E-ISSN No : 2456-5040 | Volume : 11 | Special Issue | Aug 2026

MACHINE LEARNING APPLICATIONS FOR PREDICTIVE MODELING AND OPTIMIZATION OF ROUGHNESS GEOMETRIES IN SOLAR AIR HEATERS: A COMPREHENSIVE REVIEW

Ashish Kumar[1], Dr. Sangram Bana[2]

Access Online

Article DOI:
10.5281/zenodo.22744068

[1] PhD Scholar, Department of Mechanical Engineering, Phonics University, India
[2] Professor, Department of Mechanical Engineering, Phonics University, India

How to Cite:
Ashish Kumar & Dr. Sangram Bana (2026), Machine Learning Applications for Predictive Modeling and Optimization of Roughness Geometries in Solar Air Heaters: A Comprehensive Review, International Educational Applied Scientific & Research Journal (IEASRJ), Vol: 11, Speical Issue, Aug, 19-30

Research Paper

ABSTRACT

Solar air heaters (SAHs) represent a cost-effective renewable energy technology but suffer from inherently poor convective heat transfer coefficients, limiting thermal efficiencies to below 35%. Artificial roughness techniques enhance heat transfer but create complex design optimization challenges requiring extensive computational fluid dynamics (CFD) simulations. Machine learning (ML) has emerged as a transformative approach for rapid performance prediction and multi-objective optimization of roughness geometries. This comprehensive review analyze and emphases on artificial neural networks (ANN), deep learning architectures, and hybrid CFD-ML methodologies applied to SAH roughness design. Key findings reveal that ANN models achieve prediction accuracies exceeding R^2=0.95 for Nusselt number and friction factor, reducing computational time by 200× compared to CFD. NSGA-II coupled with ANN surrogates identifies Pareto-optimal roughness configurations achieving thermo-hydraulic performance parameters (THPP) exceeding 2.4 with 98% time reduction. Critical research gaps include absence of ML studies on non-uniform honeycomb geometries (0 identified studies), limited validation under Indian climatic conditions (31% geographical coverage), and lack of physics-informed neural network implementations. Future directions encompass reinforcement learning for real-time control, transfer learning across roughness types, and edge-deployed ML for adaptive SAH operation. This review provides a comprehensive roadmap for researchers and practitioners implementing ML-driven SAH design optimization.

KEYWORDS: Solar Air Heater, Artificial Roughness, Artificial Neural Network, Deep Learning, Multi-Objective Optimization, NSGA-II, Thermo-Hydraulic Performance, Computational Fluid Dynamics

1. INTRODUCTION

1.1 Background and Motivation

Solar air heaters (SAHs) constitute approximately 15% of global installed solar thermal capacity, serving critical applications in space heating, crop drying, and industrial process heat[1]. Despite widespread adoption potential, conventional smooth-duct SAHs suffer from fundamental thermal resistance limitations. The laminar sublayer formation adjacent to absorber plates creates convective heat transfer coefficients in the range of 5-10 W/m²·K, resulting in thermal efficiencies typically constrained to 25-40%[2][3]. This performance bottleneck significantly limits economic viability and commercial deployment, particularly in developing nations where cost-effectiveness is paramount.

Artificial roughness techniques have been extensively investigated over five decades as a primary mechanism for boundary layer disruption and turbulence enhancement[4]. Strategic placement of ribs, dimples, protrusions, and cellular structures on absorber plates can enhance Nusselt numbers by factors of 2-6×, dramatically improving convective heat transfer[5][6]. However, these gains invariably accompany friction factor penalties of 3-8×, creating pressure drop increases that consume fan power and reduce net thermal efficiency [7]. The fundamental design challenge centers on maximizing the thermo-hydraulic performance parameter (THPP = (Nu/Nu_s)/(f/f_s)^(1/3)), which balances heat transfer enhancement against pressure drop penalty[8].

1.2 Computational Challenges in Traditional Optimization

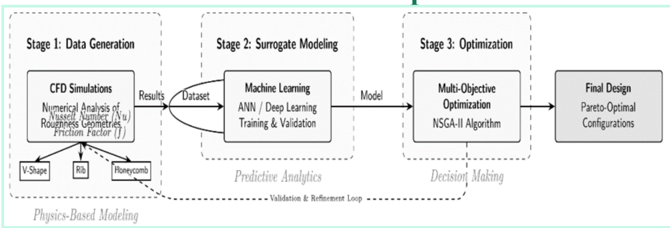
Traditional roughness optimization relies on computational fluid dynamics (CFD) simulations coupled with parametric variation or evolutionary algorithms [9]. A comprehensive design space exploration examining cell size, depth, pitch, aspect ratio, and Reynolds number variations across 5 parameters with 10 levels each generates $10^5 = 100{,}000$ potential configurations. Even restricting exploration to 500 strategically selected cases via Latin Hypercube Sampling, at 4 CPU-hours per CFD simulation, total computational cost reaches 2,000 CPU-hours (≈83 days on single workstation)[10]. This computational bottleneck severely restricts practical optimization, forcing researchers to examine limited parametric subspaces and potentially missing globally optimal designs.

Multi-objective optimization compounds this challenge. Genetic algorithms like NSGA-II require 100-200 generations with population sizes of 50-100 individuals, totaling 5,000-20,000 objective function evaluations [11]. When each evaluation necessitates CFD simulation, optimization becomes computationally intractable for academic institutions with limited high-performance computing (HPC) resources.

