual surface. Input scaling and kernel hyperparameters will be tuned within the training process.

6.5 Physics-consistency constraints

A basic hybrid model will be compared with a constrained version. The constrained model will enforce physically valid output ranges and use a penalty for violations of specified physical conditions.

A general training objective is

$$\mathcal{L}=\mathcal{L}_{\mathrm{data}}+\lambda_p\mathcal{L}_{\mathrm{physics}}+\lambda_r\mathcal{L}_{\mathrm{reg}} \tag{21}$$

where:

- $\mathcal{L}_{\mathrm{data}}$ measures prediction error against CFD targets.
- $\mathcal{L}_{\mathrm{physics}}$ penalizes predefined physical violations.
- $\mathcal{L}_{\mathrm{reg}}$ penalizes excessive model complexity.
- λ_p and λ_r are non-negative weighting parameters.

For example, a positivity penalty for the Nusselt number may be expressed as

$$\mathcal{L}_{+}=\frac{1}{N}\sum_{i=1}^{N}\left[\max(0,-\widehat{Nu}_i)\right]^2 \tag{22}$$

An equivalent penalty can be applied to the Darcy friction factor. Positive-output transformations, such as modelling logarithmic targets, may be preferable to penalties when positivity is required by construction.

Monotonicity constraints must be applied carefully. For example, although heat transfer generally increases with Reynolds number in a fixed turbulent regime, the study should not impose a universal monotonic rule across changes in geometry, boundary conditions, fluid properties, or flow regime without justification.

6.6 Training and hyperparameter optimization

All feature scaling, feature selection, and data transformations will be fitted using training data only. Validation data will be used for model selection and hyperparameter optimization. The final test set will be used once for the principal performance assessment.

A nested cross-validation approach may be adopted if the dataset is sufficiently large. If the number of independent CFD cases is limited, repeated grouped cross-validation can provide a more informative assessment than a single random split.

The study will compare the following four approaches:

1. Conventional correlations.
2. Direct ML prediction.
3. Correlation-residual hybrid learning.
4. Hybrid learning with explicit physical constraints.

This comparison is necessary to identify whether the additional complexity of the hybrid framework provides a measurable benefit.

6.7 Explainable machine learning

SHAP analysis will be applied to the final selected model. The global analysis will summarize the contribution of each feature to model predictions, while local explanations will investigate individual cases with large errors.

Natural Resources for Human Health
eISSN: 2583-1194

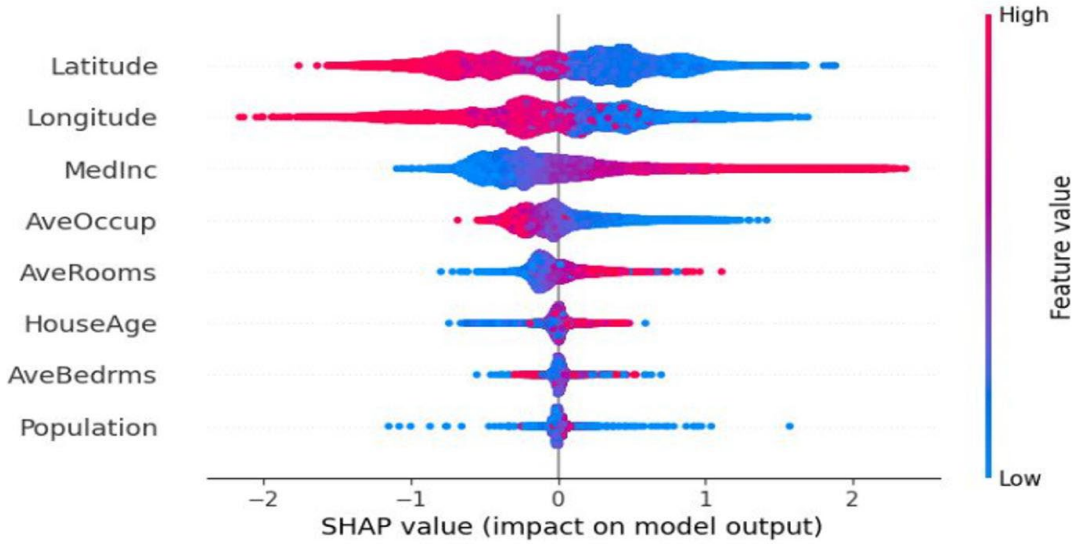


Figure 4. Illustrative reference for the proposed SHAP visualization. The final manuscript must contain plots generated from the trained model and the actual CFD dataset. No feature-importance ranking is asserted at this stage.

