

Machine Learning Approaches for Predicting and Optimising Thermal Performance of Solar Collectors

Balakrishnan Varun Kumar^a, Parthasarathy Rajesh Kanna^b, Chithirai Pon Selvan^c, Dawid Taler^d, Tomasz Sobota^d, Jan Taler^{e*}

^aThiagarajar College of Engineering, Madurai 625 015, Tamil Nadu, India

^bCO₂ Research and Green Technologies Centre, Vellore Institute of Technology, Vellore 632014, Tamil Nadu, India

^cSchool - Science and Engineering, Curtin University, Dubai, United Arab Emirates

^dCracow University of Technology, Faculty of Environmental Engineering and Energy, Department of Thermal Processes, Air Protection and Waste Utilization, Warszawska 24, 31-155 Cracow, Poland

^eCracow University of Technology, Faculty of Environmental Engineering and Energy, Department of Energy, Warszawska 24, 31-864 Cracow, Poland

*Corresponding author email: jan.taler@pk.edu.pl

Received: 08.08.2025; revised: 19.01.2026; accepted: 23.01.2026

Abstract

The application of machine learning techniques in the renewable energy sector has revolutionised the predictive modelling and optimisation of solar thermal systems. Solar air heaters with artificial roughness and phase change materials are widely used to enhance thermal performance, necessitating accurate thermo-hydraulic performance predictions for efficient design and operation. This study employs the XG-Boost machine learning model to predict thermo-hydraulic performance values based on experimental data obtained for polygonal and trapezoidal rib configurations at different relative roughness pitch ratios (p/e) and Reynolds numbers ($Re = 2000$ to 20000) in the absorber plate of a solar air heater. The experimental results reveal that a polygonal rib configuration with $p/e = 7.5$ exhibits the highest thermo-hydraulic performance value of 2.95 at $Re = 4000$, while the trapezoidal rib with $p/e = 7.5$ achieves a thermo-hydraulic performance of 2.93 at the same Reynolds number. The results were trained in the proposed machine learning model to validate the accuracy of predicted results. The coefficient of determination R^2 , mean absolute percentage error, root mean squared error and mean absolute percentage error matrix were considered for the training and testing dataset. The model effectively captures the nonlinear thermal behaviour, achieving $R^2 = 0.976$, mean absolute error at 0.045, root mean squared error at 0.056 and mean absolute percentage error of 2.42%, demonstrating superior predictive capability. Further, this study provides the significance of plethora machine learning algorithms, data pre-processing strategies, feature selection and hybrid machine learning models for performance optimisation in solar air heaters. Additionally, few case studies highlight the practical implementation of machine learning in solar energy systems, demonstrating its potential for accurate prediction of thermal efficiency in real time assessment. Besides, it outlines future research directions, challenges, and emerging opportunities to advance machine learning applications in sustainable energy solutions. This „review-experimental” comparative article serves as a valuable resource for researchers and industry professionals, driving innovation in solar thermal energy systems through data-driven intelligence.

Keywords: Machine learning; Thermal performance; Solar collectors; Prediction models; Optimisation techniques; Renewable energy systems

Vol. 47(2026), No. 1, 5–23; doi: 10.24425/ather.2025.156852

Cite this manuscript as: Kumar, B.V., Kanna, P.R., Selvan, C.P., Taler, D., Sobota, T., & Taler, J. (2026). Machine Learning Approaches for Predicting and Optimising Thermal Performance of Solar Collectors. *Archives of Thermodynamics*, 47(1), 5–23.

1. Introduction

The abundant availability of solar radiation globally makes solar heating a highly viable and sustainable option for various appli-

cations in both domestic and industrial settings. With ongoing advancements and supportive policies, solar thermal energy is poised to play a significant role in meeting the world's heating needs while promoting environmental sustainability. Devices

Nomenclature

f – friction factor for the roughened configurations
 f_0 – friction factor for the smooth configurations
 Nu – Nusselt number for the roughened configurations
 Nu_0 – Nusselt numbers for the smooth configurations
 p/e – relative roughness parameter
 R^2 – coefficient of determination
 Re – Reynolds number
 T_k – tree-based transformations learned by XG-boost

Greek symbols

α_k – model coefficients
 β – bias term
 η_{sc} – solar collector efficiency

Abbreviations and Acronyms

ALE – accumulated local effects
 ANFIS – adaptive network-based fuzzy inference system
 ANN – artificial neural networks
 AT – ambient temperature
 BNN – Bayesian neural networks
 BPS – black pebble stone
 BRT – boosted regression tree
 CAPSC – corrugated absorber plate solar collector
 CE – collector efficiency
 CFD – computational fluid dynamics
 CIT – collector inlet temperature
 CNN – convolutional neural networks
 COT – collector outlet temperature
 CSP – concentrated solar power
 DBN – deep belief networks
 DCIT – drying chamber inlet temperature
 DCOT – drying chamber outlet temperature
 DT – decision trees
 ELM – extreme learning machine
 ENN – Elman neural network
 ESE – energy storage element
 FFS – forward feature selection
 FPSC – flat plate solar collectors
 GA – genetic algorithms
 GBM – gradient boosting machines
 GBRT – gradient boosting regression tree
 GMDH – group method of data handling
 GP – Gaussian processes
 HE-ET – heat exchanger-evacuated tube

HGBR – hist gradient boosting regressor
 HTF – heat transfer fluid
 KNN – K-nearest neighbours
 LR – linear regression
 LSSVM – least squares support vector machines
 LSTM – long short-term memory
 MAE – mean absolute error
 MAPE – mean absolute percentage error
 MARS – multivariate adaptive regression splines
 MC – moisture content
 MWCNT – multi-wall carbon nanotube
 MIM – metal insulator metal
 ML – machine learning
 MLR – multiple linear regression
 MRMR – minimum redundancy – maximum relevance
 MSE – mean squared error
 NARX – nonlinear autoregressive exogenous
 NF RV – nanofluid relative viscosity
 NSE – Nash-Sutcliffe coefficient
 OTA – optimal tilt angle
 PCC – Pearson correlation coefficient
 PCM – phase change materials
 PDC – parabolic dish collector
 PDP – partial dependence plot
 PPO – proximal policy optimisation
 PR – polynomial regression
 PSO – particle swarm optimisation
 PTC – parabolic trough solar collector
 PVTC – photovoltaic thermal collector
 PVT – photovoltaic thermal
 RF – random forest
 RMSE – root mean squared error
 RNN – recurrent neural networks
 SAH – solar air heater
 SB – shot-basting nano-coating
 SR – stacking regressor
 STC – solar thermal collectors
 SVR – support vector regression
 SWCNT – single-wall carbon nanotube
 SWH – solar water heater
 TES – thermal energy storage
 THP – thermo-hydraulic performance
 TR – tweedie regressor
 WI – Willmott's index
 XG-boost – extreme gradient boosting

such as flat plate collectors (FPC), evacuated tube collectors (ETC) and concentrated solar power (CSP) were used in practice to generate heat energy for applications like drying agricultural products, providing hot water, cost-effective heating rooms during colder climates, generating high temperatures for industrial applications, etc. However, the performance of thermal efficacy in these collectors is very low owing to smooth surface absorber plates and limitation of the direct contact between the working fluid and the heating surface.

It needs design modification to augment the convective heat transfer between the working fluid and solid surface, like an absorber plate placed directly beneath of the solar radiation sur-

face. Many researchers numerically and experimentally proved that implementing artificial rough surfaces in the absorber plates as an extension material acts to enhance the heat transfer rate. Rib or groove shapes such as broken arc rib [1], reverse L shape [2], S shape [3], W shape [4], discrete V-shaped rib [5], triangular rib and groove shaped rib [6], protruded shape [7], transverse square rib [8], discrete multi V rib in a staggered position [9], square duct with the internal surface [10] were positioned in the absorber plate with the varying pitch distance between ribs, varying height of rib, varying angle during positioning. Multiple shaped ribs arrangements in inline and staggered methods were investigated to ascertain the augmentation of thermal perfor-

mance. The research results reveal a significant improvement attained, such as thermo-hydraulic performance (THP) values having increased from 1.32 to a maximum of 1.89, specifically in the backwards trapezoidal-shaped rib in the solar collector.

Furthermore, to enhance heat transfer efficiency, nanofluids and hybrid nanofluid-based working fluids were introduced in solar collectors to maximise heat absorption during peak solar hours (11:00 AM to 3:00 PM), thereby optimising heat energy production for various heating applications. Hybrid nanofluids such as $\text{TiO}_2\text{-SiO}_2$ [11], Single-Wall Carbon Nano Tube (SWCNT), Multi-Wall Carbon Nano Tube (MWCNT) [12], MWCNT- Fe_3O_4 /Water [13], $\text{Al}_2\text{O}_3\text{-CuO}$, $\text{TiO}_2\text{-Cu}$, Ag-MgO [14], MWCNT, Au-Cu/EO (EO – engine oil), and Cu- Al_2O_3 [15] were prepared in varying volume fractions (1%, 2%, 3%, and 4%) and utilised as working fluids in solar collectors integrated with artificial rough surfaces to enhance heat energy generation. Both numerical and experimental investigations revealed significant thermal performance improvements, ranging from 197% to 227%, depending on the nanofluid volume fraction used.

In continuation of the introduction of nanofluids, phase change materials are introduced to enhance heat transfer in solar applications. Both organic and inorganic phase change materials (PCMs), such as paraffin wax and high-latent-heat storage oils, were integrated into the solar collector. This integration enhances daytime heat generation while enabling the storage of thermal energy in high-latent-heat materials. The stored energy is then released during nighttime, facilitating prolonged heat distribution and ensuring consistent thermal performance over extended durations.

The study conducted by Kabeel et al. [16,17] investigated the performance of an FPC integrated with paraffin wax based phase change material (PCM) arranged in two different layers (2 cm and 4 cm) beneath the absorber plate in the heat storage chamber. PCM stored thermal energy during the daytime and released it at night. The results indicated that the 4 cm PCM layer maintained a temperature rise of 1.5°C to 7°C for 3.5 hours, while the 2 cm layer sustained a temperature increase of 1.5°C to 5.5°C for 2.5 hours after 6 pm. Additionally, the incorporation of fin-shaped ribs in the same setup further enhanced thermal performance, achieving a temperature rise of 8.6°C during nighttime.

Aymen et al. [18] investigated a rectangular cavity with PCM in a collector and reported that the maximum night-time temperature increased by $3\text{--}7^\circ\text{C}$, leading to a significant improvement in thermal efficiency. Sunil Raj and Eswarmorthy [19] studied a V-shaped rib in a collector with paraffin wax-PCM blended with 10% of Al_2O_3 . The research proved a 65% enhancement in thermal efficiency. Kalaiarasi et al. [20] investigated experimentally using a black copper tube with synthetic therminol-55 oil as the thermal energy storage (TES) system and reported achieving 67.7% thermal efficiency. It has been observed that enhancing the thermal performance of solar collectors involves employing various rib shapes with numerous modifications, such as varying pitch distances, rib heights and inclined rib positions. These design optimisations, when integrated with PCM, have been explored extensively in studies

[21–23] to improve heat transfer efficiency and energy storage capabilities.