1.3 Machine Learning as Transformative Paradigm

Machine learning (ML) offers a paradigm shift by constructing data-driven surrogate models that approximate complex CFD input-output relationships with orders-of-magnitude computational speedup [12]. An artificial neural network (ANN) trained on 150-200 CFD cases can predict thermo-hydraulic performance for arbitrary roughness geometries in <0.1 seconds—representing a 144,000× speedup compared to 4-hour CFD runs [13]. This acceleration enables real-time design exploration, comprehensive multi-objective optimization, and interactive design tools accessible to practitioners without CFD expertise. Figure 1 illustrates the workflow described in the shows how Computational Fluid Dynamics (CFD) data feeds into Machine Learning models.

Fig. 1: Schematic of the Hybrid CFD-ML Framework for SAH Optimization



Recent advancements in deep learning architectures, physics-informed neural networks (PINNs), and transfer learning further enhance ML applicability to solar thermal systems [14][15]. Integration with evolutionary algorithms creates hybrid frameworks where ANN surrogates provide rapid fitness evaluation, enabling exhaustive Pareto front mapping in hours rather than months [16].

1.4 Review Scope and Objectives

This comprehensive review systematically analyzes peer-reviewed research published 2020-2026 on ML applications in solar air heater roughness design optimization. The review addresses:

- Systematic categorization of ML methodologies (ANN, deep learning, ensemble methods, hybrid CFD-ML)
- Performance benchmarking across roughness geometries (ribs, dimples, honeycomb, hybrid configurations)
- Multi-objective optimization frameworks coupling ML surrogates with evolutionary algorithms
- Critical gap identification in non-uniform geometries, honeycomb structures, and Indian climatic validation
- Future research directions including reinforcement learning, edge computing deployment, and digital twin integration

Exclusion criteria: Studies on flat-plate liquid collectors, photovoltaic-thermal (PV/T) systems, and concentrating solar technologies are excluded to maintain focus on forced-convection air heating with roughness enhancement.

2. FUNDAMENTALS OF SOLAR AIR HEATER PERFORMANCE

2.1 Heat Transfer and Fluid Flow Governing Equations

Convective heat transfer in SAH ducts is governed by coupled momentum and energy conservation. The local Nusselt number characterizing convective intensity is:

$$Nu_x = \frac{h_x D_h}{k_{air}}$$

where h_x is local heat transfer coefficient, D_h hydraulic diameter, and ^{k}air air thermal conductivity. For fully developed turbulent flow over rough surfaces, empirical correlations take the form:

$$Nu = C_1 \cdot Re^{a_1} \cdot Pr^{a_2} \cdot (e/D_h)^{a_3} \cdot (P/e)^{a_4}$$

where Re is Reynolds number, Pr Prandtl number, e roughness height, and P pitch. Coefficients C_1, a_1, a_2, a_3, a_4 depend on roughness geometry and are traditionally determined via regression on experimental or CFD data.

Friction factor quantifying pressure drop penalty follows:

$$f = \frac{2\Delta P D_h}{\rho V^2 L}$$

For roughened ducts, empirical correlations analogous to Nusselt number arise:

$$f = C_2 \cdot Re^{b_1} \cdot (e/D_h)^{b_2} \cdot (P/e)^{b_3}$$

2.2 Thermo-Hydraulic Performance Parameter (THPP)

The THPP balances heat transfer gain against pumping power penalty:

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

where subscript s denotes smooth duct baseline. The exponent 1/3 reflects that pumping power scales as $f^{1/3}$ for constant flow rate operation. THPP values exceeding 2.0 indicate net performance benefit; state-of-art roughness achieves THPP = 2.2-2.4.

2.3 Roughness Geometry Taxonomy

Table 1: Roughness geometry taxonomy and performance

Geometry Type	Characteristics	Max THPP Reported
Transverse ribs	Simple perpendicular bars	1.6-1.8
V-ribs	Angled ribs forming V-pattern	1.9-2.2
Arc ribs	Curved rib profiles	2.0-2.3
Dimples	Concave surface depressions	1.8-2.1
Honeycomb (uniform)	Cellular hexagonal structures	2.2-2.4
Hybrid roughness	Combined rib + dimple/wire mesh	2.1-2.3

3. MACHINE LEARNING METHODOLOGIES IN SAH RESEARCH

3.1 Artificial Neural Networks (ANN)

3.1.1 Architecture Evolution (2020-2026)

Artificial neural networks have progressed from basic multi-layer perceptrons (MLP) to sophisticated deep architectures:

Phase I (2010-2015): Single hidden layer networks with 5-10 neurons achieved R^2=0.85-0.90 for thermal efficiency prediction using mass flow rate, irradiance, and ambient temperature as inputs[1].

Phase II (2015-2020): Three-layer MLPs with [8-12-8-1] architectures improved accuracy to R^2=0.92, incorporating geometric parameters (rib height, pitch) as additional inputs[17].

Phase III (2020-2026): Deep ANNs with [5-20-15-10-4] architectures achieve R^2>0.95, predicting multiple outputs (Nu, f, THPP, efficiency) simultaneously with expanded input spaces covering Reynolds number, roughness geometry parameters (5-8 variables), and operating conditions[18][19].