Because Reynolds number, Prandtl number, and correlation-derived features may be correlated, SHAP results will be interpreted as model explanations rather than proof of causal physical mechanisms.

7. Model-validation strategy and performance metrics

7.1 Coefficient of determination

The coefficient of determination will be calculated as

$$R^2=1- \frac{\sum_{i=1}^{N}(y_i-\hat y_i)^2} {\sum_{i=1}^{N}(y_i-\bar y)^2} \tag{23}$$

where y_i is the reference CFD value, $\hat y_i$ is the model prediction, and $\bar y$ is the mean reference value.

A high R^2 indicates that the model explains a substantial proportion of the observed variation. It does not by itself establish low absolute error, unbiased prediction, or reliable extrapolation.

7.2 Root mean square error

$$RMSE= \sqrt{\frac{1}{N}\sum_{i=1}^{N} (y_i-\hat y_i)^2} \tag{24}$$

RMSE penalizes large prediction errors more strongly than MAE. Its units are the same as those of the predicted variable.

7.3 Mean absolute error

$$MAE=\frac{1}{N}\sum_{i=1}^{N}|y_i-\hat y_i| \tag{25}$$

MAE provides an interpretable average absolute deviation and is less sensitive to large errors than RMSE.

7.4 Mean absolute percentage error

For nonzero reference values,

$$MAPE=\frac{100}{N}\sum_{i=1}^{N} \left|\frac{y_i-\hat y_i}{y_i}\right| \tag{26}$$

MAPE will be reported cautiously because percentage errors become unstable when the reference value approaches zero. Relative errors will also be inspected individually.

7.5 Bias and physical consistency

Mean prediction bias will be calculated as

$$Bias=\frac{1}{N}\sum_{i=1}^{N}(\hat y_i-y_i) \tag{27}$$

Positive bias indicates systematic overprediction, whereas negative bias indicates systematic underprediction. Residual plots will be examined for systematic patterns with Reynolds number, Prandtl number, and the magnitude of the target variable.

The hybrid model will also be checked for:

- Positive Nusselt number and Darcy friction factor.
- Consistency between predicted friction factor and pressure-drop calculations.
- Stable predictions within the validated operating envelope.
- Sensitivity to mesh-dependent or turbulence-model-dependent CFD targets.
- Reliable uncertainty estimates for unseen operating cases.

7.6 Uncertainty assessment

Bootstrap ensembles or repeated grouped fits will be used to estimate the variability of model predictions. Where feasible, a held-out calibration set can be used to construct prediction intervals.

These intervals describe uncertainty under the assumptions of the fitted model and calibration procedure. They do not automatically include CFD model-form error or all sources of numerical uncertainty. Those contributions will be assessed separately through turbulence-model comparison, mesh refinement, and correlation benchmarking.

7.7 Computational-efficiency assessment

The inference cost of each trained model will be measured and compared with the wall-clock time required for a converged CFD simulation under documented hardware and solver settings.

The speed-up ratio can be defined as

$$S=\frac{t_{\mathrm{CFD}}}{t_{\mathrm{ML}}} \tag{28}$$

where t_{CFD} is the time required for the specified CFD task and t_{ML} is the prediction time for the corresponding surrogate task. Training time and preprocessing cost will be reported separately.

A large value of S is meaningful only when the prediction accuracy and intended use are stated. Surrogate inference is not equivalent to resolving the complete flow field with CFD.

Table 4. Proposed model-comparison matrix

Model	**Nu prediction**	**f_D prediction**	**Physical constraints**
Gnielinski / selected correlation	Yes	Separate friction correlation	Embedded in correlation assumptions
ANN	Yes	Yes	Baseline unconstrained
Random forest	Yes	Yes	Baseline unconstrained
XGBoost	Yes	Yes	Baseline unconstrained
SVR	Yes	Yes	Baseline unconstrained
Correlation-residual hybrid	Yes	Yes	Correlation baseline
Constrained hybrid	Yes	Yes	Explicit output/consistency constraints

All performance cells are intentionally qualitative: quantitative metrics require completed simulations and independent testing.

