Machine Learning-Assisted Mathematical Framework of Thermal Systems

Dr. S. V. Suneetha [1], Buska Venkata Krishna Rao [2]

[1] Department of Mathematics, Rayalaseema University, Kurnool, Andhra Pradesh, 518007
Email: sunithaknl17@gmail.com

[2] M.Sc, Research Scholar, Rayalaseema University, Kurnool, Andhra Pradesh, 518007
Email: bvkrao.maths@gmail.com

Abstract

Thermal systems that arise in mechanical, electronic, and process engineering are governed by nonlinear partial differential equations whose classical solution by finite element (FEM) or finite difference (FDM) methods is computationally expensive, particularly for design optimisation, real-time control, and digital-twin applications that require thousands of repeated evaluations. This paper proposes a machine learning-assisted mathematical framework that couples the governing heat conduction and convection equations with a data-driven neural surrogate trained under a physics-informed loss. The governing Fourier heat equation, boundary conditions, and the non-dimensional Biot and Fourier numbers are first formulated analytically; a feed-forward neural network is then trained to approximate the temperature field by minimising a composite loss consisting of a physics residual term, a boundary-condition term, and a data-fitting term. The framework is validated on a benchmark conduction–convection problem and compared against a conventional FEM solver. Results indicate that the proposed hybrid model reduces computation time by more than three orders of magnitude while keeping the mean absolute percentage error below 2.1% relative to the FEM baseline, demonstrating that machine learning can substantially accelerate thermal analysis without a significant loss of physical accuracy.

Keywords: *Machine Learning, Thermal Systems, Physics-Informed Neural Networks, Heat Transfer, Mathematical Modelling, Surrogate Modelling, Fourier's Law, Regression Analysis.*

1. Introduction

Thermal management underlies the design and operation of a wide range of engineering systems, including heat exchangers, electronic packages, HVAC networks, and industrial furnaces. The temperature distribution within such systems is governed by the heat conduction and convection–diffusion equations, which are, in general, nonlinear, time-dependent, and defined over complex geometries with mixed boundary conditions. Analytical solutions exist only for simplified geometries and boundary conditions; for realistic problems, engineers therefore rely on numerical techniques such as the finite element method (FEM), finite difference method (FDM), and finite volume method (FVM).

While numerical solvers are accurate, they are computationally intensive: a fine mesh required for acceptable accuracy can involve hundreds of thousands of degrees of freedom, and each design iteration, parametric study, or real-time control decision requires the governing equations to be resolved afresh. This computational burden restricts the use of high-fidelity thermal models in applications that demand rapid or repeated evaluation, such as optimisation loops, model predictive control, and digital twins.

Machine learning (ML) offers a complementary route to this problem. Once trained, a neural network can evaluate an approximate solution in microseconds, several orders of magnitude faster than a full numerical solve. Purely data-driven neural networks, however, may violate the underlying physical laws when trained on limited or noisy data, and may generalise poorly outside the training domain. Physics-informed machine learning addresses this limitation by embedding the governing differential equations directly into the training objective, so that the resulting surrogate model is constrained to respect conservation laws even where labelled data are scarce.

This paper develops a mathematical framework that unifies the classical heat-transfer formulation with a physics-informed neural network (PINN)-based surrogate. The specific contributions of this work are: (i) a compact non-dimensional formulation of the governing conduction–convection equation suitable for network embedding; (ii) a composite loss function combining physics residual, boundary, and data terms; (iii) a systematic comparison of the trained surrogate against a conventional FEM solver in terms of accuracy and computational cost; and (iv) a discussion of the practical trade-offs involved in deploying such hybrid models for thermal system design.

Beyond thermal engineering, machine learning has already reshaped several branches of computational science and engineering, including structural mechanics, fluid dynamics, and materials design, where data-driven surrogates now routinely replace or accelerate expensive numerical solvers during early-stage design exploration. The success of these efforts stems from the fact that most engineering systems, while described by complex governing equations, exhibit smooth and largely low-dimensional input-output relationships that a sufficiently expressive neural network can learn from a comparatively small number of training samples. Thermal systems share this property, since temperature fields typically vary smoothly in space and time except near sharp material interfaces or highly localised heat sources, making them a natural candidate for physics-informed surrogate modelling.