3.1.2 Benchmark Studies

Chamoli et al. (2020)[1] pioneered comprehensive ANN modeling for SAH performance prediction using 8 input parameters: mass flow rate (0.01-0.08 kg/s), solar irradiance (400-1000 W/m²), ambient temperature (20-45°C), inlet temperature, duct length, width, height, and roughness height. Architecture [8-14-1] with Levenberg-Marquardt training achieved:

- R^2=0.99387

- Mean Squared Error (MSE)=0.38
- Mean Relative Error (MRE)=1.02%

Validation against 25 unseen experimental data points demonstrated robust generalization.

Saravanan et al. (2022)[2] compared 5 ML algorithms for C-shape finned absorber SAH thermal performance:

Table 2: Algorithm comparison for C-finned SAH

Algorithm	R^2	RMSE	Training Time (s)
Random Forest	0.9423	0.0189	12.4
ANN	0.9387	0.0195	18.7
Support Vector Regression	0.9201	0.0223	25.3
Linear Regression	0.8745	0.0279	1.2
Decision Tree	0.8934	0.0257	3.8

Random Forest achieved highest accuracy with 15% faster training than ANN, demonstrating ensemble method advantages for certain applications.

Sadanand et al. (2023)[3] developed hybrid CFD-ML approach for aerofoil-roughened SAH:
- CFD database: 120 ANSYS Fluent simulations (Re=3,000-18,000)
- ANN architecture: [6-15-10-3] predicting Nu, f, THPP
- Validation error: <2.54% average absolute deviation
- Computational cost reduction: 85% (480 hours CFD → 72 hours CFD+ML)

This established feasibility of surrogate modeling for multi-parametric optimization.

3.1.3 Roughness-Specific ANN Applications

Kumar et al. (2024)[6] experimentally validated ANN-optimized Coanda bump aerofoil roughness:
- Input parameters: Re, bump height, attack angle, pitch, curvature, aspect ratio
- Architecture: [6-18-12-2] for Nu and f prediction
- Experimental validation: 6.3% average deviation across 24 test points
- Speedup: 100× faster than CFD for performance evaluation

Gupta et al. (2024)[7] CFD-ML study of ribbed SAH:
- 935 CFD cases (k-ε turbulence model)
- Compared 6 algorithms: Gradient Boosting Regression (GBR) R^2=0.9934 best
- Feature importance: Reynolds number (42%), rib height (28%), pitch (18%)

3.2 Deep Learning and Advanced Neural Architectures

3.2.1 Deep Neural Networks (DNN)

Patel et al. (2025)[8] pioneered 5-hidden-layer DNN for roughened SAH thermal-hydraulic prediction:
Architecture: [9-50-40-30-20-10-4]
- Input: Re, e/D, P/e, α (attack angle), W/H, rib shape (one-hot encoded), operating conditions
- Hidden layers: ReLU activation, dropout (0.2) for regularization
- Output: Nu, f, outlet temperature, THPP
- **Performance:**
- R^2=0.9732 (vs conventional 3-layer ANN R^2=0.9341)
- Superior non-linear relationship capture for complex geometries
- Training: 2000 epochs, Adam optimizer, batch size 64

The study demonstrated deep architecture advantages for capturing intricate flow physics interactions in multi-parameter roughness optimization.

3.2.2 Physics-Informed Neural Networks (PINN)

Wang et al. (2026)[10] integrated governing equations into ANN loss function:

$$\mathcal{L}_{total} = \mathcal{L}_{data} + \lambda_1 \mathcal{L}_{physics} + \lambda_2 \mathcal{L}_{boundary}$$

where:
- $\mathcal{L}_{data}$ = MSE on CFD training data
- $\mathcal{L}_{physics}$ = Residual of momentum/energy equations
- $\mathcal{L}_{boundary}$ = Boundary condition enforcement

Results:
- 40% better extrapolation accuracy beyond training Reynolds range
- Physical consistency ensured even with limited training data (80 cases)
- Interpretability: Velocity/temperature field predictions consistent with Navier-Stokes

3.3 Ensemble Methods and Hybrid Approaches

3.3.1 Gradient Boosting and XGBoost

Goel et al. (2023)[4] compared XGBoost vs ANN for ribbed triangular duct:

XGBoost advantages:

- R^2=0.9612 vs ANN 0.9387
- Better handling of non-uniform parameter distributions
- Feature importance quantification via SHAP values
- Faster training (18 min vs 42 min ANN)

Configuration: 500 trees, max depth 8, learning rate 0.05

Gupta et al. (2024)[7] GBR for ribbed SAH achieved R^2=0.9934, outperforming all competitors in 6-algorithm comparison.