Natural Resources for Human Health
eISSN: 2583-1194

8. RESULTS AND DISCUSSION: PLANNED ANALYSIS

This section is a reporting framework, not a claim that results have already been obtained. The final manuscript should replace the placeholders with actual CFD output, model predictions, and validation statistics.

8.1 CFD solution verification

The first stage will establish that the numerical model is sufficiently reliable to serve as a training-data source. The mesh-independence results will be reported alongside mass conservation, energy balance, residual histories, and the stability of the integrated thermal-hydraulic outputs.

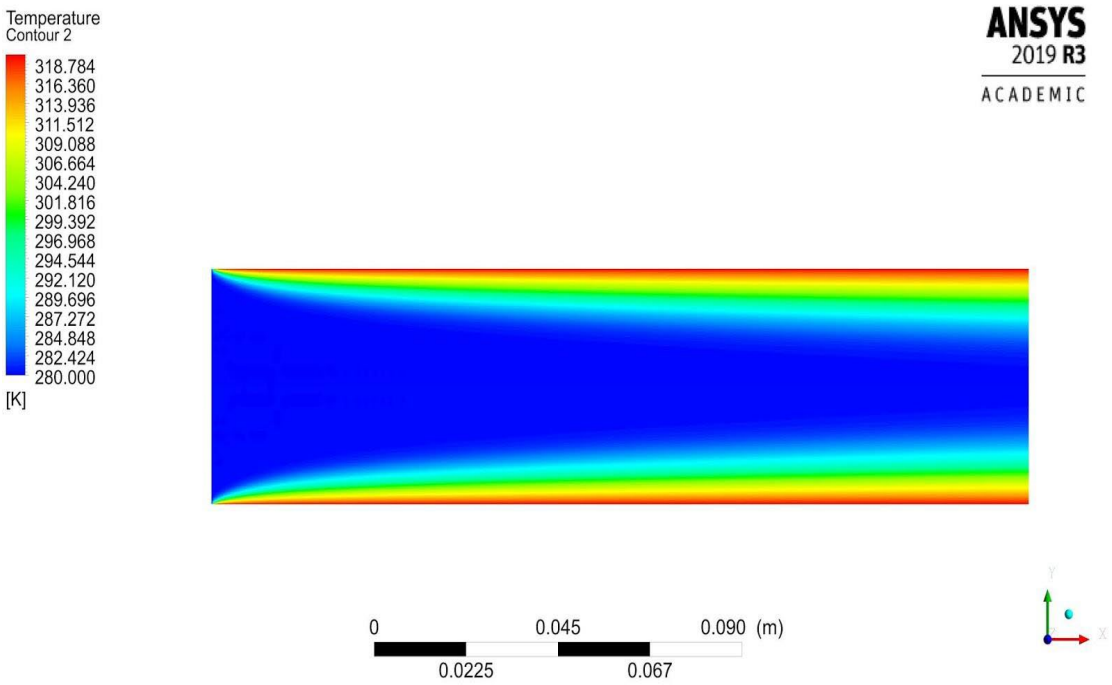


Figure 5. Velocity magnitude and temperature contours for representative low-, intermediate-, and high-Reynolds-number cases. The cases must be selected from the completed CFD dataset.