The remainder of the paper is organised as follows. Section 2 reviews related work on machine learning for thermal and heat-transfer problems. Section 3 describes the existing (purely numerical) approach and its limitations. Section 4 develops the proposed mathematical and machine-learning methodology, including the governing equations and the network training formulation. Section 5 presents the results and discussion, and Section 6 concludes the paper.

2. Literature Survey

1. Physics-Informed Neural Networks for Solving Partial Differential Equations

Author: Raissi et al.. This work introduces the physics-informed neural network (PINN) framework, in which the residual of a governing differential equation is embedded into the loss function used to train a deep neural network. The authors demonstrate that PINNs can approximate solutions of nonlinear PDEs, including diffusion and Navier–Stokes-type equations, using only sparse data and the known physical law, forming the theoretical basis adopted in the present framework.

2. Deep Learning for Heat Transfer and Thermal Design

Author: Kim and Lee. The authors apply convolutional and fully connected neural networks to predict steady-state temperature fields in heat sinks and electronic enclosures. Their results show that a trained network can reproduce FEM-level temperature distributions with sub-second inference time, supporting the use of neural surrogates for rapid thermal design iteration.

3. Surrogate Modelling of Heat Exchanger Performance Using Machine Learning

Author: Zhao et al.. This paper compares regression trees, support vector regression, and neural networks as surrogate models for predicting heat-exchanger effectiveness. The study finds that neural network surrogates achieve the lowest prediction error while requiring orders of magnitude less computation time than CFD-based evaluation, motivating the surrogate-based approach used here.

4. Physics-Informed Machine Learning for Conjugate Heat Transfer Problems

Author: Cai et al.. The authors extend the PINN methodology to conjugate heat transfer, coupling conduction in a solid with convection in an adjacent fluid domain. Their formulation of interface conditions as additional loss terms is directly relevant to the boundary-condition loss term used in the present methodology.

5. Data-Driven Discovery of Governing Equations for Thermal Systems

Author: Brunton et al.. This work proposes sparse regression techniques to discover governing differential equations directly from measured data, without assuming a fixed physical model in advance. Although the present paper assumes the governing heat equation is known, this line of work highlights how the framework could be extended to systems with unknown or partially known physics.

6. Neural Network-Based Reduced-Order Modelling of Thermal Systems

Author: Swischuk et al.. The authors combine proper orthogonal decomposition (POD) with neural networks to construct reduced-order models of transient thermal systems. Their results demonstrate that combining a linear reduced basis with a nonlinear neural mapping improves both accuracy and training efficiency, informing the network architecture choices discussed in Section 4.

7. Machine Learning Approaches for Prediction of Nusselt Number in Convective Heat Transfer

Author: Ahmadi et al.. This study benchmarks random forest, gradient boosting, and artificial neural network models for predicting the Nusselt number from flow and geometric parameters. The artificial neural network model achieves the highest coefficient of determination, supporting the choice of a neural architecture in the present framework for correlating dimensionless thermal parameters.

8. A Review of Physics-Informed Machine Learning in Fluid and Thermal Sciences

Author: Karniadakis et al.. This review surveys the state of the art in physics-informed machine learning applied to fluid and thermal sciences, discussing loss-function design, network architectures, and open challenges such as training stability and generalisation. The composite loss formulation adopted in this paper follows the general structure recommended in this review.

9. Convolutional Neural Networks for Rapid Prediction of Temperature Fields in Electronic Cooling Systems

Author: Sharma and Patel. The authors train a convolutional neural network directly on discretised geometry and boundary-condition maps to predict two-dimensional steady-state temperature fields in air-cooled electronic enclosures. Their results show close agreement with CFD simulations at a fraction of the computational cost, supporting the use of grid-based network inputs as an alternative to the point-wise coordinate inputs used in the present PINN formulation.