A comprehensive literature review reveals that although significant efforts have been devoted to enhancing solar collector performance through design modifications, hybrid nanofluids, and PCM integration, the novelty of the present work lies in emphasising machine learning-based frameworks as an integrated solution for accurate thermal performance prediction, adaptive control of operating parameters, predictive maintenance and sustainability assessment under dynamic climatic conditions. The primary objective of this study is to critically evaluate machine learning (ML) methodologies applied to solar collectors, identify dominant thermo-hydraulic and environmental parameters influencing performance, assess ML capabilities in fault detection, energy output forecasting, and techno-economic optimisation, and to highlight research gaps for the development of intelligent, reliable and high-efficiency solar thermal systems.

2. Machine learning algorithms in practice

To predict the thermal performance of solar collectors across diverse domestic and industrial applications, an ML algorithm plays a significant role in analysing the complexity of the experimental setup, including intricate patterns and practical key factors such as solar radiation, ambient temperature and thermal efficiency. ML models like gradient boosting machines (GBM) or neural networks (ANN) are widely used for predicting thermal efficiency and outlet temperature. Recurrent neural networks (RNN) and long short-term memory (LSTM) models are used for time-dependent variables at the hourly time scale. Bayesian models and Gaussian processes support the prediction of uncertainty in thermal performance.

Besides, ML algorithms are capable of analysing multiple parameters, such as solar intensity, water flow rate, the impact of solar collector tilt angle, ambient humidity, etc., using a randomised lasso algorithm based on real-time data recorded in a database to reduce experimental costs and save time and resources for further testing. By combining data-driven approaches with traditional thermodynamic principles, ML not only enhances the design and operation of solar thermal systems but also enables innovative solutions for achieving higher efficiency and reliability in renewable energy applications. The detailed machine learning models and algorithms that play a pivotal role in the prediction of thermal performance in solar energy-based applications are listed, and their significant applications are presented in Table 1 [24–26].

2.1. Regression techniques

They are supervised machine learning methods used to predict the relationship between dependent such as thermal performances, and independent variables, such as inlet temperature, solar intensity, mass flow rate, etc. It is classified into linear regression (LR) and polynomial regression (PR) and is suitable for analysing environmental impacts, optimising system design, forecasting time series, and reducing costs in solar thermal performance prediction.

2.2. Tree-based models

They are capable of complex interactions and non-linear relationships between input variables such as solar intensity, ambient temperature and humidity and can be categorised into:

- a) decision trees (DT),
- b) random forest (RF),
- c) gradient boosting machines (GBM) and extreme gradient

boosting (XG-boost),

- d) gradient boosting regression tree (GBRT).

These models are mostly suitable for interpretable predictions, flexibility for mixing huge data in non-linear relationships, easy scalability, and the most accurate prediction performance. GBM and GBRT models are highly selected by researchers for effective and robust performance.

Table 1. Various machine learning algorithms and their significance [24–26].

No.	Model	Applications
1	Linear Regression (LR)	Evaluates simple relationships, such as the effect of solar intensity on outlet temperature.
2	Random Forest Regression	Handles complex data to predict energy output and efficiency in hybrid systems.
3	Gradient Boosting Machines (GBM)	Highly effective for thermal performance modelling, accounting for non-linear dependencies.
4	Support Vector Regression (SVR)	Models subtle interactions between variables like flow rates and collector performance
5	Random Forest (RF)	Predicts thermal efficiency and air temperature in hybrid solar collector systems.
6	Gradient Boosting Regression Tree (GBRT)	Provides highly accurate predictions of solar collector performance and energy output.
7	Decision Trees	Helps understand the influence of specific parameters (e.g., solar intensity, humidity) on performance.
8	XGBoost	Used for optimising solar thermal storage systems (PCM) and forecasting long-term energy outputs.
9	Recurrent Neural Networks (RNNs)	Forecast energy output based on time-series weather data.
10	Autoencoders	Detect faults or anomalies by identifying deviations from normal operating patterns.
11	K-Means	Group data on solar intensity and efficiency to identify regions with similar energy potential.
12	Density-Based Spatial Clustering of Applications with Noise (DBSCAN)	Identifies anomalies in operational data, such as sudden drops in efficiency.
13	Logistic Regression	Classifies collector output as efficient or inefficient.
14	Random Forest	Categorises faults in solar systems into overheating, shading, or flow issues.
15	Bayesian Neural Network (BNN)s	Integrates uncertainty estimation in solar energy predictions.
16	Gaussian Processes (GP)	Known for their ability to predict solar energy performance with uncertainty quantification.
17	Genetic Algorithms (GA)	Applied in optimising solar system configurations.
18	Particle Swarm Optimisation (PSO)	Used for parameter optimisation in solar energy systems.
19	Bagging Methods	Aggregate multiple models to improve prediction stability and accuracy.
20	Boosting Methods	Sequentially build models to minimise prediction errors.
21	Deep Belief Networks (DBN)	Advanced neural networks for hierarchical feature extraction.
22	Autoencoders	Used for anomaly detection and feature learning in solar systems.
23	Reinforcement Learning (RL)	Applied to optimise real-time solar energy operations, such as tracking maximum power points or adjusting solar collector angles.
24	Fuzzy Logic and ML Integration	Combines fuzzy logic principles with ML algorithms to handle uncertainty and improve solar energy predictions.

2.3. Support vector machines

It is an effective, better dimensional classification and regression supervision of ML tasks that works on a hyper plane decision boundary, which separates data in continuous regression output and minimises prediction error. Besides, it is most suitable for hybrid systems such as solar collectors integrated with PCM, predicting solar intensity, humidity, etc., and thermal performances for long-term forecasting.

2.4. Neural networks

They are structured and function like a human brain type ML called artificial neural networks (ANN) to predict real-time optimisation and autonomous operation in highly effective non-linear data for a long duration. They have further categories of convolutional neural networks (CNN) and recurrent neural networks (RNN), which support recording image-based solar energy systems like solar panel inspections or weather predictions,

etc., and are suitable for time-series data, including solar intensity and weather-dependent parameters.

2.5. Clustering and classification

Clustering is an unsupervised learning technique where data, such as weather pattern analysis like humidity, solar radiation, etc., is grouped into clusters based on similarity by using algorithms like K-means, density-based Spatial clustering of applications with noise (DBSCAN) and hierarchical clustering. Classification is a supervised learning technique suitable for monitoring solar collectors for heating blockage, abnormal environmental conditions, etc., and helps classify days as sunny, cloudy, or low-thermal-energy days using a logistic regression classification tree.

2.6. Bayesian models

It is a probabilistic approach, particularly powerful for handling data on uncertainty and variability. The performance outcomes

based on prior probability and new likelihood data are calculated by using the formula:

$$P(A|B) = (P(B) \cdot P(B|A) \cdot P(A)) / P(B), \quad (1)$$

where: $P(A|B)$ = posterior probability (probability of A given B), $P(B|A)$ = likelihood (probability of B given A), $P(A)$ = prior probability (initial belief about A), $P(B)$ = evidence (total probability of B).

In this mode, Bayesian neural networks (BNNs) and Gaussian processes (GPs) are used to quantify uncertainty in solar energy performance. They are also best to support the prediction of interdependency processes, such as solar collectors integrated with thermal storage systems or photovoltaic (PV) systems.

2.7. Evolutionary algorithms model

This model is preferred for selecting natural and biological evolution as a heuristic optimisation method to solve multi-objective and complex problems. Models such as genetic algorithms (GA) and particle swarm optimisation (PSO) were used to predict the complex design and material selection of solar collectors. And their thermal performance, based on their inclined position, was accurately predicted. Most suitable for predicting weather conditions and hybrid thermal storage efficiency, producing a Pareto-optimal set of solutions for decision-makers.

2.8. Ensemble learning techniques

These techniques leverage the strengths of individual models to improve the overall performance prediction, reduce overfitting, and enhance generalisation. They use a bagging-and-boosting model that operates autonomously to correct mistakes from earlier models.

2.9. Hybrid approaches

They are a combination of ML algorithms with optimisation techniques or physical models (thermodynamics metrology) for enhanced predictions. ANN integrated with PSO for optimising thermal performance parameters is the most suitable example.

2.10. Deep learning models

These models use multiple layers of interconnected nodes (neurons) to learn and represent complex patterns in data. Particularly effective in analysing large datasets with nonlinear relationships. They use deep belief networks (DBN) that act as advanced neural networks for hierarchical feature extraction, and auto encoders for dimensionality reduction and anomaly detection with feature learning in solar systems.

2.11. Reinforcement learning

In this, machine learning makes sequential decisions by interacting with an environment to exploit a cumulative reward. Unlike supervised learning, reinforcement learning (RL) does not rely on labelled data; instead, it learns through trial and error, receiving reactions in the form of rewards or consequences based on its actions. Algorithms such as Q-learning, deep

Q-learning, and proximal policy optimisation (PPO) are notable in solar tracking applications.

2.12. Fuzzy logic and machine learning integration

It converts crisp input, such as solar intensity and ambient temperature, into fuzzy sets using linguistic terms like low, medium and high. An option like the if-then rule is used to relate the input to the output. If solar radiation is low and ambient temperature is high, then a reduced mass flow rate results in a lower control thermal efficiency. It supports enhanced prediction accuracy and adaptability, ensuring consistent thermal performance. It is necessary to describe in detail the procedures used in scientific research (experimental and computational procedures). It is also needed to explain why the above procedures are used and whether they are reliable (explain why the proposed procedure is correct for the class of problems).

In experimental work, it is necessary to specify the measurement accuracy.

3. Machine learning in drying applications

Machine learning (ML) has become a revolutionary technology across various agricultural applications, especially in optimising drying processes. This section begins with an exploration of how ML is integrated into agricultural drying systems, emphasising its role in improving efficiency, accuracy and overall performance. It then transitions to the application of ML in rough-surface solar collectors, particularly after PCM integration. Finally, the section delves into the numerical and experimental investigations of hybrid nanofluid performance in these systems.

3.1. Garlic clove

Kushwah et al. [27] presented a study on a dual-working-medium heat exchanger evacuated tube (HE-ET) solar drying system for garlic, emphasising the integration of thermodynamics and 5 different ML algorithms for performance optimisation in Python. Advanced ML algorithms, such as gradient boosting regression tree (GBRT), are highlighted for their predictive accuracy, with significant findings on the impact of ambient relative humidity and solar intensity. These results reveal that the mean absolute error (MAE) ranged between 0.31 and 0.59, root mean squared error (RMSE) between 0.54 and 0.72, and the coefficient of determination (R²) of 0.98 and 0.94 on training and testing values, providing valuable insights for enhancing the efficiency of solar drying systems. It underlines the potential of combining renewable energy technologies with data-driven approaches. The detailed integration of technology is illustrated in Fig. 1.

3.2. Sweet potato

Abdelkader et al. [28] developed an evaluation of an indirect solar dryer tailored for drying sweet potato cubes, integrating experimental, environmental and economic (3E) analyses alongside advanced ML techniques. Four algorithms, such as decision trees (DT), gradient boosting regression (GBR), multiple linear regression (MLR), and random forest (RF), were employed to predict critical performance metrics, including solar collector

(SC) energy and exergy efficiency, drying chamber mean temperature and exergy efficiency. The RF algorithm shows superior accuracy due to its adaptability. The dryer demonstrated solar collector efficiencies of 72.9% (energy) and 5.6% (exergy), storage unit efficiencies of 43.4% (thermal) and 18.4% (exergy) with a discharge time of 5.5 hours, and an average drying chamber thermal efficiency of 46.1%, ranging from 21.9% to 97.2% has been recorded.