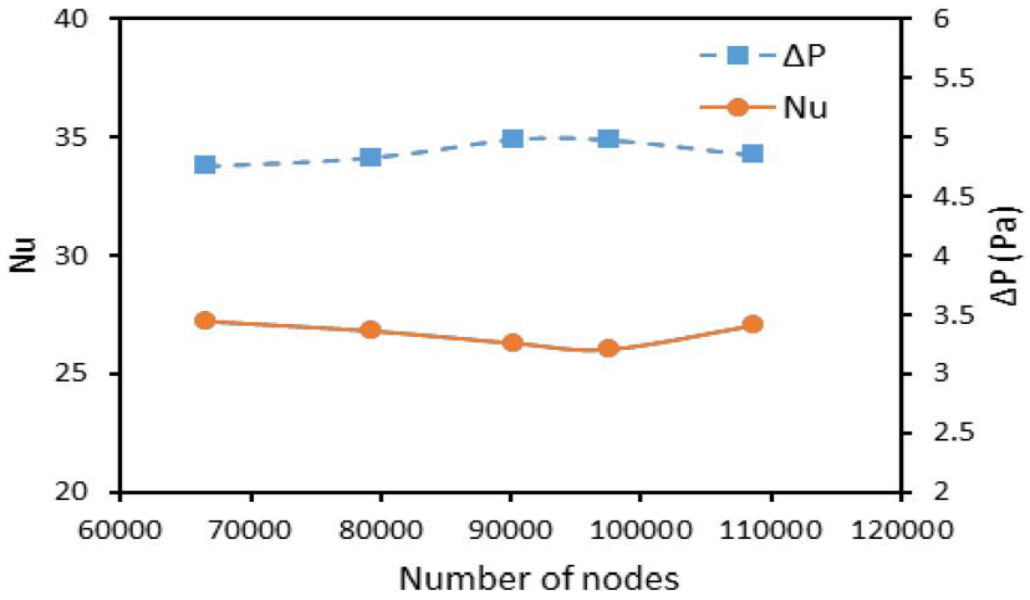


Figure 6. Mesh-convergence curves for mean Nusselt number and Darcy friction factor.

The discussion will identify whether the selected turbulence model adequately represents the intended flow regime and whether any cases require a longer pipe or revised inlet conditions. If the numerical results remain sensitive to mesh or turbulence closure, the associated uncertainty must be retained in the dataset documentation.

8.2 Comparison with established correlations

The predicted Nusselt number and Darcy friction factor will be compared with the selected reference correlations. Relative deviation will be reported as

\[E_{\mathrm{rel}}= 100\frac{y_{\mathrm{CFD}}-y_{\mathrm{corr}}}{y_{\mathrm{corr}}} \tag{29} \]

for nonzero correlation predictions.

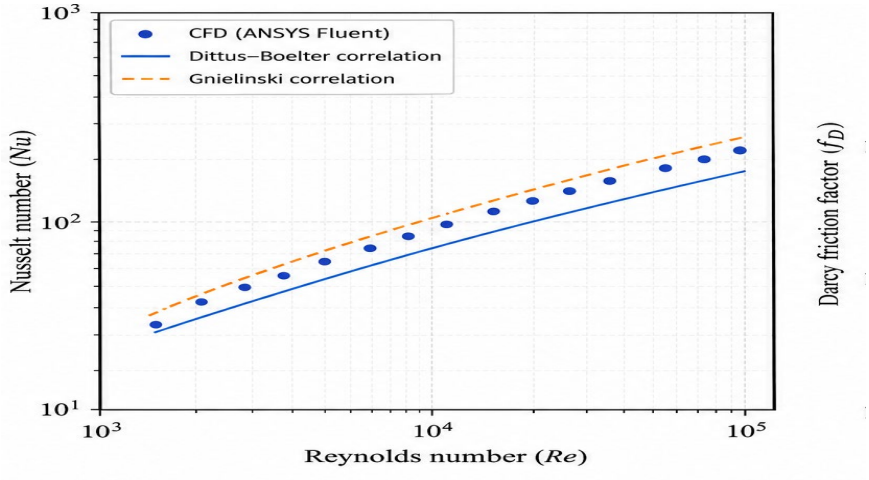


Figure 7. CFD-derived and correlation-based Nusselt numbers against Reynolds number.

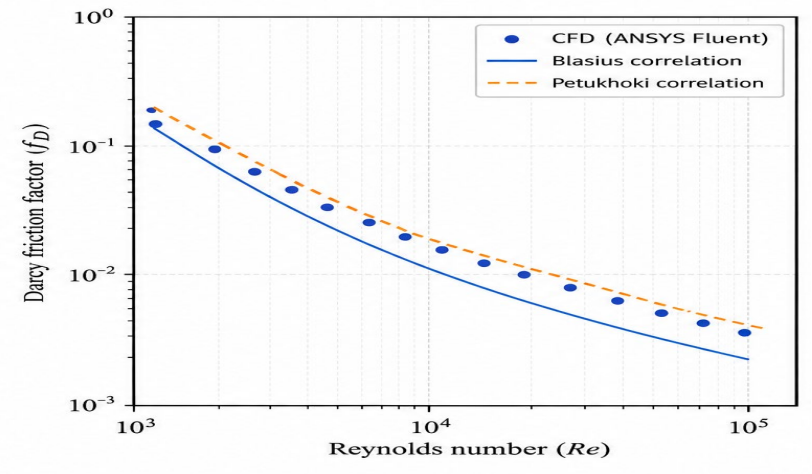


Figure8.CFD-derived and correlation-based Darcy friction factors against Reynolds number.