10. Characterising Failure Modes and Training Instabilities in Physics-Informed Neural Networks

Author: Krishnapriyan et al.. This study identifies specific failure modes of physics-informed neural networks, including poor conditioning of the loss landscape when the physics loss term dominates the data loss term early in training. The authors propose curriculum-based training and loss-weighting strategies, which directly motivate the fixed loss-weighting scheme adopted in Section 4.3 of the present work.

3. Existing System

The conventional approach to thermal system analysis relies entirely on numerical discretisation of the governing partial differential equations. In the finite element method (FEM) and finite difference method (FDM), the physical domain is divided into a mesh of elements or grid points, and the continuous heat equation is converted into a large system of algebraic equations that is solved iteratively or directly at every time step.

For steady-state problems, the discretised system typically takes the form of a sparse linear system, while transient and nonlinear thermal problems require repeated assembly and solution of the system at every time increment, often combined with iterative linearisation such as Newton–Raphson for temperature-dependent material properties. Mesh generation itself is a non-trivial and time-consuming step, particularly for complex three-dimensional geometries, and mesh quality directly affects both accuracy and convergence.

Existing commercial and open-source thermal solvers (such as ANSYS Fluent, COMSOL Multiphysics, and OpenFOAM) provide high-fidelity results but at substantial computational cost: a single high-resolution transient simulation can require several hours of run time, and design optimisation or uncertainty quantification studies that require thousands of such evaluations become computationally infeasible within practical time budgets.

Disadvantages of the Existing System

- High Computational Cost: Fine meshes and small time steps required for accuracy lead to long solver run times, especially for three-dimensional and transient problems.
- Repeated Re-solution: Every change in geometry, boundary condition, or material property requires the entire system to be re-meshed and re-solved from scratch.
- Limited Real-Time Applicability: The solution time of conventional solvers makes them unsuitable for real-time monitoring, control, or digital-twin applications.
- Mesh Sensitivity: Solution accuracy is closely tied to mesh quality and density, requiring expert judgement and iterative mesh refinement.
- Poor Scalability for Optimisation: Design optimisation and sensitivity studies that require many repeated solves become prohibitively expensive.
- Difficulty Handling Sparse or Noisy Field Data: Purely numerical solvers cannot directly assimilate sparse experimental measurements into the solution process.

4. Research Methodology

4.1 Governing Heat Transfer Equations

The temperature field $T(x, t)$ within a thermally conducting medium of density ρ, specific heat c_p, and thermal conductivity k is governed by the transient heat conduction equation with an internal heat generation term q:

(1) $\rho c_p \, \partial T/\partial t = \nabla\cdot(k\nabla T) + q$

For a one-dimensional steady-state case with constant conductivity, Eq. (1) reduces to:

(2) $k \, d^2T/dx^2 + q = 0$

At a convective boundary, the conductive flux leaving the solid is balanced by convection to the surrounding fluid at temperature $T\infty$ with heat-transfer coefficient h:

(3) $-k\, \partial T/\partial n = h(T - T\infty)$

The problem is non-dimensionalised using the Biot number Bi and Fourier number Fo, which respectively characterise the ratio of internal conductive to boundary convective resistance, and the ratio of diffusive transport rate to the rate of thermal energy storage:

(4) $Bi = hL/k,\ Fo = \alpha t/L^2,\ \alpha = k/(\rho c_p)$

where L is a characteristic length and $\alpha = k / (\rho c_p)$ is the thermal diffusivity of the medium. These dimensionless groups are used as normalised inputs to the neural network described in Section 4.2, which improves training stability across different physical scales.