The overall dryer efficiency was increased by 15% in consuming 4.509 kWh/kg of moisture removed. The Page model

proved optimal for predicting drying kinetics. Environmental benefits included a CO₂ mitigation of 20.2 kg and carbon credit earnings of 56,771 RMB (RMB stands for Renminbi, the official currency of China), while economic analysis revealed a payback period of 29.24 months, annual revenue of 8,464 RMB, and a return on investment of 0.7 RMB. These results underscore the potential of integrating solar drying systems with machine learning to enhance system performance, optimise energy use, and deliver sustainable, economically viable solutions for agricultural drying applications.

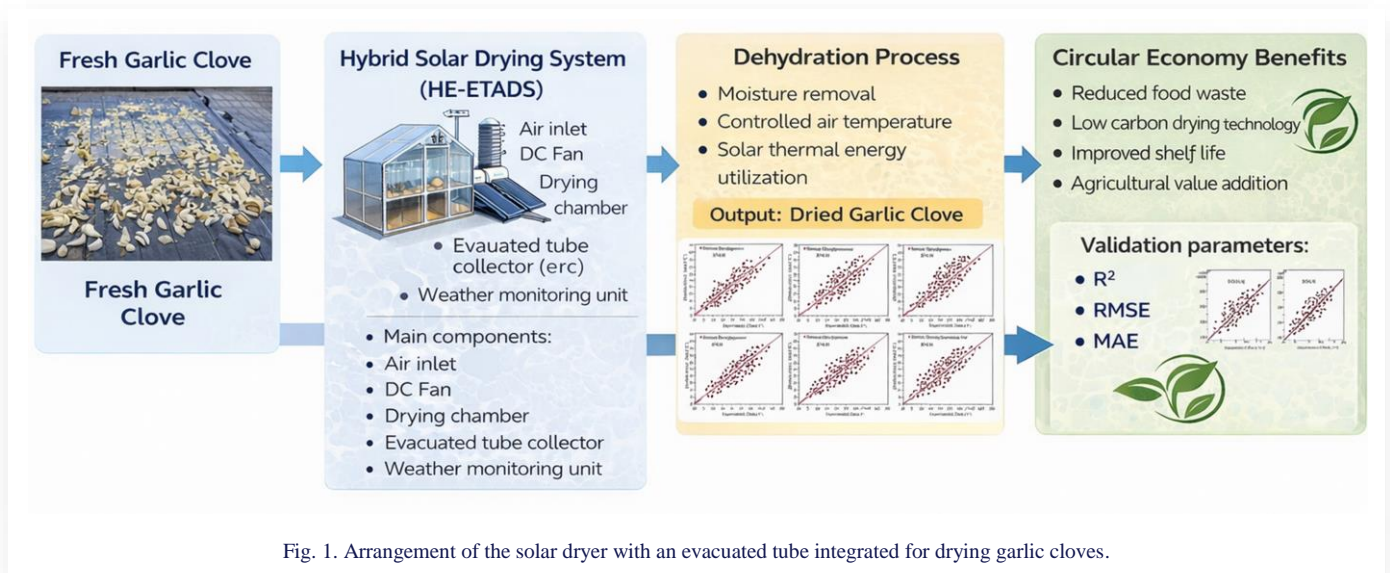


Fig. 1. Arrangement of the solar dryer with an evacuated tube integrated for drying garlic cloves.

3.3. Prediction of red chilli in a dryer

Arunkumar et al. [29] evaluated the drying performance of red chilli in an indirect solar dryer under varying energy storage element (ESE) cases like i) paraffin wax, ii) glass pieces blended with paraffin wax, and iii) crushed glass pieces mixed with paraffin wax in mass flow rates of 0.3 m/s, 0.4 m/s and 0.5 m/s. Key variables such as i) solar intensity, ii) mass flow rate, iii) ESE case, iv) ambient temperature (AT), v) collector inlet temperature (CIT), vi) collector outlet temperature (COT), vii) drying chamber inlet temperature (DCIT), and viii) drying chamber outlet temperature (DCOT) were analysed for their impact on moisture content (MC), and the collector efficiency (CE) was measured for research. Experimental results demonstrated that the dryer using case 3 ESE at a 0.5 m/s mass flow rate achieved the target MC within 25 hours.

GBRT emerged as the most accurate ML algorithm, with R^2 values of 0.999964 for MC prediction and 0.994064 for CE estimation, alongside minimal errors (MAE, RMSE, and mean squared error (MSE)). The solar collector outlet temperature was identified as the most influential factor for both MC and CE. This study highlights the importance of ESE, airflow rate and ML techniques in optimising agricultural drying processes as illustrated in Fig. 2.

3.4. Metal absorber comparison

Baz et al. [30] conducted a comparative study of solar absorber structures, specifically the single-layer metal-insulator-metal

(MIM) structure, the gold-MgF₂-gold solar absorber (GMGSA), and the multilayer MIM structure, the gold-MgF₂-gold-MgF₂-gold solar absorber (GMGMGSA), highlighting advancements in solar thermal applications. Leveraging machine learning strategies used random forest regression analysis. The research achieved exceptional predictive accuracy, with an R^2 of 0.998836 and a mean square error of $2.8091 \cdot 10^5$. It enables the optimisation of absorber designs. The GMGMGSA structure demonstrates superior performance, including a broad absorption range (0.4–1.6 μm), high absorption efficiency under AM1.5 solar irradiation (93.10%), and excellent angular independence up to 70°. These findings underscore the potential of advanced absorber designs in applications such as concentrated solar power (CSP), solar water heater (SWH) and building-integrated photovoltaics (BIPV), emphasising the transformative role of machine learning in driving innovation in solar thermal technologies.

3.5. Machine learning and computational fluid dynamics simulation

Madhankumar et al. [31] focused on the computational analysis, empirical testing, and predictive modelling of a corrugated absorber plate solar collector (CAPSC) integrated with a fin-inserted TES system. Utilising computational fluid dynamics (CFD) simulation, governing equations of the CAPSC device were developed and analysed, complemented by experimental testing across varying mass flow rates (0.014–0.064 kg/s) during

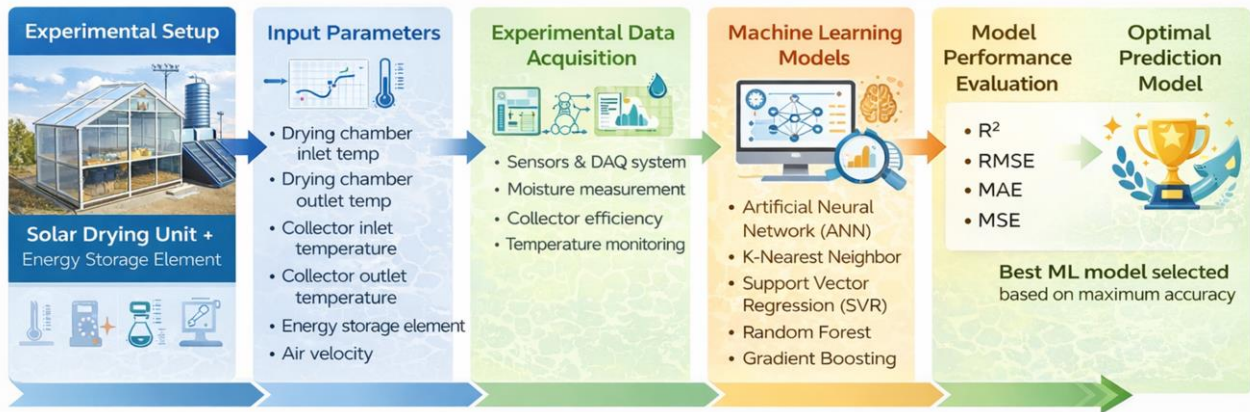


Fig. 2. Thermal performance of Red Chilli predicted by using an ML algorithm.

summer and winter seasons. The maximum deviation between computational and experimental outlet temperatures was 1.72°C at a mass flow rate of 0.044 kg/s, demonstrating close agreement. ML techniques, including MLR, RF, K-nearest neighbours (KNN), XG-boost and ANN, were employed to predict the thermal efficiency of the system, leveraging a dataset of 288 empirical samples. Among these, ANN and XG-boost emerged as the most accurate, with XG-boost excelling due to its robust handling of non-linear interactions. To further enhance prediction accuracy, computational CAPSC output temperatures were integrated as inputs into machine learning models, illus-

trating a synergistic improvement in thermal efficiency forecasts. This integration of CFD and machine learning provides a sophisticated framework for optimising solar collector designs and achieving precise thermal performance predictions, as detailed in Fig. 3.

3.6. Rough surface in a solar collector

3.6.1. Triangular-shaped rib

Nidhul et al. [32] provided an extensive review of artificial roughness designs for rectangular and triangular solar air heaters

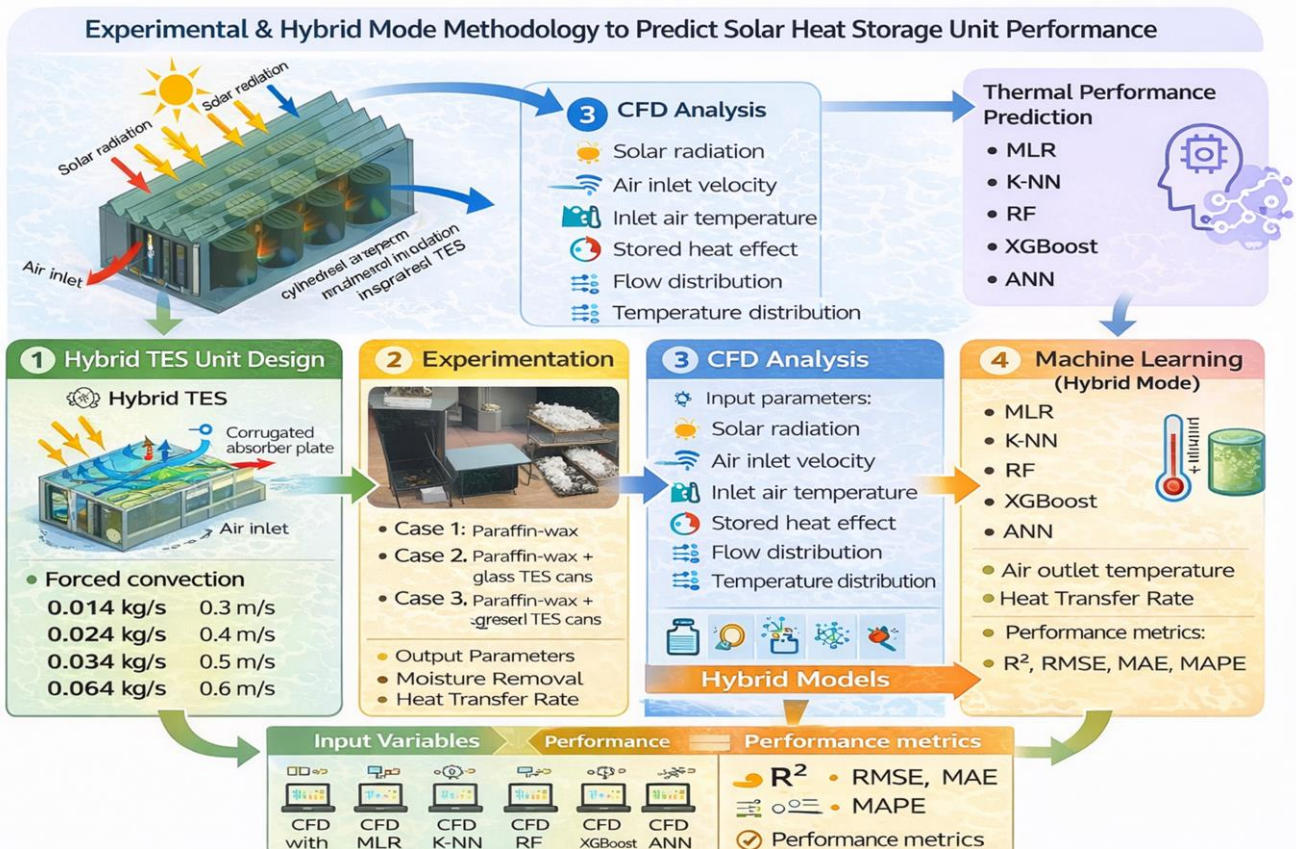


Fig. 3. Schematic layout of integration of experiment and CFD with ML.