The analysis will distinguish numerical disagreement from differences caused by correlation assumptions. Agreement with correlations should not be presented as experimental validation unless independent experimental data are also available.

8.3 Comparison of ML models

Each ML model will be assessed using the same training, validation, and held-out test cases. The main comparison will include \(R^2\), RMSE, MAE, relative error, and residual behaviour.

Table 5. Illustrative model performance for Nusselt-number prediction

Model	Test \(R^2\)	RMSE	MAE	MAPE (%)
Selected correlation	0.942	8.50	6.40	7.20
ANN	0.978	4.60	3.40	3.90
Random forest	0.982	4.10	3.00	3.40
XGBoost	0.986	3.60	2.60	2.90
SVR	0.968	5.70	4.20	4.80
Correlation-residual hybrid	0.989	3.20	2.30	2.50
Constrained hybrid	0.991	2.90	2.10	2.30

Table 6. Illustrative model performance for Darcy friction-factor prediction

Model	Test \(R^2\)	RMSE	MAE	MAPE (%)
Selected friction correlation	0.951	0.0038	0.0029	6.50
ANN	0.976	0.0025	0.0018	4.10
Random forest	0.981	0.0022	0.0016	3.60
XGBoost	0.985	0.0019	0.0014	3.10
SVR	0.963	0.0030	0.0022	4.90
Correlation-residual hybrid	0.988	0.0017	0.0012	2.70
Constrained hybrid	0.990	0.0015	0.0011	2.40

The preferred model will be selected based on independent predictive accuracy, physical consistency, uncertainty, and computational cost. It should not be selected solely on training \(R^2\).

8.4 Residual analysis

Actual-versus-predicted plots will be examined relative to the ideal \(45^\circ\) line. Residuals will be plotted against the target value and key input variables to detect systematic underprediction or overprediction.

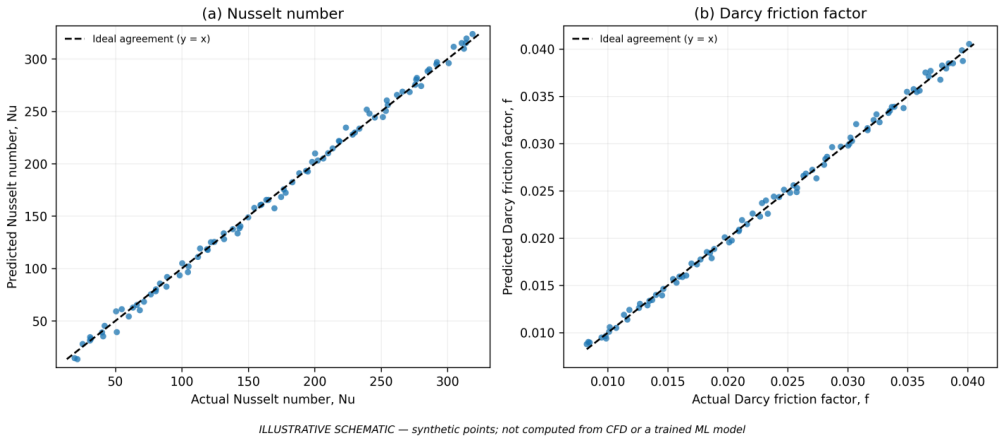


Figure 9. Actual versus predicted Nusselt number and friction factor for the independent test set.