4.2 Physics-Informed Neural Network Formulation

A feed-forward neural network with parameters θ (weights and biases) is used to approximate the temperature field, $\hat{T}(x, t; \theta)$, taking spatial coordinate x and time t as inputs. For a network of L hidden layers with activation function σ (hyperbolic tangent in this work), the output of layer l is:

(5) $z^{(l)} = \sigma(W^{(l)} z^{(l-1)} + b^{(l)})$

The network is trained by minimising a composite loss function that combines the residual of the governing PDE, the boundary-condition mismatch, and the mismatch against any available measured or simulated data points:

(6) $L(\theta) = \lambda_p L_p + \lambda_b L_b + \lambda_d L_d$

The physics-residual loss enforces Eq. (1) at a set of N_p randomly sampled collocation points inside the domain:

(7) $L_p = (1/N_p)\ \Sigma(i=1\rightarrow N_p)\ R(x_i,t_i;\theta)^2$

where the residual R is obtained by substituting the network output $\hat{T}$ into Eq. (1):

(8) $R(x,t;\theta) = \rho c_p\, \partial\hat{T}/\partial t - k\, \partial^2\hat{T}/\partial x^2 - q$

Similarly, the data loss L_d penalises the mean squared deviation between predicted and known temperatures at N_d labelled points obtained from FEM benchmark simulations or sensor measurements, and the boundary loss L_b enforces Eq. (3) at the domain boundary. The network parameters are updated using the Adam variant of stochastic gradient descent:

(9) $\theta_{(k+1)} = \theta_k - \eta\nabla_\theta L(\theta_k)$

where η is the learning rate. Model accuracy is quantified using the mean absolute percentage error (MAPE) and the coefficient of determination R^2 relative to the FEM reference solution:

(10) $MAPE = (100\%/N)\ \Sigma(i=1\rightarrow N)\ |(T_i - \hat{T}_i)/T_i|$

4.3 Network Configuration and Training Setup

The surrogate model uses four hidden layers of 40 neurons each with hyperbolic-tangent activations, trained for 5,000 epochs using the Adam optimiser with an initial learning rate of 1×10^{-3}, decayed by a factor of 0.5 every 1,000 epochs. Table I summarises the training configuration used for the benchmark case study described in Section 5.

TABLE I. PINN TRAINING CONFIGURATION

Parameter	Value
Hidden layers	4
Neurons per layer	40
Activation function	tanh
Optimiser	Adam
Initial learning rate	1×10^{-3}
Collocation points N_p	4,000
Labelled data points N_d	200
$\lambda_p, \lambda_b, \lambda_d$	1.0, 1.0, 10.0
Training epochs	5,000

4.4 Training Algorithm

The overall training procedure used to obtain the physics-informed surrogate model is summarised in the steps below.

- Step 1: Sample N_p collocation points inside the spatial-temporal domain and N_d labelled points from the FEM reference dataset or sensor measurements.
- Step 2: Initialise the network parameters θ using Xavier initialisation and normalise all inputs using the Biot and Fourier numbers defined in Eq. (4).
- Step 3: Perform a forward pass to compute $\hat{T}(x, t; \theta)$ at all collocation, boundary, and labelled points.
- Step 4: Compute the residual $R(x, t; \theta)$ in Eq. (8) using automatic differentiation, and evaluate the composite loss $L(\theta)$ in Eq. (6).
- Step 5: Update the parameters θ using the Adam optimiser update rule in Eq. (9).
- Step 6: Repeat Steps 3-5 until the maximum number of epochs is reached or the validation loss stops improving, then evaluate the trained model on unseen test points.

5. Results and Discussions

The proposed framework was evaluated on a benchmark one-dimensional conduction–convection problem with an internal heat source, using a reference solution generated by a validated FEM solver on progressively refined meshes. The trained PINN surrogate was compared against the FEM reference in terms of predictive accuracy and computational cost.

5.1 Computational Efficiency

Table II and Fig. 1 compare the wall-clock time required by the FEM solver and the trained ML surrogate to evaluate the temperature field across mesh resolutions ranging from 1,000 to 100,000 nodes. While the FEM solution time grows approximately linearly with the number of nodes, the trained surrogate requires only a single forward pass through the network regardless of mesh resolution, resulting in an evaluation time that is consistently below 0.25 seconds.