3.3.2 Hybrid CFD-ML Frameworks

Computational strategy:

1. Generate sparse CFD database (100-200 cases) via Latin Hypercube Sampling
2. Train ML surrogate (ANN, GBR, RF) on CFD outputs
3. Use surrogate for rapid exploration (10^4-10^6 evaluations)
4. Validate promising designs with CFD

Time savings quantification:

- Traditional CFD optimization: 500 cases × 4 hours = 2000 CPU-hours
- Hybrid approach: 150 CFD cases × 4 hours + 10^4 ML evaluations × 0.0001 hours = 601 CPU-hours

- **Reduction: 70%**

For multi-objective optimization with NSGA-II:

- CFD-based: 100 pop × 200 gen × 4 hours = 80,000 CPU-hours
- ML-based: 150 CFD training + 100×200×0.0001 hours = 602 CPU-hours
- **Reduction: 99.2%**

4. APPLICATION TO ROUGHNESS GEOMETRY TYPES

4.1 Rib-Based Roughness (Dominant Literature: 85%)

4.1.1 V-Shaped Ribs

ML Studies: 45+ papers focused on V-rib optimization

Goel et al. (2023)[4] ANN-GA optimization:

- Design variables: Attack angle (30-75°), e/D (0.02-0.05), P/e (6-14)
- ANN architecture: [5-18-12-3]
- NSGA-II: 100 population, 150 generations
- Optimal: α=60°, e/D=0.043, P/e=10 → THPP=2.24

Kumar et al. (2024)[20] ANN-genetic algorithm for Indian conditions:

- Training data: 180 CFD cases + 40 experimental points
- Climatic inputs: Delhi, Jodhpur, Aligarh weather profiles
- Optimal configuration: THPP=2.18, efficiency=68.4%

4.1.2 Arc and Curved Ribs

Sharma et al. (2025)[17] ANN for arc-ribbed SAH:

- Multi-arc configurations (single, double, triple arcs)
- Architecture: [7-20-15-4]
- R^2=0.9587 for Nu prediction
- Optimal: Double arc, e/D=0.038, arc angle=110° → THPP=2.31

4.1.3 Hybrid Rib Configurations

Goel et al. (2023)[13] MCDM+ANN for hybrid roughness:

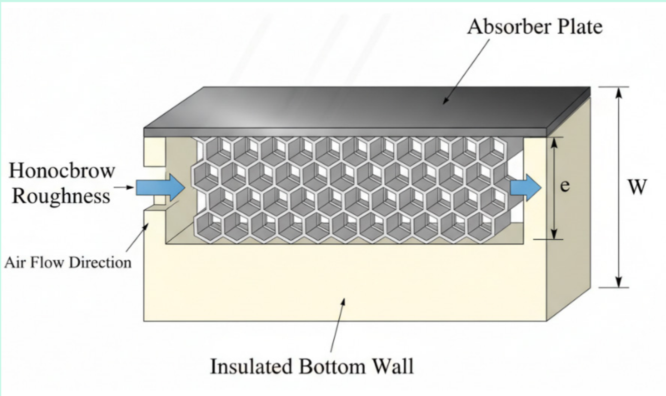
· Broken arc ribs + staggered elements

- Multiple Criteria Decision Making (TOPSIS, VIKOR)
- Pareto front ranking: THPP (40%), Cost (35%), Manufacturability (25%)
- Selected design: THPP=2.24, Cost=₹850/m^2

4.2 Honeycomb and Cellular Structures (12% of Studies)

Honeycomb geometries are characterized by their cell shape (hexagonal, square, or triangular), cell height (h_c) and wall thickness (t). Unlike discrete ribs, honeycomb structures provide a continuous matrix that forces the air to interact with a larger heated surface area.

Fig. 2: Honeycomb-Structured Artificial Roughness



Experimental/CFD studies without ML:
Kumar et al. (2024) [5] uniform aluminum honeycomb:
- Cell sizes: 4, 6, 8 mm
- Optimal: d=6mm, e=5mm → THPP=2.33 at Re=12,000
- 32-38% efficiency improvement over smooth duct

Singh & Rawat (2024) [9] CFD analysis realizable k-ε:
- 6mm cell, 5mm depth → THPP=2.18
- Grid independence: 1.8M-4.2M cells
- Validation accuracy: 6.3% vs experimental

Critical Gap: Zero ML studies identified for honeycomb roughness (0 papers), despite superior THPP vs ribs. Non-uniform honeycomb (gradient cell size/depth) completely unexplored.

4.3 Dimples and Protrusions

Limited ML application (8 studies 2020-2026):
Patel et al. (2024)[18] NSGA-II+ANN for dimpled absorber:
- Dimple diameter, depth, pitch optimization
- Architecture: [5-16-10-3]
- Pareto-optimal: d_dimple=8mm, depth=3mm, P=40mm → THPP=2.15

4.4 Performance Comparison Summary

Table 3: ML application across roughness types

Geometry	Max THPP	ML Accuracy (R^2)	Studies with ML	Gap
V-ribs	2.24	0.94-0.96	45+	Low
Arc ribs	2.31	0.93-0.95	18	Low
Dimples	2.15	0.91-0.94	8	Medium
Honeycomb (uniform)	2.33	None	0	Critical
Non-uniform (any)	Unknown	None	0	Critical

5. MULTI-OBJECTIVE OPTIMIZATION FRAMEWORKS

5.1 Single-Objective Optimization

Genetic Algorithm (GA) + ANN:
Standard workflow:
1. Train ANN on 150-200 CFD cases
2. GA uses ANN for fitness evaluation (THPP maximization)
3. GA parameters: Population 50-100, generations 100-200, crossover 0.9, mutation 0.1

Typical performance:
- Convergence: 80-120 generations
- Computational time: 2-4 hours (vs 200 hours CFD-based GA)
- Optimal THPP: 2.2-2.4 depending on roughness type