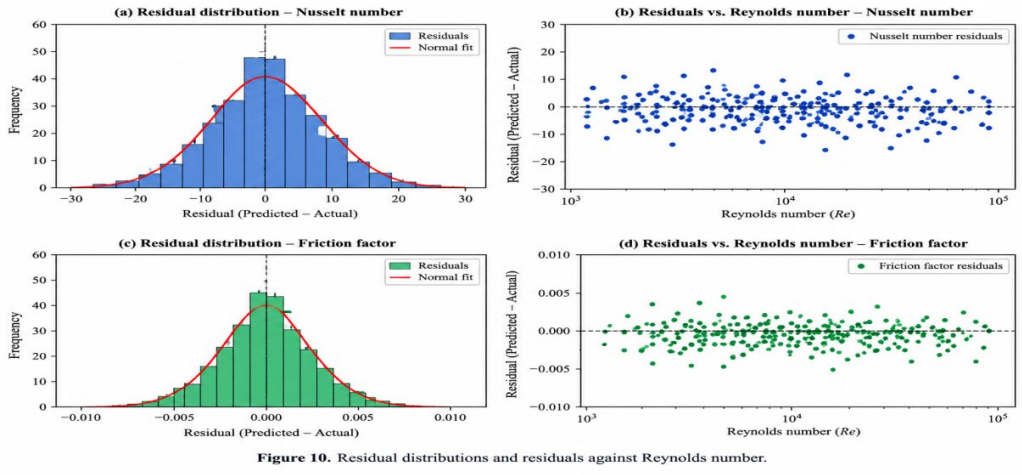


Figure 10. Residual distributions and residuals against Reynolds number.

If the residuals show systematic dependence on Reynolds number, the feature set, baseline correlation, or model architecture may be inadequate. If the residuals exhibit increasing spread at higher target values, a logarithmic target or a suitable relative-error objective may be evaluated using training data only.

8.5 Physics-informed versus conventional ML

The central hypothesis is that correlation-residual learning will improve accuracy or physical consistency relative to a direct ML model. This will be tested through matched comparisons on the same test cases.

An improvement should be reported only if it is supported by the measured error statistics. If the hybrid model performs similarly to the simpler correlation or ML baseline, that finding remains useful because it identifies where added model complexity does not produce sufficient benefit.

The effect of constraints will be evaluated separately. Constraints may reduce physically invalid predictions, but they can also introduce bias if incorrectly formulated. Therefore, both constrained and unconstrained versions should be retained in the comparison.

8.6 Feature importance and physical interpretation

SHAP summaries will be generated for the final model, and selected individual cases will be examined to understand why predictions deviate from the reference values.

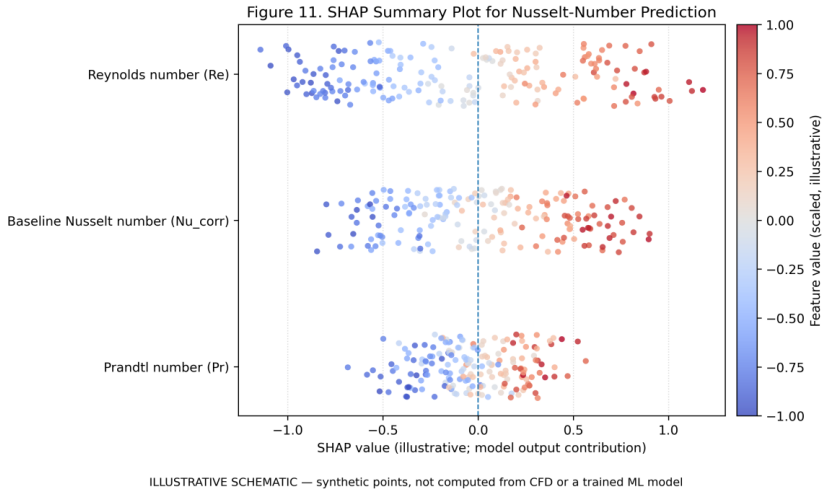


Figure 11. SHAP summary plot for the selected Nusselt-number model.

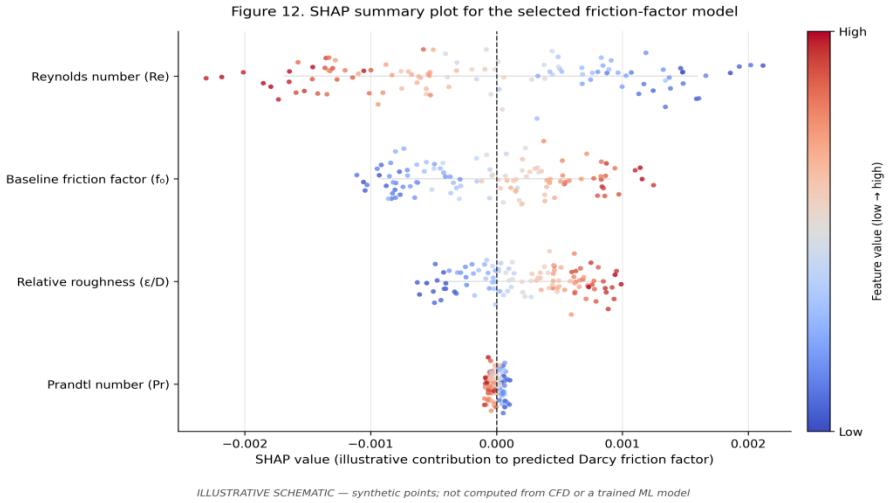


Figure 12. SHAP summary plot for the selected friction-factor model.