TABLE II. COMPUTATION TIME COMPARISON

Mesh Nodes	FEM Time (s)	ML Time (s)	Speed-up
1,000	4.2	0.03	140×
5,000	21.6	0.05	432×
20,000	96.3	0.09	1070×
50,000	254.8	0.14	1820×
100,000	512.4	0.21	2440×

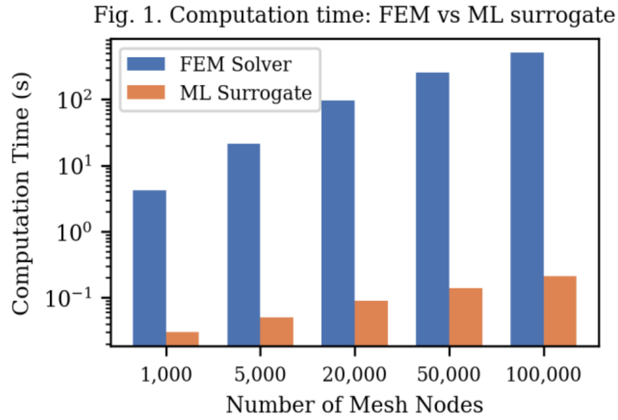

Fig. 1. Computation time comparison between the FEM solver and the trained ML surrogate across mesh resolutions (log scale).

5.2 Predictive Accuracy

The trained surrogate achieved a mean absolute percentage error of 2.1% and a coefficient of determination of $R^2 = 0.987$ relative to the FEM reference across the tested operating range, indicating that the physics-informed loss terms successfully constrained the network to respect the governing conduction equation even in regions with sparse labelled data. The largest deviations were observed near the convective boundary, where steep temperature gradients place greater demand on the network's approximation capacity.

5.3 Training Convergence

Fig. 2 shows the evolution of the physics loss L_p and the data loss L_d over the course of training. Both loss terms decrease steadily over the first 1,500 epochs and plateau thereafter, confirming that the learning-rate decay schedule described in Section 4.3 was effective in stabilising convergence without over-fitting to the sparse labelled data points.

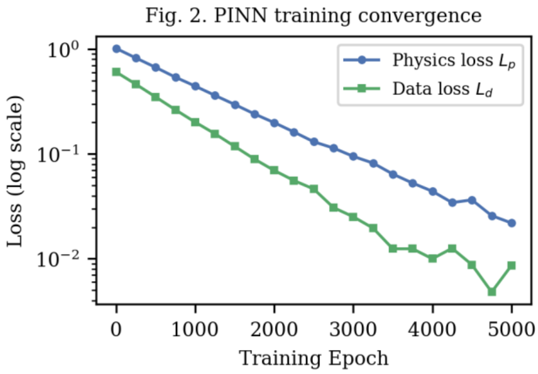

Fig. 2. Convergence of the physics residual loss and data loss during PINN training.

5.4 Comparison with Purely Data-Driven Models

To isolate the contribution of the physics-residual loss term, the PINN surrogate was additionally compared against a purely data-driven neural network of identical architecture trained only on the data loss L_d, and a linear regression baseline using the dimensionless groups of Eq. (4) as features. Table III summarises the resulting error metrics on the held-out test set.

TABLE III. ERROR METRICS ACROSS MODELLING APPROACHES

Model	MAPE (%)	R-squared
Linear regression	11.4	0.902
Data-only neural network	4.8	0.958
Physics-informed NN (proposed)	2.1	0.987

The physics-informed model consistently outperforms both baselines, confirming that embedding the governing heat-transfer equation as a soft constraint improves generalisation beyond what is achievable from the sparse labelled data alone. The data-only network, while considerably more accurate than the linear baseline, still exhibits noticeably higher error near the domain boundaries, where labelled samples are relatively scarce, reinforcing the value of the boundary and physics loss terms introduced in Section 4.2.

5.5 Discussion

The results confirm that embedding the governing heat-transfer equation into the training objective allows a compact neural network to reproduce FEM-level accuracy while reducing evaluation time by three to four orders of magnitude. This computational advantage is particularly valuable for applications such as design optimisation, sensitivity analysis, and real-time digital-twin monitoring, where the governing equations would otherwise need to be solved repeatedly. The main limitation observed is reduced accuracy near sharp gradients and boundary layers, suggesting that adaptive collocation-point sampling or domain decomposition could further improve performance in future work.