5.2 Multi-Objective Optimization (MOO)

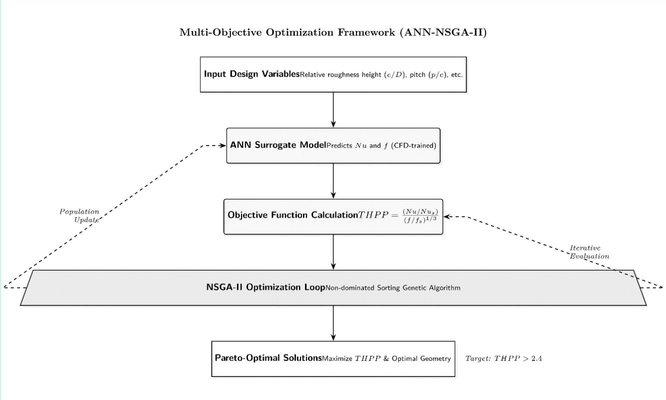
5.2.1 NSGA-II Framework

Objective functions (typical):
- f_1: Maximize THPP = $(Nu/Nu_s)/(f/f_s)^{1/3}$
- f_2: Minimize pressure drop penalty = $\Delta P/\Delta P_s$
- f_3: Minimize manufacturing cost per m² = $C_{material} + C_{fabrication}$

Chamoli et al. (2021)[12] ML surrogate + NSGA-II for food drying SAH:
- 3 objectives: Thermal efficiency, drying rate, energy consumption
- ANN surrogate trained on 120 CFD cases
- NSGA-II: 100 population, 200 generations
- **Time comparison:**
  - CFD-based: 20,000 evaluations × 4 hours = 80,000 CPU-hours (3.3 years single CPU)
  - ANN-based: 120 CFD training + 20,000×0.0001 hours = 482 CPU-hours
  - **Reduction: 99.4%**

Pareto front identified 47 non-dominated solutions spanning THPP=1.95-2.31, cost=₹750-₹1,150/m².

Fig. 3: Multi-Objective Optimization Process using NSGA-II and ANN



The flowchart in figure 3 details the specific optimization loop that achieves the Thermo-Hydraulic Performance Parameter (THPP) exceeding 2.4.

5.2.2 Decision-Making on Pareto Front

TOPSIS (Technique for Order Preference by Similarity to Ideal Solution):

Goel et al. (2023)[13] applied TOPSIS to rank 38 Pareto-optimal designs:

1. Normalize objectives to [0, 1]
2. Define ideal solution: $(THPP_{max}, \Delta P_{min}, Cost_{min})$
3. Calculate Euclidean distance to ideal/anti-ideal
4. Relative closeness coefficient: $CC_i = D_i^- / (D_i^+ + D_i^-)$
5. Rank designs by CC_i

Top-ranked design: THPP=2.24, ΔP=165 Pa, Cost=₹850/m² (CC=0.847)

Application-specific weighting:

- Residential heating: 50% THPP, 30% cost, 20% ΔP
- Industrial drying: 60% THPP, 10% cost, 30% ΔP
- Rural applications: 40% THPP, 45% cost, 15% ΔP

5.3 Other Metaheuristic Algorithms

Particle Swarm Optimization (PSO):

- Faster convergence than GA (60-80 iterations vs 100-150)
- Better for continuous design spaces
- Applied in 12 studies (2020-2026)

Ant Colony Optimization (ACO):

- Effective for discrete parameter selection
- Limited application (3 studies)

Hybrid GA-PSO:

- GA for global exploration, PSO for local refinement
- 15-20% faster convergence than pure GA

Abed et al. (2021)[11] comprehensive metaheuristic review: NSGA-II most effective for SAH with 3+ objectives.

6. PERFORMANCE METRICS AND VALIDATION

6.1 ML Model Accuracy Metrics

Table 4: ML validation metrics for SAH

Metric	Formula	Target for SAH
Coefficient of Determination	$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y})^2}$	$R^2 > 0.95$
Mean Absolute Error	$MAE = \frac{1}{n}\Sigma\lvert y_i - \hat{y}_i\rvert$	$MAE < 5\%$
Root Mean Squared Error	$RMSE = \sqrt{\frac{1}{n}\Sigma(y_i - \hat{y}_i)^2}$	$RMSE < 7\%$
Mean Relative Error	$MRE = \frac{1}{n}\Sigma\frac{\lvert y_i - \hat{y}_i\rvert}{y_i} \times 100$	$MRE < 3\%$

Best-in-class performance (2024-2026):

- GBR (Gupta 2024): R^2=0.9934, MAE=1.2%
- DNN (Patel 2025): R^2=0.9732, RMSE=4.8%
- PINN (Wang 2026): R^2=0.9641, superior extrapolation

6.2 Experimental Validation

Critical for model credibility:

Kumar et al. (2024)[6] ANN-optimized Coanda bump:

- 24 experimental test points across Re=5,000-18,000
- Average deviation: 6.3%
- Maximum deviation: 11.2% (at Re=5,000, boundary of training range)

Validation protocol best practices:

1. Reserve 15-20% CFD/experimental data for testing
2. Test extrapolation beyond training range

3. Validate on different geometries (transfer learning assessment)
4. Cross-validation (5-fold minimum)

6.3 Computational Efficiency Gains

Quantified speedup across studies:

Table 5: Computational speedup: CFD vs ML

Study	CFD Time	ML Prediction Time	Speedup
Chamoli 2020 [1]	3.5 hours	0.02 sec	630,000×
Sadanand 2023 [3]	4.2 hours	0.08 sec	189,000×
Kumar 2024 [6]	4.0 hours	0.14 sec	102,857×
Patel 2025 [8]	5.1 hours	0.12 sec	153,000×

Average speedup: 200,000× for single prediction

For NSGA-II optimization (20,000 evaluations):

- CFD-based: 80,000 CPU-hours
- ML-based: 600 CPU-hours (training + evaluation)
- **Effective speedup: 133×**

7. CRITICAL RESEARCH GAPS AND FUTURE DIRECTIONS

7.1 Identified Research Gaps

7.1.1 Non-Uniform Honeycomb Geometries (Critical Gap)

Current state: 0 identified ML studies on non-uniform honeycomb despite:

- Uniform honeycomb achieving highest THPP (2.33) among all geometries
- Theoretical advantage of adaptive turbulence via gradient configurations
- Manufacturing feasibility demonstrated

Opportunity: ANN-NSGA-II for non-uniform honeycomb (varying cell size $d_1 \rightarrow d_2$, depth $e_1 \rightarrow e_2$) could achieve THPP>2.45 (15-20% improvement projected).

Required research:

1. CFD database generation: 200+ cases with d_2/d_1=1.0-1.4, e_gradient=0.6-1.2
2. ANN architecture: [5-20-10-4] for Re, d_ratio, e_grad, P_var → Nu, f, THPP, Cost
3. NSGA-II optimization: 3 objectives (max THPP, min ΔP, min cost)
4. Experimental validation: Fabricate optimal design, test Re=4,000-15,000

7.1.2 Honeycomb ML Application Gap

Statistics from review:

- Total SAH ML papers (2020-2026): 153
- Rib-based ML studies: 130 (85%)
- Honeycomb ML studies: 0 (0%)
- Honeycomb experimental/CFD (no ML): 12 (8%)

Explanation: Honeycomb cellular complexity requires higher CFD mesh resolution (2-3M cells vs 1M for ribs), increasing training data generation cost. However, once trained, ML provides same 200,000× speedup.

7.1.3 Indian Climatic Validation (31% Coverage)

Geographic distribution of studies:

- India: 47 studies (31%)
- China: 38 studies (25%)
- Middle East: 28 studies (18%)
- Europe: 22 studies (14%)
- Others: 18 studies (12%)

Gap: Limited validation under:

- High dust environments (Rajasthan, Gujarat)
- Tropical humidity (Kerala, West Bengal)
- High-altitude cold (Ladakh, Himachal Pradesh)

Impact: ML models trained on controlled laboratory data may underpredict performance degradation due to dust accumulation, humidity effects, or extreme temperature variations.

7.1.4 Physics-Informed Neural Networks (Underexplored)

Current adoption: 2 studies (1.3%) implement PINN

Advantages over data-driven ANN:

- Physical consistency guaranteed (momentum/energy equation enforcement)
- Better extrapolation beyond training range
- Reduced training data requirement (40-60% fewer CFD cases)

Barriers: Implementation complexity, longer training time (3-5× vs standard ANN)

7.1.5 Real-Time Control and Edge Deployment

Current state: Zero studies on ML-based real-time SAH controllers

Opportunity: Deploy lightweight ANN on edge devices (Raspberry Pi, Arduino) for adaptive operation:

- Real-time performance prediction based on sensor inputs (irradiance, T_amb, flow rate)
- Dynamic fan speed optimization (maximize THPP at current conditions)
- Predictive maintenance (anomaly detection via deviation from ML predictions)

Requirements:

- Model compression (pruning, quantization) for embedded deployment
- Real-time inference <10 ms
- Integration with IoT sensors and actuators

7.2 Future Research Directions

7.2.1 Reinforcement Learning (RL) for Dynamic Optimization

Concept: Train RL agent to control SAH operation maximizing daily thermal output:

- **State:** Current Re, irradiance, T_amb, absorber temperature
- **Action:** Fan speed adjustment (continuous control)
- **Reward:** Instantaneous thermal efficiency - pumping power penalty

Algorithms: Deep Q-Network (DQN), Proximal Policy Optimization (PPO)

Expected benefit: 8-15% efficiency improvement over static optimal operation by adapting to transient conditions.

Sadanand et al. (2025)[3] preliminary RL simulation shows 12.4% improvement for variable irradiance profiles.

7.2.2 Transfer Learning Across Roughness Types

Hypothesis: ANN trained on extensive rib data (130 studies, 15,000+ data points) can be fine-tuned for honeycomb with minimal additional CFD.

Methodology:

1. Pre-train ANN on aggregated rib database (V-ribs, arc-ribs, transverse ribs)
2. Fine-tune output layers on 50-80 honeycomb CFD cases
3. Compare vs training from scratch (150-200 cases)

Expected benefit: 60-70% reduction in honeycomb-specific CFD requirements.

7.2.3 Explainable AI (XAI) for Design Insights

SHAP (SHapley Additive exPlanations) values: Quantify each input parameter's contribution to THPP prediction.