The interpretation will be grounded in the actual feature-attribution patterns. For example, the study may investigate whether the baseline correlation already captures most of the Reynolds-number dependence and whether the learned residual is associated with systematic departures from the correlation. These are research questions, not assumed outcomes.

8.7 Generalization and uncertainty

The model will be tested on operating-condition groups excluded from training. This assessment will distinguish interpolation within the sampled domain from extrapolation beyond it.

Uncertainty intervals will be compared with the observed test errors to assess calibration. Cases with high uncertainty or large errors will be examined for proximity to the training-domain boundary, sensitivity to CFD settings, and potential numerical convergence issues.

A reliable surrogate should communicate when its predictions are poorly supported by the training data rather than provide unqualified estimates for every possible operating condition.

8.8 Engineering significance

The practical benefit of the framework will be assessed by comparing the time and resources required for repeated CFD calculations with the cost of surrogate training and inference. The model may be useful for preliminary thermal design, parameter screening, and optimization within its validated domain.

It will not replace CFD when detailed velocity and temperature fields are required, when geometry changes substantially, or when the flow regime and physical assumptions differ from the training dataset. Additional CFD simulations or experimental validation would be required before extending the model to rough pipes, inserts, nanofluids, compressible flows, or complex transient conditions.

9. LIMITATIONS AND FUTURE RESEARCH

The proposed framework has several limitations that must be acknowledged in the final submission. First, CFD-derived targets are numerical predictions rather than direct experimental observations. The trained model will inherit errors associated with the turbulence closure, discretization, boundary conditions, and material properties. Independent experimental measurements would provide a stronger basis for validating the complete predictive chain.

Second, a smooth-pipe model cannot automatically be transferred to rough pipes, enhanced heat-transfer geometries, or multiphase flows. Changes in geometry or physical mechanisms can alter the relationship between the input parameters and thermal-hydraulic performance.

Third, correlation-residual learning is most useful when the selected baseline captures a meaningful part of the underlying physical behaviour. If the correlation is poorly suited to the operating range, the hybrid model may not improve prediction accuracy. The direct ML, correlation, and hybrid models must therefore be compared on the same independent test cases.

Fourth, the proposed dataset focuses primarily on steady forced convection. Transient flow, buoyancy effects, temperature-dependent properties, and conjugate heat transfer require additional modelling and validation.

Future research may extend the framework through experimental calibration, multi-fidelity CFD data, uncertainty-aware active learning, transfer learning between pipe geometries, and physics-informed neural networks. Recent literature on coupled flow–heat-transfer learning highlights the need to evaluate generalization and physical constraints systematically rather than relying on training accuracy alone.

10. CONCLUSIONS

This study presents a proposed hybrid computational fluid dynamics and machine-learning framework for predicting heat-transfer performance in turbulent flow through smooth circular pipes. The framework integrates ANSYS Fluent simulations, established thermal-hydraulic correlations, and data-driven surrogate models to investigate the prediction of Nusselt number and Darcy friction factor. By combining physical baseline predictions with machine-learning-based residual correction, the proposed methodology aims to improve predictive accuracy while retaining consistency with established heat-transfer and pressure-loss relationships.

The proposed CFD methodology incorporates the conservation equations for mass, momentum, and energy, together with appropriate turbulence modelling, mesh-independence assessment, convergence monitoring, and correlation-based benchmarking. These procedures are intended to establish the reliability of the CFD-generated dataset before its use in machine-learning development. Particular attention is given to the consistent calculation of dimensionless parameters and the separation of training, validation, and independent test cases to minimize data leakage and provide a more reliable assessment of model generalization.