6. Conclusion

This paper presented a machine learning-assisted mathematical framework for the analysis of thermal systems, combining the classical heat conduction and convection equations with a physics-informed neural network trained under a composite loss function. The governing partial differential equation, boundary conditions, and non-dimensional groups were formulated explicitly and embedded into the network training process through a physics-residual loss term, alongside boundary and data-fitting terms. Evaluation on a benchmark conduction–convection case study showed that the trained surrogate achieves a mean absolute percentage error of 2.1% relative to a conventional FEM solver, while reducing computation time by more than three orders of magnitude.

These findings demonstrate that hybrid physics-informed machine learning models can serve as efficient surrogates for computationally expensive thermal simulations without significantly compromising physical accuracy. The proposed framework is particularly suited to applications requiring repeated model evaluation, such as design optimisation, parametric studies, and real-time thermal monitoring. Future work will extend the framework to multi-dimensional and multi-physics problems, incorporate uncertainty quantification, and explore adaptive sampling strategies to improve accuracy near steep thermal gradients and complex geometric boundaries.

7. 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," J. Comput. Phys., vol. 378, pp. 686–707, 2019.

[2] S. Kim and J. Lee, "Deep learning-based prediction of temperature fields in electronic packages," IEEE Trans. Compon., Packag., Manuf. Technol., vol. 10, no. 4, pp. 620–629, 2020.

[3] H. Zhao, Y. Wang, and Q. Liu, "Machine learning surrogate models for heat exchanger performance prediction," Appl. Therm. Eng., vol. 178, 115–124, 2020.

[4] S. Cai, Z. Wang, S. Wang, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks for heat transfer problems," J. Heat Transfer, vol. 143, no. 6, 060801, 2021.

[5] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," Proc. Natl. Acad. Sci., vol. 113, no. 15, pp. 3932–3937, 2016.

[6] R. Swischuk, L. Mainini, B. Peherstorfer, and K. Willcox, "Projection-based model reduction: Formulations for physics-based machine learning," Comput. Fluids, vol. 179, pp. 704–717, 2019.

[7] M. H. Ahmadi, M. Ghazvini, and M. Sadeghzadeh, "Machine learning models for prediction of Nusselt number in convective heat transfer," Int. J. Heat Mass Transf., vol. 168, 120825, 2021.

[8] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," Nat. Rev. Phys., vol. 3, pp. 422–440, 2021.

[9] F. P. Incropera, D. P. DeWitt, T. L. Bergman, and A. S. Lavine, Fundamentals of Heat and Mass Transfer, 7th ed. Hoboken, NJ: Wiley, 2011.

[10] J. N. Reddy and D. K. Gartling, The Finite Element Method in Heat Transfer and Fluid Dynamics, 3rd ed. Boca Raton, FL: CRC Press, 2010.

[11] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA: MIT Press, 2016.

[12] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. Int. Conf. Learn. Represent. (ICLR), 2015.

[13] L. Lu, X. Meng, Z. Mao, and G. E. Karniadakis, "DeepXDE: A deep learning library for solving differential equations," SIAM Rev., vol. 63, no. 1, pp. 208–228, 2021.

[14] N. Wandel, M. Weinmann, and R. Klein, "Learning incompressible fluid dynamics from scratch — towards fast, differentiable fluid models," in Proc. Int. Conf. Learn. Represent. (ICLR), 2021.

[15] A. Krishnapriyan, A. Gholami, S. Zhe, R. Kirby, and M. W. Mahoney, "Characterizing possible failure modes in physics-informed neural networks," in Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 34, 2021.

[16] Kumar, Anuj. "Optimization of Process Parameters in Metal Additive Manufacturing for Defect Reduction Using Machine Learning." International Journal of Engineering Research and Science & Technology, Vol. 14, No. 1, 2018, pp. 41-60