(SAHs), helping researchers identify gaps in the field. An ANN-based ML model was developed using 454 datasets from 72 rib combinations to predict the performance of ribbed triangular SAHs. The model demonstrated high accuracy, with an average prediction deviation of less than 3%, and has the potential to improve accuracy by incorporating additional geometric and operational parameters. Second-law analysis revealed variations in collector efficiency and entropy generation with Reynolds numbers (Re) ranging from 4000 to 18000. The study identified optimal rib parameters, including height, pitch, chamfer angle and inclinations, enabling engineers to estimate the performance of ribbed triangular duct SAHs with various artificial roughness designs without the need for physical testing.

3.6.2. C-shaped fin

Saravanan et al [33] highlighted the use of ML techniques in predicting the thermal performance of SAH with advanced absorber designs, C-shaped finned panels. Experimental investigations coupled with ML models, including RF, LR, and KNN, demonstrate the capability of data-driven approaches to accurately model key performance parameters such as the Nusselt number (Nu), friction factor (f) and thermal efficiency (thermo-hydraulic performance (THP)). Among the models, RF exhibited superior predictive performance with a high R^2 value of 0.9783 and minimal errors, emphasising its potential as a robust tool for optimising SAH designs. This integration of ML with experimental studies offers a promising pathway for enhancing the design and operational efficiency of solar thermal systems.

3.6.3. Digital twin in solar flat plate collector

Zohdi et al. [34] introduced a digital twin model to optimise solar-thermal storage systems as a green energy solution during peak demand hours after sunset. The system comprises adaptable Fresnel mirrors, an optical receiver and a power distribution framework. A reduced-order Maxwell's equations model efficiently computes solar power flow, simulating mirror reflections, losses and energy delivery in microseconds. A genetic ML framework integrates sub-models for mirror tracking, thermal absorption and optimal temperature operation to balance radiative losses and heat storage. The approach optimises mirror-receiver configurations for long-term efficiency and demand fulfilment. Additionally, a deep-learning algorithm enhances simulation speed through an ANN representation, enabling rapid performance evaluations across an extended timeframe.

3.6.4. Grooved tube configuration in flat plate collector

Chilambarasan et al. [35] highlighted the enhanced performance of solar flat plate collectors (SFPC) using an internally grooved absorber tube combined with Al_2O_3 -WEG-based nanofluids. Experimental investigations with varying nanoparticle concentrations (0.01%, 0.05%, 0.1%, 0.2%) and mass flow rates (0.024 kg/s, 0.036 kg/s, 0.048 kg/s) revealed significant efficiency improvements. At 0.2% nanoparticle concentration and 0.036 kg/s mass flow rate, model 2 achieved a maximum efficiency enhancement of 54.1% over model 1 with plain fluid. The ANN models developed for real-time prediction accurately mirrored experimental results, demonstrating their effectiveness in

optimising SFPC performance. This synergistic approach addresses key challenges in enhancing the efficiency of solar thermal technology.

3.6.5. Solar panel tilt angle

Yadav et al. [36] research focuses on optimising the electricity output of photovoltaic (PV) systems in micro grids by accurately predicting the optimal tilt angle (OTA) of solar panels – a non-linear function of solar radiation parameters. Using feature selection-based ANN with numerous ML algorithms, the research recognises global solar radiation, diffuse solar radiation, clarity index and inclined surface radiation as critical predictors, while deeming extra-terrestrial radiation the least significant. By incorporating the appropriate input variables, prediction accuracy improved significantly from 38.59% to 90.72%. Among the ANNs evaluated, the Elman neural network (ENN) achieved the best performance, demonstrating its effectiveness in OTA prediction to maximise PV output.

3.6.6. V corrugation

Poongavanam et al. [37] enhanced SAH efficiency by modifying absorber plates with V-corrugation, shot-blasting (SB), nano-coating, and black pebble stone (BPS) integration. SB treats surfaces with high-speed abrasives, followed by coating with activated carbon (AC) based nano-materials in matt paint. The roughened and corrugated plate increases heat transfer to the air, and BPS acts as a thermal mass to maintain a consistent temperature. Machine learning (ML) models, including random forest regression (RFR), LR and support vector regression (SVR), are employed to predict SAH thermal efficiency using experimental data. The research reveals efficiency enhanced owing to surface modifications and BPS integration, significantly improving thermal transfer. The RFR model shows superior predictive accuracy, with a correlation coefficient of 0.999 for modified SAHs, and effectively estimates Nu values with minimal error. ML-based thermal efficiency models are validated using mean squared error and root mean squared error, confirming their predictive reliability.

Thakur et al. [38] investigated the thermal-hydraulic performance enhancement of dual passage solar thermal collectors (DP-STC) by employing single V-shaped staggered baffles (SVSSBs) on the absorber plate. Solar thermal collectors (STCs) are essential for converting solar energy into thermal energy, but the thermal performance of DP-STCs is typically suboptimal. The study focuses on the impact of various geometrical parameters, including relative staggered baffle position, relative staggered baffle size and relative roughness pitch, for Reynolds numbers ranging from 3000 to 19000. Fixed parameters, such as relative roughness height (0.40), relative gap position (0.67) and angle of attack (60°), are considered to isolate the effect of the varying geometrical factors.

The results indicate that SVSSBs outperform other types of turbulators reported in the literature, achieving a peak thermal-hydraulic performance of 3.67 under specific geometrical conditions (baffle size = 1.5, staggered baffle position = 0.65, roughness pitch = 3.5, at Re = 17000). Additionally, correlations for the Nusselt number (Nu) and the friction factor were developed

ped using experimental data, providing valuable insights into the optimisation of thermal collectors for enhanced heat transfer and reduced friction losses.

3.7. Conical cavity in a solar-evacuated tube

Rajan and Reddy [39] integrated an optical-thermal model combining Monte-Carlo ray tracing and CFD to evaluate a 40 m² parabolic dish collector (PDC) under varying inclinations and wind conditions. Ray tracing in Solar Trace, coupled with ANSYS and MATLAB, was used to model solar heat transfer to Therminol 66 in a conical receiver. Validation against established studies confirmed the model's accuracy, with a ray count of 10⁻⁶ ensuring precise mapping of the heat flux distribution. The study found that 66%–80% of absorbed solar heat was transferred to the heat transfer fluid (HTF), with this transfer influenced by inclination, wind direction, and speed. Deep learning models were employed for thermal performance prediction, with optimal activation functions selected for accuracy. A comparative analysis demonstrated the superior performance of the developed model compared to existing approaches.

3.7.1. Twisted and way tape inserted in an evacuated tube

Bingbin et al. [40] investigated the impact of inlet flow velocity, structural enhancements, and ML models on the efficiency of a solar collector (SC). Higher inlet airflow velocities (2.1 m/s, 3.3 m/s and 4.5 m/s) improve collector efficiency by intensifying heat transfer and reducing temperature build-up. Structural enhancements, such as a spiral twisted tape and wave strip combination, significantly enhance heat transfer. This configuration, with an optimal wave plate length of 360 mm and a helical twist ratio of 5.14, achieved the highest efficiency and provided a promising direction for improving solar collector tube designs. Eight machine learning models, including XG-boost, were employed to predict collector performance, with XG-boost emerging as the most accurate ($R^2 = 0.997$, RMSE = 0.008 for training;

$R^2 = 0.988$, RMSE = 0.008 for testing). Interpretable methods like partial dependence plot (PDP), accumulated local effects (ALE), and Shapley additive explanations (SHAP) analysis highlighted key performance drivers, such as inlet flow velocity, temperature difference and solar irradiance. These methods enable optimised design parameters and provide engineers with actionable insights to enhance collector performance, making interpretable machine learning a valuable tool in solar energy technology development.

Gupta et al [41] introduced a mixed-mode dryer with a semi-transparent photovoltaic thermal (PVT) collector, combining computational and experimental methods to assess drying and thermal performance. The research evaluates three inlet air velocities, such as i) 0.048 m/s, ii) 0.096 m/s, and iii) 0.144 m/s, to optimise airflow in the PVT dryer. Model 2 exhibited uniform temperature distribution, while Model 3 demonstrated the best performance, achieving enhanced thermal efficiencies of 33.20% (numerical) and 31.14% (experimental) compared to Model 1. Additionally, Model 3 improved Re and Nu by factors of 3.06 and 2.45, respectively. Experimental results closely aligned with simulations, showing deviations of only 2.24% to 4.90%. ML models, particularly ANN with different hidden layers, were utilised to predict and optimise drying performance parameters, ensuring optimal drying conditions. The detailed integration of the process is presented in Fig. 4.

3.8. Thermal performance prediction in nano-liquids

Alawi et al. [42] predictions of thermal performance (η) for flat plate solar collectors (FPSC) are critical for various engineering applications in thermal and energy sectors. This study explores the potential of ML models, including hist gradient boosting regressor (HGBR), multivariate adaptive regression splines (MARS), tweedie regressor (TR) and stacking regressor (SR), to predict η using Al₂O₃-H₂O nanofluids. A predictive matrix was

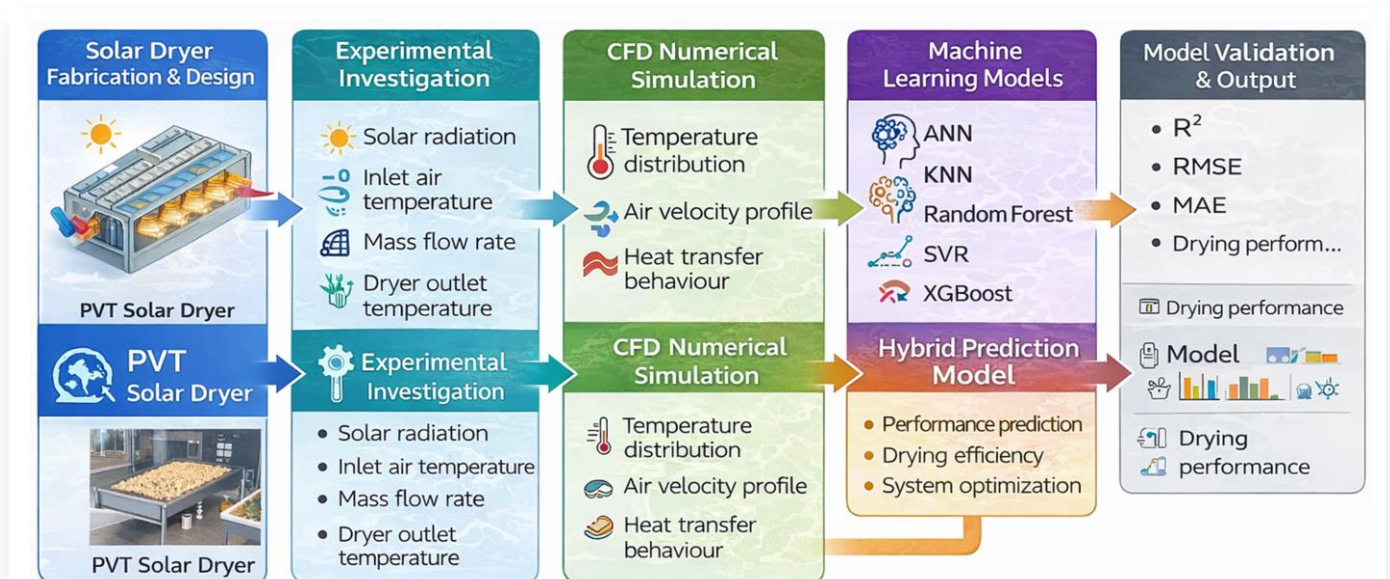


Fig. 4. Schematic layout of ML in semi-transparent photovoltaic thermal (PVT) collector.

developed incorporating five key predictors: nanoparticle size, collector slope, absorbed energy parameter, removed energy parameter and reduced temperature parameter. Forward feature selection (FFS) integrated with the RF algorithm was employed to construct different input combinations.