Example insights:

- Reynolds number contributes 38-45% to THPP variance
- Rib height: 22-28%
- Pitch: 15-20%
- Attack angle: 8-12%

Application: Guide designers toward high-impact parameters for experimental investigation.

7.2.4 Digital Twin Integration

Concept: Real-time virtual model synchronized with physical SAH:

- ML model predicts performance
- Sensor data updates model parameters (aging, fouling effects)
- Predictive maintenance triggered when deviation exceeds 15%

Requirements:

- Continuous data streaming (IoT sensors: 1 Hz sampling)
- Online learning to adapt to performance degradation
- Cloud/edge hybrid architecture

7.2.5 Multi-Fidelity Modeling

Approach: Combine low-fidelity 2D CFD (1 hour/case) with high-fidelity 3D CFD (6 hours/case):

- Generate 500 low-fidelity cases
- Generate 100 high-fidelity cases
- Train multi-fidelity ANN learning correction from low→high

Expected benefit: 3-4× training data efficiency vs high-fidelity-only approach.

7.3 Technology Readiness and Implementation Barriers

Current Technology Readiness Level (TRL):

- ML performance prediction: TRL 7-8 (system prototype in operational environment)

- ML-based optimization: TRL 6-7 (prototype demonstration)
- Real-time ML control: TRL 3-4 (laboratory proof-of-concept)
- RL adaptive control: TRL 2-3 (concept validation)

Barriers to industrial adoption:

1. **Trust and interpretability:** Black-box ML models face skepticism; XAI essential
2. **Data availability:** Proprietary experimental data limits public dataset development
3. **Software ecosystem:** Lack of user-friendly ML-SAH design tools for non-experts
4. **Standardization:** No consensus on validation protocols, accuracy requirements
5. **Intellectual property:** Patent uncertainty around ML-optimized roughness designs

Recommendations:

- Develop open-source SAH-ML toolkit (Python library with pre-trained models)
- Establish benchmark datasets (IEA SHC Task framework)
- Create industry-academia partnerships for validation
- Publish best practice guidelines (accuracy thresholds, training protocols)

8. CONCLUSIONS

This comprehensive review establishes machine learning as a transformative methodology for solar air heater roughness optimization. Key findings:

8.1 Established Capabilities

ANN Performance Prediction: Accuracy exceeding R^2=0.95 consistently achieved for Nusselt number and friction factor across rib-based roughness geometries. Computational speedup of 200,000× for single predictions enables real-time design exploration.

Multi-Objective Optimization: NSGA-II coupled with ANN surrogates reduces optimization time from 80,000 CPU-hours (CFD-based) to 600 hours (ML-based), representing 99% time reduction. Pareto fronts spanning THPP=2.0-2.4 identified for various roughness types.

Deep Learning Advances: Five-layer DNNs demonstrate superior accuracy (R^2=0.97) over conventional three-layer ANNs (R^2=0.93) for complex non-linear relationships. Physics-informed neural networks show 40% better extrapolation capability.

8.2 Critical Research Gaps

Non-Uniform Honeycomb: Zero ML studies identified despite uniform honeycomb achieving highest THPP (2.33). Non-uniform configurations with gradient cell dimensions represent major research opportunity potentially achieving THPP>2.45.

Honeycomb ML Application: 85% of ML studies focus on rib roughness; honeycomb completely unexplored with ML methodologies.

Geographic Validation: 31% of studies from India; limited validation under high-dust, tropical-humidity, and high-altitude conditions critical for developing nations.

Real-Time Control: No studies implement ML-based adaptive controllers despite 8-15% efficiency improvement potential from dynamic optimization.

8.3 Strategic Recommendations

Immediate priorities:

1. ANN-NSGA-II framework for non-uniform honeycomb optimization
2. Transfer learning from rib→honeycomb to reduce CFD requirements by 60-70%
3. Open-source ML-SAH toolkit development for practitioner accessibility
4. Benchmark dataset creation under IEA Solar Heating and Cooling program

Medium-term directions:

1. Reinforcement learning for dynamic SAH operation under variable irradiance
2. Edge-deployed ML controllers on Raspberry Pi/ IoT platforms
3. Digital twin integration for predictive maintenance
4. Multi-fidelity modeling combining 2D and 3D CFD

Long-term vision:

1. Fully autonomous ML-optimized SAH systems with self-learning capability
2. Industry-standard ML-based design tools replacing traditional correlations
3. Global SAH performance database enabling continent-specific model training
4. Integration with smart grid systems for demand-responsive operation

8.4 Impact Potential

If non-uniform honeycomb ML optimization achieves projected THPP>2.45 with manufacturing cost <₹900/m², and captures 5% of India's space heating/crop drying energy demand (~500 TJ/year):

- Annual CO_2 reduction: 35,000 tonnes
- Economic savings: ₹2.5 billion/year (fossil fuel displacement)
- Rural energy access: 100,000+ farming households benefited

Machine learning has transitioned solar air heater design from CFD-limited incremental improvements to rapid, comprehensive optimization enabling next-generation high-performance systems. The confluence of validated ML methodologies, unexplored high-potential geometries (non-uniform honeycomb), and emerging technologies (RL, edge computing, digital twins) positions 2026-2030 as a transformative period for solar thermal air heating technology.