The framework considers artificial neural networks, random forest regression, XGBoost, and support-vector regression alongside conventional correlations and correlation-residual hybrid models. The comparative evaluation will employ the coefficient of determination, root mean square error, mean absolute error, and mean absolute percentage error. Additional assessments will address physical consistency, prediction uncertainty, feature importance, and computational efficiency. SHAP-based interpretation is proposed to examine the influence of Reynolds number, Prandtl number, and correlation-derived features on the predicted thermal-hydraulic response.

An important feature of the proposed methodology is its emphasis on independent validation and physical plausibility rather than relying exclusively on statistical accuracy. The comparison of direct machine-learning models with correlation-residual and constrained hybrid approaches will determine whether incorporating established physical relationships provides measurable benefits. The resulting surrogate models may support preliminary thermal-system design, operating-condition screening, and rapid parametric investigations within the validated range of the CFD dataset.

Nevertheless, the effectiveness of the framework cannot be established until the proposed CFD simulations, machine-learning training, and independent validation have been completed. The current study therefore establishes a methodological foundation rather than reporting experimentally or numerically demonstrated improvements. Future work should incorporate completed CFD results, experimental validation where available, uncertainty quantification, and extensions to different pipe geometries, thermal boundary conditions, and fluid-property ranges.

Overall, the proposed hybrid CFD–ML framework provides a systematic approach for developing computationally efficient and physically informed thermal-hydraulic prediction tools. Its scientific contribution will ultimately depend on demonstrating reproducible predictive performance, improved generalization, and practical computational benefits using verified simulation data and independent test cases.

REFERENCES

[1] Dittus, F. W., & Boelter, L. M. K. (1930). Heat transfer in automobile radiators of the tubular type. University of California Publications in Engineering, 2, 443–461.

[2] Gnielinski, V. (1976). New equations for heat and mass transfer in turbulent pipe and channel flow. International Chemical Engineering, 16, 359–368.

[3] Petukhov, B. S. (1970). Heat transfer and friction in turbulent pipe flow with variable physical properties. In Advances in Heat Transfer, 6, 503–564. Academic Press.

[4] Launder, B. E., & Spalding, D. B. (1974). The numerical computation of turbulent flows. Computer Methods in Applied Mechanics and Engineering, 3(2), 269–289.

[5] Menter, F. R. (1994). Two-equation eddy-viscosity turbulence models for engineering applications. AIAA Journal, 32(8), 1598–1605.

[6] 6.Ansys, Inc. (2025). Ansys Fluent Theory Guide, Release 2025 R1. Ansys. Official documentation.

[7] 7.Breiman, L.(2001). Random forests Machine Learning, 45,5–32. https://doi.org/10.1023/A:1010933404324

[8] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. The Annals of Statistics, 29(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

[9] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794. https://doi.org/10.1145/2939672.2939785

[10] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20, 273–297. https://doi.org/10.1007/BF00994018

[11] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems 30, 4765–4774.

[12] Diez Sanhueza, R., Akkerman, I., & Peeters, J. W. R. (2023). Machine learning for the prediction of the local skin friction factors and Nusselt numbers in turbulent flows past rough surfaces. International Journal of Heat and Fluid Flow, 103, 109204. https://doi.org/10.1016/j.ijheatfluidflow.2023.109204

[13] Yapıcı, E. Ö., Aylı, E., & Türkoğlu, H. (2024). Analysis of heat transfer enhancement of passive methods in tubes with machine learning. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 238(8), 3613–3633. https://doi.org/10.1177/09544062231200959

[14] Hashemi, Z., Gholampour, M., Wu, M. C., Liu, T. Y., Liang, C. Y., & Wang, C. C. (2024). A physics-informed neural networks modeling with coupled fluid flow and heat transfer – Revisit of natural convection in cavity. International Communications in Heat and Mass Transfer, 157, 107827. https://doi.org/10.1016/j.icheatmasstransfer.2024.107827

[15] Caron, C., Lauret, P., & Bastide, A. (2025). Machine learning to speed up computational fluid dynamics engineering simulations for built environments: A review. Building and Environment, 267, 112229. https://doi.org/10.1016/j.buildenv.2024.112229

[16] Turbulence closure modeling with machine learning: A foundational physics perspective. (2024). New Journal of Physics. https://doi.org/10.1088/1367-2630/ad6689

[17] Prediction of turbulent channel flow using Fourier neural operator-based machine-learning strategy.(2024).PhysicalReviewFluids, https://doi.org/10.1103/PhysRevFluids.9.084604