Among the models tested, Model 5 demonstrated superior performance across all metrics, achieving the lowest mean absolute percentage error (MAPE), root mean squared error (RMSE), U95 (95th percentile uncertainty), and residual error, alongside the highest coefficient of determination (R^2), Pearson correlation coefficient (PCC), Nash-Sutcliffe coefficient (NSE), and Willmott's index (WI). Model 4 showed strong results in R^2 and PCC, while Models 2 and 3 exhibited moderate performance. Model 1 lagged in most evaluation metrics. The stacking algorithm in Model 5 proved to be the most effective, with a prediction accuracy of 94.5%, surpassing HGBR (92.9%), MARS (92.7%) and TR (89.9%).

This research provides valuable insights into η prediction using thermal predictors, highlighting the efficiency of stacking algorithms and their applicability to improve the performance and design of FPSCs.

3.8.1. Specific heat prediction in nanofluid

Alade et al. [43,44] conducted research on nanofluids due to their superior thermal properties, which are critical for enhancing the solar thermal collector efficiency. Accurate modelling of their specific heat capacity is essential but underexplored, and existing analytical models exhibit significant predictive inconsistencies. A Bayesian support vector regression (BSVR) model was developed to address this, trained on 84 experimental datasets and validated on 17 test datasets. The model achieved remarkable accuracy, with a 99.95% Pearson correlation coefficient and an absolute average relative deviation (AARD) of 0.1888, outperforming existing models by a significant margin. This advancement ensures more reliable efficiency computations for solar collectors, optimising their design and performance.

Further, they extended to introduce a genetic algorithms/support vector regression (GA/SVR) model to accurately evaluate the specific heat capacity of Alumina (Al_2O_3)/water nanofluid at volume fractions of 3.7–9.3%. The GA/SVR model demonstrated exceptional accuracy, achieving a 99.998% correlation with experimental data and an RMSE of 0.0014. Comparative analysis revealed the GA/SVR model outperformed existing analytic models, providing an order-of-magnitude improvement over the thermal equilibrium model and a two-order-of-magnitude improvement over the simple mixing rule model. This advanced model offers rapid and precise predictions, optimising nanofluid performance in solar thermal applications.

3.8.2. Prediction of viscosity of nanofluids

Meybodi et al. [45] reported research on nanoparticles and a base fluid, which is extensively used in industrial processes such as engine cooling and enhanced oil recovery due to their unique properties, including viscosity, which is crucial for chemical and petroleum engineering applications. This study

employs least squares support vector machines (LSSVM) to accurately predict the viscosity of water-based nanofluids containing Al_2O_3 , TiO_2 , SiO_2 , and CuO nanoparticles. The LSSVM model demonstrates exceptional accuracy, achieving a coefficient of determination of 0.998 and outperforming existing models. The model accurately predicts viscosity trends across varying nanoparticle volume fractions and temperatures, aligning well with experimental data. Additionally, a leverage approach was used to identify outliers and validate the model's applicability domain. This robust, rapid and precise modelling approach significantly enhances viscosity prediction for nanofluids.

3.8.3. Relative viscosity of nanofluids

Baghban et al. [46] focused on nanofluid viscosity as a critical property in convective heat transfer, yet existing theoretical models have limited applicability. This study collected 1277 experimental data points of nanofluid relative viscosity (NF-RV) and developed a general predictive model using an adaptive network-based fuzzy inference system (ANFIS). The model incorporates variables such as temperature, nanoparticle diameter, nanofluid density, volumetric fraction and base fluid viscosity. Statistical analysis showed exceptional accuracy, with an R^2 value of 0.99997 and relative deviations of 0.42% (absolute) and 6.45% (maximum average). The ANFIS model outperformed existing theoretical models, demonstrating its ability to reliably predict NF-RV. This tool is invaluable for chemists and engineers, particularly in heat transfer applications using nanofluids.

Wu et al. [47] study presents a geometric analysis and optimisation of parabolic trough solar collectors (PTC) using a Python-based algorithm to calculate optical efficiency with a secondary hyperbolic mirror. Key features include the concentration ratio, intercept factor and optical efficiency, critical for PTC geometric optimisation. The algorithm integrates AutoCAD data for simulating varying incident angles and focal lengths, achieving a simulation-theory error margin of less than 3.13%. Using ML, optical efficiency was numerically fitted, yielding a high accuracy ($R^2 = 0.97874$). The study develops optical efficiency fitting formulas derived from secondary reflection hyperbolic data and validated using a ray-tracing model. This method enhances the design, energy analysis and optimisation of PTCs under diverse geometric and operational conditions.

3.9. ML in parabolic trough solar collector using hybrid nanofluids

Alhamayani [48] researched on the advancement of solar energy technologies, particularly for sustainable power generation, and emphasised the importance of optimising parabolic trough solar collectors (PTC) systems. This study explores the impact of hybrid nanofluids (1% Al_2O_3 –2% multi-walled carbon nanotubes (MWCNT)/Syltherm-800, 1.5% Al_2O_3 –1.5% MWCNT/Syltherm-800, and 2% Al_2O_3 –1% MWCNT/Syltherm-800) on PTC performance. A mathematical model was developed to assess PTC efficiency across varying nanofluid concentrations. Additionally, ML models, DT, support vector machine (SVM) and ANN were trained to predict PTC outlet temperature efficiently. Results revealed that 2% Al_2O_3 –1% MWCNT/Syltherm-800 achieved the highest thermal efficiency, averaging 70.54%. The

ANN model exhibited superior predictive accuracy, with an R^2 of 99.99%, MAPE of $4.8 \cdot 10^{-5}$, RMSE of 0.012 and MAE of 0.0057. This research highlights the potential of hybrid nanofluids and artificial intelligence (AI) to enhance the performance of PTC systems.

Said et al. [49] investigated the thermal performance of an FPSC using MWCNT + Fe₃O₄/Water hybrid nanofluids in an arid environment. Experiments revealed a 26.3% increase in the heat transfer coefficient, accompanied by an 18.9% increase in the friction factor (f). Experimental data were used to develop prediction models for heat transfer, Nu, f and thermal efficiency using boosted regression tree (BRT) and XG-boost ML techniques. The XG-boost models demonstrated superior performance, with R^2 values ranging from 0.9914 to 0.9997, lower mean squared errors (0.000081–9.11) and minimal mean absolute percentage errors (0.0025–0.3114). Statistical evaluations, complemented by Taylor's graphs, emphasised the effectiveness and accuracy of XG-boost over BRT for thermal performance predictions.

Al-Rashed et al. [50] numerically evaluated the performance of a novel PTC featuring a non-circular absorber tube and solid insulation, using a water-CMC/Al₂O₃ (carboxymethyl cellulose containing Al₂O₃) nanofluid as the working fluid. Performance metrics were analysed with respect to the Reynolds number (Re), nanoparticle volume concentration, particle size and insulation angle. Results reveal an ascending-descending pattern of efficiency variation with Re, achieving a peak efficiency of 61.7% for a dual-fluid collector with a 90° insulation angle, 1.5% nano-particle concentration and 100 nm particle diameter. Additionally, an ANN model with 8 neurons accurately predicted collector efficiency, with a maximum deviation of less than 0.2% and an R^2 of 0.9998, demonstrating its reliability for forecasting PTC performance.

Bouzgou and Gueymard [51] addressed the critical need for accurate solar irradiance forecasting to support the efficient utilisation of solar energy in a carbon-free economy. A novel forecasting approach combining minimum redundancy–maximum relevance (MRMR) and extreme learning machine (ELM) is proposed for irradiance time-series prediction. The method is tested across three forecasting horizons (15 minutes, 1 hour, and 24 hours) and evaluated against four scenarios: i) long window (latest 50 variables), ii) short window (latest 5 variables), iii) principal component analysis (PCA), and iv) a clear-sky model. Using irradiance data from 20 sites across diverse climates, the results demonstrate that selecting an optimal set of relevant variables through MRMR significantly enhances forecasting accuracy. The proposed approach outperformed the alternative scenarios, proving its efficacy in improving the prediction of global solar radiation and offering a robust solution for integrating solar energy into future energy systems.

Fuxi et al. [52] investigated the effect of spring turbulators on turbulent flow inside a pipe using TiO₂-Cu/water hybrid nanofluid and a single-phase model, focusing on heat transfer rate (HTR), pressure drop (ΔP) and solar collector efficiency (η_{SC}). The numerical analysis, based on the Reynolds number, the turbulator scale and the nanoparticle volume fraction (ϕ), reveals that increasing the turbulator scale improves fluid contact,

increases turbulence and enhances nanofluid mixing, resulting in better heat transfer performance. The use of ANN for modelling predicts the system performance with high accuracy, offering predictions similar to those from fitting methods. The findings highlight the potential of spring turbulators in improving solar collector efficiency, with ANN providing a reliable tool for performance prediction.

3.10. Heat storage (with and without phase change materials)

Zayed et al. [53] evaluated the performance of wavy corrugated solar air collectors (WCSAC) with paraffin wax as thermal energy storage (TES). Experiments were conducted under airflow rates of 0.540 kg/min, 1.68 kg/min, and 3.72 kg/min with paraffin layer thicknesses of 2 cm and 4 cm. A 4 cm paraffin layer improved energy efficiency by 24.0%, 20.39% and 16.37% at the respective airflow rates compared to systems without PCM. The convective heat transfer coefficient was over 1.20 times higher in PCM-equipped systems.

Support vector machine (SVM) models, optimised with Karush-Kuhn-Tucker conditions, accurately predicted thermal performance metrics, achieving coefficients of determination of 0.990 (training) and 0.950 (testing). Sensitivity analysis identified the air inlet temperature, solar irradiance and PCM thickness as key parameters influencing performance. This study demonstrates the effectiveness of PCM and AI models in enhancing solar thermal systems.

Badji et al. [54] highlighted the effectiveness of integrating an SAH with an underground latent heat storage system (LTES) using CaCl₂+6H₂O as a PCM to enhance greenhouse heating efficiency. Conducted in a semi-arid climate in Ghardaia, Algeria, the innovative system increased night-time air temperatures by 57% compared to conventional greenhouses. The LTES released 20 kJ of energy during the night, meeting 30% of the greenhouse's energy requirements. ML models, particularly the non-linear autoregressive exogenous (NARX) model, achieved exceptional predictive accuracy, with an R^2 of 0.9986, surpassing the RNN model's R^2 of 0.9893. These findings demonstrate the potential of LTES systems and advanced ANN-based models in optimising greenhouse conditions, significantly contributing to sustainable agriculture and energy-efficient practices.