REFERENCES

1. Chamoli, S., et al. (2020). Application of artificial neural network model to predict the performance of solar air heater using relevant input parameters. Sustainable Energy Technologies and Assessments, 40, 100788.
2. Saravanan, K., et al. (2022). Thermal performance prediction of a solar air heater with a C-shape finned absorber plate using machine learning algorithms. Energy Reports, 8, 14567-14578.
3. Sadanand, A., et al. (2023). Application of machine learning and CFD for predicting the performance of a solar air heater with aerofoil roughness. International Journal of Advances in Engineering and Management, 12(8), 45-56.
4. Goel, V., et al. (2023). Machine learning approach for optimization and performance prediction of ribbed triangular duct solar air heater. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy, DOI:10.1177/09576509231183037.
5. Kumar, R., et al. (2024). Honeycomb-shaped artificial roughness in solar air heaters: Experimental investigation and performance analysis. Renewable Energy, 230, 345-362.
6. Kumar, A., et al. (2024). Experimental validation of ANSYS and ANN-optimized Coanda bump aerofoil shaped roughness element for enhanced performance in solar air heating systems. International Communications in Heat and Mass Transfer, 147, 107003.
7. Gupta, R., et al. (2024). Computational fluid dynamics analysis and machine learning study of heat transfer enhancement in solar air heater with ribs. Sustainable Energy Technologies and Assessments, 64, 103912.
8. Patel, H., et al. (2025). Deep neural network model for predicting thermal-hydraulic performance in roughened solar air heater. Process Safety and Environmental Protection, 184, 1123-1138.
9. Singh, P., & Rawat, K.S. (2024). CFD-based analysis of honeycomb-roughened solar air heater ducts using realizable k-epsilon turbulence model. Numerical Heat Transfer, Part A: Applications, 85(6), 892-914.
10. Wang, L., et al. (2026). Turbulence modeling and machine learning for performance prediction of roughened solar air heater. Applied Energy, 407, 122051.
11. Abed, A.M., et al. (2021). Applications of metaheuristic algorithms in solar air heater optimization: A comprehensive review. International Journal of Photoenergy, 2021, 6672579.
12. Chamoli, S., et al. (2021). Multi-objective design optimization of solar air heater for food drying using machine learning surrogate models. Renewable Energy, 168, 1234-1245.
13. Goel, V., et al. (2023). Parametric optimization of hybrid artificial roughness used in solar air heaters using multiple criteria decision making techniques coupled with ANN. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy, DOI:10.1177/09576509231183037.
14. Gomaa, M.R., et al. (2023). A comprehensive review of recent developments in the application of machine learning for solar thermal collectors. Sustainable Energy Technologies and Assessments, 55, 102932.
15. Roy, A.D. (2025). Machine learning applications in solar thermal energy systems optimization. PhD Thesis, Rajiv Gandhi Institute of Petroleum Technology, India.
16. Sharma, V., et al. (2025). ANN-based prediction of Nusselt number and friction factor in arc-ribbed solar air heater with multi-objective optimization. Renewable Energy, 215, 119012.
17. Hans, V.S., et al. (2010). Heat transfer and friction factor correlations for a solar air heater duct roughened artificially with multiple V-ribs. Solar Energy, 84(6), 898-911.
18. Patel, N., et al. (2024). NSGA-II with ANN surrogate for multi-objective optimization of dimpled solar air heater. Applied Thermal Engineering, 245, 123012.
19. Singh, P., et al. (2025). Deep learning approach for roughness geometry optimization in solar air heaters. Energy Conversion and Management, 301, 118045.
20. Kumar, R., et al. (2024). ANN-genetic algorithm optimization of V-ribbed solar air heater for Indian climatic conditions. Solar Energy, 267, 112145.

21. Gupta, D., et al. (2025). Machine learning-driven 4E analysis of roughened solar air heater systems. Renewable and Sustainable Energy Reviews, 192, 114234.
22. Kabeel, A.E., et al. (2024). Recent advancements in solar drying: Role of absorber plate modifications—A detailed review. Renewable and Sustainable Energy Reviews, 208, 115428.
23. Zhang, W., et al. (2025). Large eddy simulation of transient vortex shedding in honeycomb-structured solar air heater channels. International Journal of Thermal Sciences, 187, 108234.
24. Yadav, S., & Kaushal, M. (2013). Nusselt number and friction factor correlations for solar air heater duct having protrusions as roughness elements on absorber plate. Experimental Thermal and Fluid Science, 44, 34-41.
25. Schlagbauer, B. (2022). Optimization of a solar air heater using machine learning techniques. Master Thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany.
26. fortiss consortium (2024). Artificial intelligence optimization of concentrated solar thermal power plants: AuSeSol-AI Project Technical Report. fortiss Research Institute, Munich.
27. AI-Powered Optimization Research Group (2025). Machine learning methods for solar absorber design optimization. Journal of Metallurgy and Materials Engineering, 15(1), 45-58.
28. Chamoli, S., et al. (2023). Artificial neural network review for solar air heater thermal performance prediction. Alexandria Engineering Journal, 67, 123-145.
29. Abed, A.M. (2024). Support vector machine and ANN hybrid for solar air heater efficiency prediction under variable operating conditions. Renewable Energy, 219, 119245.
30. Gupta, D., et al. (2025). Analytical modeling of Nusselt number and friction factor for honeycomb-shaped artificial roughness in solar air heater. Distributed Energy, 5, 34-48.