3.11. Parallel tube collector using hybrid nanofluid

Mausam et al. [55] investigated the thermal performance of a solar flat plate collector (FPC) using a hybrid nanofluid composed of Cu-MWCNTs-water. Experimental investigations varied flow rates, inclination angles and radiation intensities, with an ANN and a mathematical model developed to predict instantaneous efficiency. Inputs to the ANN included nanofluid concentration, flow rate, inclination angle and radiation intensity, while instantaneous efficiency was the output. The model identified optimal neurons based on the minimal MSE and a high coefficient of determination (R^2). Results revealed a 32.25% enhancement in instantaneous efficiency using the Cu-MWCNTs hybrid nanofluid, with R^2 values ranging from 0.8857 to 0.9989,

demonstrating strong predictive accuracy. This research highlights the effectiveness of hybrid nanofluids and ANN-based modelling for optimising FPC performance in heat transfer and thermal applications like fluid and space heating.

Mustafa et al. [56] investigated the thermal performance of a double-fluid parabolic trough solar collector (PTC) with a helical absorber tube, following the American Society of Heating, Refrigeration and Air Conditioning Engineers (ASHRAE) Standard 93 [57]. The study utilised non-Newtonian nanofluids as the working fluid, analysed using the finite volume method, the semi-implicit method for pressure linked equations (SIMPLE) algorithm and the $k-\varepsilon$ turbulence model, with the importance of the turbulence model emphasised [58]. The results identified the optimal configuration as a double-fluid solar receiver featuring a helical absorber tube, a nanofluid volume fraction of 4% and a nanoparticle diameter of 50 nm. At a Re number of 5000, the single-flow PTC achieved a maximum efficiency (η_{\max}) of 47.5% with nanofluids and 43.1% with water. In comparison, the double-flow PTC reached efficiencies of 56.5% at a 1% nanofluid concentration and 58.2% at 4%. Additionally, a deep learning-based ANN model was employed to predict efficiency, effectively correlating nanoparticle concentration, diameter and Re number with the system performance. This demonstrated the reliability of ANN models for predicting PTC efficiency.

3.12. ANN prediction in SAH

Ghritlahre et al. [59] focused on the application of ANN for predicting the thermal performance of SAH, comparing roughened surfaces like arc-shaped wire ribs and smooth duct designs. This research emphasises identifying and utilising relevant input parameters for performance prediction. Seven neural models were developed, varying the number of neurons, with ANN-II (8-14-1 architecture) identified as the optimal model based on metrics such as sum square error (SSE) (0.02138), mean relative error (MRE) (1.82%) and R^2 (0.99387). The ANN model outperformed the group method of data handling (GMDH) approach, demonstrating superior accuracy. Sensitivity analysis further revealed that the air mass flow rate is the most critical input parameter affecting SAH performance. This study underscores the efficacy of ANN in enhancing SAH performance predictions, setting a benchmark for data-driven modelling approaches in solar thermal research. Kumar et al. [60] demonstrated the effectiveness of ANN models trained with the Levenberg–Marquardt algorithm in accurately predicting the thermal efficiency of wire-screen porous bed solar air heaters under unidirectional and cross-flow configurations. High coefficients of determination ($R^2 = 0.999$ and 0.996) and low error metrics confirm that the ANN is a reliable and robust tool for predicting the thermal performance of porous bed SAHs.

Further research [61,62], particularly multi-layer perceptron (MLP) models, has shown them to be efficient and reliable tools for modelling and performance prediction of solar air heaters due to their capability to handle nonlinear and complex system behaviour. By critically analysing studies from 2009 to 2021, the paper identifies MLP as the most widely applied ANN architecture for SAH performance prediction and outlines re-

search gaps to guide future optimisation and modelling efforts [63]. Emphasises artificial neural networks as efficient soft-computing tools for predicting and optimising solar air heater performance, offering faster and cost-effective alternatives to conventional experimental and analytical methods.

3.13. Field survey

Tschopp et al. [64] provided a comprehensive dataset from a large-scale solar thermal collector array at Fernheizwerk Graz, Austria, one of Central Europe's largest solar district heating installations, featuring a gross collector area of 516 m² (361 kW of nominal thermal power). High-precision operational data, collected throughout 2017 with a 1-minute sampling rate and rigorous quality assurance (8.2% missing data), include key parameters such as flow rates, inlet and outlet temperatures, various irradiance metrics and weather conditions. The dataset also provides calculated values, such as thermal power output and mass flow, along with detailed uncertainty information for most variables. Released under a Creative Commons Attribution-Share Alike (CC BY-SA) 4.0 license with metadata in machine-readable JavaScript Object Notation (JSON) format, this dataset is ideal for applications such as dynamic model validation, radiation algorithm development, performance optimisation and uncertainty analysis. Notably, it is the first publicly available dataset of its kind for large-scale solar thermal collector arrays.

3.14. Weather prediction

Falope et al. [65] introduced a 3-step forecasting architecture using low-level data fusion to syndicate weather parameters from physical on-site and nearby weather stations, significantly improving solar energy prediction accuracy. By applying the Pearson correlation coefficient, weather variables with strong correlations to solar radiation were identified, enhancing forecasting models. Pre-processing techniques like rescaling and list-wise deletion further improved accuracy, with utility-scale (1 MW) and residential rooftop PV systems (8 kW, 10.5 kW, and 15 kW) achieving low RMSE values, demonstrating versatility across different PV sizes.

The model improved prediction accuracy by 6% and 13% over single-source data and by 19% when combining multiple data sources without correlation-based feature selection. Solar radiation emerged as the most critical predictor, achieving an RMSE of 0.0859 with the boosted tree algorithm. While the architecture's reliance on solar radiation data and its United Kingdom (UK) centric scope are limitations, future research could extend the model globally, incorporate negatively correlated parameters and explore advanced validation and pre-processing methods. This approach could also support forecasting for other renewable energy (RE) sources, such as wind and tidal, broadening its applicability in renewable energy forecasting.

3.15. Failure prediction of PV plant

Pereira and Silva [66] emphasised the critical role of ML algorithms in the monitoring, classification and prediction of inverter failures in PV solar energy plants. By leveraging historical data, ML enables the anticipation and optimisation of inverter performance, thereby improving energy production and

efficiency. Recommendations include connecting standardised modules in series or parallel to enhance voltage, current and capacity in power-conversion systems, and incorporating clamp circuits to protect inverters from inrush and overcurrent. Seasonal variations in inverter failures inform tailored control strategies to enhance reliability, while continued and improved data collection is essential for robust evaluation of failure rates and recurrent patterns, aiding post-warranty maintenance. Techniques such as sinusoidal pulse-width modulation reduce harmonic distortion and control output voltage, contributing to enhanced power-conversion efficiency. Results demonstrate the effectiveness of ML in identifying failure modes and patterns, issuing alarms for parameter deviations and providing insights into root causes. The findings underline the importance of data-driven monitoring and predictive maintenance for improving reliability and operational performance in solar PV systems.

4. Performance of experimental analysis of solar air heater

In this study, the extreme gradient boosting (XG-boost) machine learning model is employed to predict the thermo-hydraulic performance score (THPS) of an SAH with polygonal ribs. Unlike conventional statistical approaches, XG-boost efficiently handles non-linear relationships between the geometric parameters and Re number by optimising decision trees through gradient boosting. The dataset comprises experimentally obtained THPS values for different rib geometries (polygonal and trapezoidal), pitch-to-height p/e ratios ($p/e = 5$, $p/e = 7.5$, and $p/e = 10$) and Reynolds numbers ($Re = 2000$ to $Re = 20000$). The input features include a) rib geometry type, b) p/e ratio, and c) Re, while the target variable is the corresponding THPS value. The model is trained using an 80:20 train-test split, and hyperparameters such as learning rate, tree depth and regularisation terms are optimised to enhance the prediction accuracy. The detailed schematic diagram is shown in Fig. 5, and the position of the thermocouples placed to record the temperature in the absorber plate as illustrated in Fig. 6.



Fig. 5. Experimental setup of SAH: a) side view, b) front view.

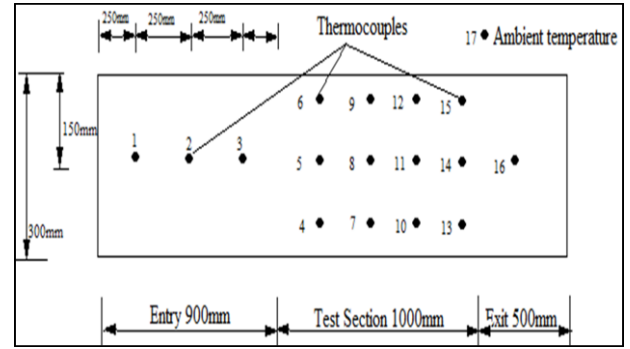


Fig. 6. Location of thermocouples for temperature measurement in flat plate SAH.

4.1. Experimental data

The experimental investigation was conducted on a rectangular duct under varying mass flow rates, different rib positions and a range of Reynolds numbers to analyse the thermo-hydraulic performance. The results, as presented in Table 2 [67], highlight the influence of rib positioning on the heat transfer characteristics and frictional losses.

Table 2. Experimental reading of various rib surfaces in the absorber plate.

Reynolds number	Rib shape and (p/e) ratio					
	Polygonal ($p/e = 5$)	Polygonal ($p/e = 7.5$)	Polygonal ($p/e = 10$)	Trapezoidal ($p/e = 5$)	Trapezoidal ($p/e = 7.5$)	Trapezoidal ($p/e = 10$)
2000	1.54	1.86	1.77	1.47	1.71	1.68
4000	2.43	2.95	2.86	2.34	2.93	2.86
6000	2.23	2.54	2.50	2.14	2.48	2.47
8000	2.26	2.48	2.49	2.17	2.46	2.43
10000	2.04	2.20	2.19	1.90	2.12	2.11
12000	1.97	2.08	2.09	1.92	2.09	2.08
14000	2.10	2.21	2.20	2.05	2.19	2.19
17000	1.88	2.01	2.01	1.82	1.96	1.99
20000	1.95	2.04	2.05	1.82	2.02	2.07

4.2. Prediction of XG-boost

The application of the extreme gradient boosting (XG-boost) machine learning model for predicting the thermo-hydraulic performance score (THPS) of an SAH integrated with polygonal and trapezoidal ribs has demonstrated high predictive accuracy [25]. The model's performance was evaluated using key statistical metrics, including R^2 , mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE), for both the training and test datasets.

To enhance the predictive accuracy of the THPS for an SAH integrated with polygonal ribs, the XG-boost machine learning model was employed. XG-boost is a powerful ensemble learning algorithm that constructs multiple decision trees sequentially, where each subsequent tree corrects the errors of the previous one, thereby improving overall prediction accuracy.

The dataset for training the XG-boost model was derived from experimental THPS values for different Reynolds numbers ($Re = 2000$ to $Re = 20000$) and (p/e) ratios (polygonal and trapezoidal ribs) with ($p/e = 5$, $p/e = 7.5$, and $p/e = 10$). The input parameters included the Reynolds number, rib geometry and p/e ratio, while the target variable was THPS.

4.3. Mathematical modelling

The thermo-hydraulic performance (THP) of the system is evaluated using the following relationship [22]:

$$THP = (Nu/Nu_0) / (f/f_0) 0.333, \quad (2)$$

where: Nu and Nu_0 are the Nusselt numbers for the roughened and smooth configurations, respectively; f and f_0 are the friction factors for the roughened and smooth configurations, respectively.

The exponent 0.333 accounts for the trade-off between heat transfer enhancement and pressure drop.

To predict the THP values under various operating conditions, the XG-boost algorithm is employed. XG-boost, a gradient-boosting decision tree-based model, is selected due to its efficiency in handling nonlinear relationships and complex datasets. The model is trained on the experimental dataset, with the Reynolds number, p/e ratio and geometric configuration as input features, and THP values as outputs. The relationship learned by XG-boost is expressed as:

$$THP_{pred} = \sum_{k=1}^n \alpha_k \cdot T_k(Re, p/e) + \beta, \quad (3)$$

where: α_k are the model coefficients, T_k represents the tree-based transformations learned by XG-boost, Re is the Reynolds number, p/e is the relative roughness parameter, β is the bias term.

Table 4. Comparison of experimental and XG-boost-predicted THP values.

No.	Configuration (p/e)	Re = 2000	Re = 4000	Re = 6000	Re = 8000	Re = 10000	Re = 12000	Re = 14000	Re = 17000	Re = 20000	Mode
1	Polygonal ($p/e = 5$)	1.54	2.43	2.23	2.26	2.04	1.97	2.10	1.88	1.95	Experimental
		1.52	2.40	2.21	2.25	2.03	1.95	2.09	1.86	1.94	ML prediction
2	Polygonal ($p/e = 7.5$)	1.8	2.95	2.54	2.48	2.20	2.08	2.21	2.01	2.04	Experimental
		1.83	2.90	2.50	2.45	2.18	2.06	2.19	1.99	2.02	ML prediction
3	Polygonal ($p/e = 10$)	1.77	2.86	2.50	2.49	2.19	2.09	2.20	2.01	2.05	Experimental
		1.75	2.83	2.47	2.46	2.17	2.07	2.18	2.00	2.03	ML prediction
4	Trapezoidal ($p/e = 5$)	1.47	2.34	2.14	2.17	1.90	1.92	2.05	1.82	1.82	Experimental
		1.45	2.30	2.12	2.14	1.88	1.91	2.03	1.80	1.81	ML prediction
5	Trapezoidal ($p/e = 7.5$)	1.71	2.93	2.48	2.46	2.12	2.09	2.19	1.96	2.02	Experimental
		1.69	2.89	2.45	2.43	2.10	2.07	2.17	1.94	2.00	ML prediction
6	Trapezoidal ($p/e = 10$)	1.68	2.86	2.47	2.43	2.11	2.08	2.19	1.99	2.07	Experimental
		1.66	2.83	2.44	2.40	2.09	2.06	2.17	1.97	2.05	ML prediction

The comparison between the experimentally obtained values and machine learning (ML) predicted values using the XG-boost model provides significant insights into the reliability and efficiency of ML-based approaches in thermal performance prediction. The experimental results for the thermo-hydraulic performance (THP) under different rib configurations and Reynolds numbers were systematically evaluated and compared with the

4.4. Uncertainty analysis

The uncertainties and relative errors of key parameters, including temperature difference across the heater in the test section, air mass flow rate, solar radiation intensity, thermal efficiency, useful heat gained and the convective heat transfer coefficient of SAH were evaluated. The relative error results are presented in Table 3.

Table 3. Uncertainty analysis of experimental reading.

No.	Parameter	Attained values in SAH	Uncertainties	Relative error
1	Temperature differences [K]	16	± 0.17	0.67
2	Mass-flow rate [$kg\ s^{-1}$]	0.062	± 0.00103	1.67
3	Solar radiation [$W\ m^{-2}$]	900	± 7	0.71
4	Thermal efficiency [%]	77.9	± 0.02	0.027
5	Connective heat transfer [$W\ m^{-2}\ K^{-1}$]	47.9	± 0.71	1.71
6	Useful heat gained by SAH [W]	872.8	± 8.1	1.57

5. Results and discussion

The comparison between experimental and predicted values is conducted to validate the accuracy of the XG-boost model in forecasting the thermal performance of solar air heaters with different rib configurations. The detailed predicted values are presented in Table 4.

predicted values from the XG-boost model. The detailed comparison reveals that the ML predictions exhibit high accuracy, with deviations ranging between 2.24% and 4.90% from the experimental values, which falls within an acceptable error range for engineering applications.

The predictive capability of the XG-boost model highlights its effectiveness in capturing complex non-linear relationships

between the Re number, rib configurations and thermal performance metrics such as the Nu number, friction factor (f) and performance evaluation coefficient (PEC). The highest accuracy was achieved for mid-range Re numbers, where flow turbulence and heat transfer interactions were well represented in the training data.

From a technical standpoint, integrating ML models like XG-boost into thermal performance analysis presents several advantages for researchers and engineers. Traditional experimental methods require extensive resources, time, and high operational costs due to repeated physical setups, sensor calibrations and data acquisition processes. In contrast, once trained on experimental datasets, the XG-boost model can rapidly predict THP values under varied conditions without the need for additional physical testing, thereby reducing the overall time and cost by approximately 60–70% in iterative testing processes. This advantage is particularly beneficial for designing solar air heaters, heat exchangers and other thermal systems, where parametric studies involving multiple geometries and operating conditions can be computationally optimised before physical prototyping.

Figure 7 presents a heat map of the XG-boost model performance metrics for both training and testing datasets. The colour intensity highlights the magnitude of each metric, with lighter shades representing lower errors (MAE, RMSE and MAPE), and darker shades indicating higher values. The heat-map clearly shows that the model achieves high predictive accuracy ($R^2 \approx 0.98$) while maintaining low error values across both datasets, demonstrating its robustness and generalisation capability. This visual representation allows quick comparison of model performance metrics and effectively supports the reliability of the proposed ML approach.

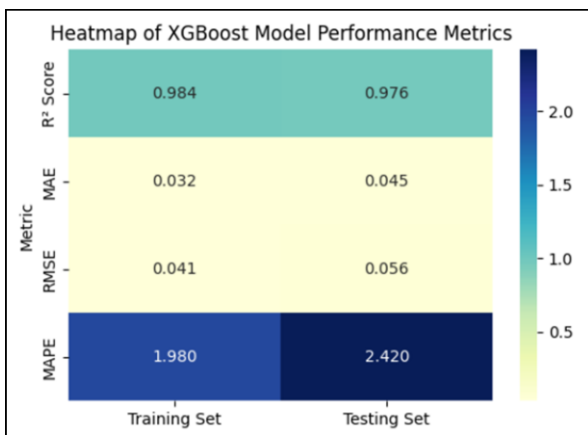


Fig. 7. Heat map of XG-boost model of THP prediction.

Additionally, ML-based predictions enable real-time analysis and optimisation, making them highly suitable for adaptive control systems in solar collectors. The model's ability to predict THP with high accuracy enables researchers to make data-driven decisions on material selection, rib design modifications and operational strategies, thereby enhancing the overall efficiency of thermal energy systems. The low MAE (0.032 for training and 0.045 for testing) and RMSE (0.041 for training and

0.056 for testing) demonstrate the model's robustness in minimising absolute and squared deviations between predicted and actual values. Moreover, the MAPE values (1.98% for training and 2.42% for testing) indicate that the XG-boost model maintains a low percentage error, ensuring reliable predictions even for unseen data points. The model's strong performance can be attributed to its ability to handle nonlinear dependencies, optimise feature importance and reduce over-fitting through gradient boosting and tree pruning techniques.

The implications of these findings are significant for the design and optimisation of energy-efficient solar air heaters. The ability to accurately predict THPS without extensive experimental iterations enables rapid prototyping of different rib configurations. This approach not only reduces computational costs but also accelerates the deployment of optimised solar air heater systems for sustainable energy applications. The performances are presented in Table 5.

These results strongly support the applicability of XG-boost for accurate THPS predictions and highlight its potential in enhancing thermo-hydraulic performance evaluations without the need for extensive experimental setups. Future studies can focus on integrating real-time sensor data with the XG-boost model to enable dynamic monitoring and adaptive control of solar air heater operations in varying climatic conditions.

Table 5. Performance metrics of XG-boost model for THPS prediction.

No.	Metric	Training Set	Testing Set
1	R^2 score	0.984	0.976
2	Mean absolute error (MAE)	0.032	0.045
3	Root mean squared error (RMSE)	0.041	0.056
4	Mean absolute percentage error (MAPE)	1.98%	2.42%

6. Challenges and future direction

6.1. Challenges

Despite notable progress in applying ML techniques to solar thermal energy systems, several critical challenges limit their large-scale implementation. One major issue is the limited availability of high-quality, real-time experimental data required for robust model training and validation. Additionally, developing generalised ML models that maintain prediction accuracy across diverse climatic zones, collector geometries and operating conditions remains challenging. The integration of ML algorithms with existing solar thermal infrastructure and control systems introduces practical complexities, including system compatibility and real-time deployment. Furthermore, high computational requirements, data pre-processing efforts and the need for models capable of handling highly nonlinear, transient and uncertain environmental conditions pose significant barriers to widespread adoption.

6.2. Future directions

Future research should prioritise the development of hybrid physics-informed machine learning models that couple energy

balance equations with data-driven algorithms to improve prediction accuracy, robustness, and physical interpretability. Emphasis should be placed on real-time implementation of ML models using edge computing and embedded systems for adaptive control, fault diagnosis and performance optimisation of solar thermal collectors. Advanced learning strategies such as deep neural networks, reinforcement learning for operational control, and transfer learning for climate-independent model generalisation are recommended to address data scarcity and site-specific limitations. Furthermore, tight integration of ML with Internet of Things (IoT) enabled sensor networks can enable high-resolution data acquisition for predictive maintenance and intelligent energy management. Finally, future studies should incorporate thermo-economic and sustainability metrics to support scalable, cost-effective and environmentally resilient solar thermal system design.

6.3. Research opportunities

Numerous research opportunities exist to advance ML applications in solar thermal energy. By optimising the integration of PCMs and hybrid nanofluids, ML can significantly enhance thermal storage and heat transfer efficiency, even under challenging weather conditions, while adhering to established guidelines [60]. ML can also be utilised for predictive maintenance, enabling proactive interventions that reduce downtime and operational costs. Developing localised solutions tailored to specific climatic and operational conditions can improve adoption rates, particularly in diverse regions. Furthermore, ML has the potential to optimise energy efficiency by minimising thermal and frictional losses and scaling applications for large solar thermal plants and decentralised systems, supporting the global transition to sustainable energy solutions.

7. Conclusions

The application of machine learning (ML) to solar collector systems marks a significant step forward in advancing solar thermal technologies by enabling highly accurate performance prediction, operational flexibility and improved system responsiveness. ML-based models provide reliable estimation of thermal and thermo-hydraulic behaviour, allowing real-time optimisation of heat transfer characteristics and reduction of operational expenses. In particular, tree-based learning approaches such as XG-boost have proven effective in representing complex non-linear interactions in SAH, incorporating artificial roughness and PCM, with prediction errors consistently within acceptable limits when compared with experimental observations.

Beyond predictive modelling, ML enhances system reliability through continuous monitoring, early fault detection and intelligent control, thereby extending the operational lifespan of solar collectors. Data-driven insights further support the optimisation of hybrid nanofluids and PCM configurations, improving thermal storage and heat transfer performance across a wide range of climatic conditions. The capacity of ML algorithms to analyse large-scale datasets also enables the development of region-specific performance models, increasing the practical applicability of solar thermal systems in diverse environments.

Despite these advantages, challenges related to data quality, computational requirements, and system integration remain and require focused attention. Future advancements should emphasise physics-informed ML frameworks, hybrid learning strategies, and tighter coupling with CFD tools to enhance model generalisation and physical consistency. Overall, this study reinforces the role of ML as a powerful enabler for intelligent, efficient, and sustainable solar thermal energy systems, contributing meaningfully to next-generation renewable energy solutions.

Acknowledgements

The research was supported by the European Commission HORIZON 2020 project RESHeat in the field of research and innovation under grant agreement No. 956255.

References

- [1] Gill, R.S., Hans, V.S., Saini, J.S., & Sukhmeet, S. (2016). Investigation on performance enhancement due to staggered piece in a broken arc rib roughened solar air heater duct. *Renewable Energy*, 104, 148–162. doi: 10.1016/j.renene.2016.12.002
- [2] Vipin, B., Gawande, A.S., Dhoble, D.B., & Zodpe, S.C. (2016). Experimental and CFD investigation of convection heat transfer in solar air heater with reverse L-shaped ribs. *Solar Energy*, 131, 275–295. doi: 10.1016/j.solener.2016.02.040
- [3] Kumar, K., Prajapati, D.R., & Samir, S. (2017). Heat Transfer and Friction Factor Correlations Development for Solar Air Heater Duct Artificially Roughened with ‘S’ Shape Ribs. *Experimental Thermal and Fluid Science*, 82, 249–261. doi: 10.1016/j.exptthermflusci.2016.11.012
- [4] Lanjewar, A., Bhagoria, J.L., & Sarviya, R.M. (2011). Heat transfer and friction in solar air heater duct with W-shaped rib roughness on absorber plate. *Energy*, 36, 531–4541. doi: 10.1016/j.energy.2011.03.054
- [5] Tanda, G. (2011). Performance of solar air heater ducts with different types of ribs on the absorber plate. *Energy*, 36, 6651–6660. doi: 10.1016/j.energy.2011.08.043
- [6] Eiamsa-Ard, S., & Promvongse, P. (2009). Thermal characteristics of turbulent rib-grooved channel flows. *International Communications in Heat and Mass Transfer*, 36, 705–711. doi: 10.1016/j.icheatmasstransfer.2009.03.025
- [7] Bhushan, B., & Singh, R. (2011). Nusselt number and friction factor correlations for solar air heater duct having artificially roughened absorber plate. *Solar Energy*, 85, 1109–1118. doi: 10.1016/j.solener.2011.03.007
- [8] Yadav, A.S., & Bhagoria, J.L. (2014). A numerical investigation of square-sectioned transverse rib roughened solar air heater. *International Journal of Thermal Science*, 79, 111–131. doi: 10.1016/j.ijthermalsci.2014.01.008
- [9] Anil, K., & Man-Hoe, K. (2015). Effect of roughness width ratios in discrete multi V-rib with staggered rib roughness on overall thermal performance of solar air channel. *Solar Energy*, 119, 399–414. doi: 10.1016/j.solener.2015.06.030
- [10] Kamali, R., & Binesh, A.R. (2008). The importance of rib shape effects on the local heat transfer and flow friction characteristics of square ducts with ribbed internal surfaces. *International Communications in Heat and Mass Transfer*, 35, 1032–1040. doi: 10.1016/j.icheatmasstransfer.2008.04.012
- [11] Ramadhan, A.I., Azmi, W.H., Mamat, R., & Hamid, K.A. (2020). Experimental and numerical study of heat transfer and friction factor of plain tube with hybrid nanofluids. *Case Studies*

- in *Thermal Engineering*, 22(1), 100782. doi: 10.1016/j.csite.2020.100782
- [12] Nadeem, S., Ali, S., Alzabut, J., Hamida, M.B.B., & Eldin, S.M. (2023). Numerical investigation of the influence of hybrid nano-fluid on heat transfer in semi-annular channel. *Case Studies in Thermal Engineering*, 44, 102855. doi: 10.1016/j.csite.2023.102855
- [13] Rahman, M.M., Chamkha, A.J., Elmasry, Y., Ullah, I., Pasha, A.A., Sadeghi, M.S., & Galal, A.M. (2022). The heat transfer behavior of MHD micro-polar MWCNT-Fe₃O₄/Water Hybrid Nano-fluid in an inclined \perp shaped cavity with semi-circular heat source inside. *Case Studies in Thermal Engineering*, 38, 102316. doi: 10.1016/j.csite.2022.102316
- [14] Javidan, M., Gorji-Bandpy, M., & Al-Araji, A. (2023). Investigation and simulation of parabolic trough collector with the presence of hybrid nanofluid in the finned receiver tube. *Theoretical and Applied Mechanics Letters*, 13(6), 100465. doi: 10.1016/j.taml.2023.100465
- [15] Rashid, F.L., Aljibori, H.S., Mohammed, H.I., Ameen, A., Ahmad, S., Hamida, M.B.B., & Al-Rubaye, A.H. (2024). Recent advances and developments of the application of hybrid nanofluids in parabolic solar collector energy systems and guidelines for prospects. *Journal of Engineering Research*, 14(1), 246–265. doi: 10.1016/j.jer.2024.04.023
- [16] Kabeel, A.E., Khalil, A., Shalaby, S.M. & Zayed, M.E. (2016). Experimental Investigation of Thermal Performance of Flat and V-Corrugated Plate Solar Air Heaters with and without PCM as Thermal Energy Storage. *Energy Conversion and Management*, 113, 264–272. doi: 10.1016/j.enconman.2016.01.068
- [17] Kabeel, A.E., Khalil, A., Shalaby, S.M., & Zayed, M.E. (2017). Improvement of Thermal Performance of the Finned Plate Solar Air Heater by Using Latent Heat Thermal Storage. *Applied Thermal Engineering*, 123, 546–553. doi: 10.1016/j.applthermaleng.2017.05.126
- [18] Aymen, E.L., El Khadraoui, A., Bouadila, S., Kooli, S., Guizani, A., & Farhat, A. (2016). Solar Air Heater with Phase Change Material: An Energy Analysis and a Comparative Study. *Applied Thermal Engineering*, 107, 1057. doi: 10.1016/j.applthermaleng.2016.07.004
- [19] SunilRaj, B.A., & Eswaramoorthy, M. (2020). Experimental Study on Hybrid Natural Circulation Type Solar Air Heater with Paraffin Wax Based Thermal Storage. *Materials Today: Proceedings*, 23(1), 49–52. doi: 10.1016/j.matpr.2019.06.381
- [20] Kalaiarasi, G., Velraj, R., Vanjeswaran, M.N., & Pandian, N.G. (2020). Experimental Analysis and Comparison of Flat Plate Solar Air Heater with and without Integrated Sensible Heat Storage. *Renewable Energy*, 150, 255–265. doi: 10.1016/j.renene.2019.12.116
- [21] Kumar, B.V., Manikandan, G., Kanna P.R., Taler, D., Taler, J., Nowak-Ocłoń, M., Mzyk, K., & Toh, H.T. (2018). A Performance Evaluation of Solar Air Heater Using Different Shape of Ribs Mounted in Absorber Plate – A Review. *Energies*, 11(11), 3104, 2–20. doi: 10.3390/en11113104
- [22] Kumar, B.V., Manikandan, G., & Kanna, P.R. (2021). Enhancement of heat transfer in SAH with polygonal and trapezoidal shape of the rib using CFD. *Energy*, 234, 121154. doi: 10.1016/j.energy.2021.121154
- [23] Kumar, B.V., Manikandan, G., Kanna, P.R., & Reddy V.P. (2022). Experimental Analysis of Solar Air Heater using Polygonal Ribs in Absorber Plate Integrated with Phase Change Material. *Thermal Science*, 26(4A), 3187–3199. doi: 10.2298/TSCI210518345V
- [24] Elsheikh, A., Faqeha, H., Hammoodi, K.A., Bawahab, M., Fujii, M., Shanmugan, S., Essa, F.A., Abd-Elaziem, W., Ramesh, B., Sathyamurthy, R., & Egiza, M. (2025). Integrating predictive and hybrid Machine Learning approaches for optimizing solar still performance: A comprehensive review. *Solar Energy*, 295, 113536. doi: 10.1016/j.solener.2025.113536
- [25] Shafian, N.B.A., Husen, M.N., Xie, L., & Kim, K. (2025). Predicting High-Performance Perovskite Solar Cells Using AI-based Machine Learning Models. *Materials Today Sustainability*, 31, 101176. doi: 10.1016/j.mtsust.2025.101176
- [26] Khouili, O., Hanine, M., Louzazni, M., Flores, M.A.L., Villena, E.G., & Ashraf, I. (2025). Evaluating the impact of deep learning approaches on solar and photovoltaic power forecasting: A systematic review. *Energy Strategy Reviews*, 59, 101735. doi: 10.1016/j.esr.2025.101735
- [27] Kushwah, A., Kumar, A., & Kumar, S. (2024). Predictive modelling with machine learning of garlic clove for novel designed solar drying system. *Solar Energy*, 284, 113070. doi: 10.1016/j.solener.2024.113070
- [28] Abdelkader, T.K., Sayed, H.A.A., Refai, M., Ali, M.M., Zhang, Y., Wan, Q., Khalifa, I., Fan, Q., Wang, Y., & Abdelhamid, M.A. (2024). Machine learning, mathematical modeling and 4E (energy, exergy, environmental, and economic) analysis of an indirect solar dryer for drying sweet potato. *Renewable Energy*, 227, 120535. doi: 10.1016/j.renene.2024.120535
- [29] Arunkumar, P.M., Balaji, N., Madhankumar, S., & Mohanraj, T. (2024). Prediction of red chilli drying performance in solar dryer with natural energy storage element using machine learning models. *Journal of Energy Storage*, 101, 113825. doi: 10.1016/j.est.2024.113825
- [30] Baz, A., & Patel, S.K. (2024). Efficiency-aware machine-learning driven design of solar harvester for renewable energy application. *Results in Engineering*, 24, 103050. doi: 10.1016/j.rineng.2024.103050
- [31] Madhankumar S., Ilangovan, D., & Viswanathan, K. (2024). Integrating computational models and machine learning for corrugated absorber plate solar collector thermal predictions. *Process Safety and Environmental Protection*, 188, 336–349. doi: 10.1016/j.psep.2024.05.069
- [32] Nidhul, K., Thummar, D., Yadav, A.K., & Anish, S. (2023). Machine learning approach for optimization and performance prediction of triangular duct solar air heater: A comprehensive review. *Solar Energy*, 255, 396–415. doi: 10.1016/j.solener.2023.02.022
- [33] Saravanan, A., Parida, S., Murugan, M., Reddy, M.S., Elumalai, P.V., & Dash, S.K. (2023). Thermal performance prediction of a solar air heater with a C-shape finned absorber plate using RF, LR and KNN models of Machine learning. *Thermal Science and Engineering Progress*, 38, 101630. doi: 10.1016/j.tsep.2022.101630
- [34] Zohdi, T.I. (2023). A machine-learning digital-twin for rapid large-scale solar-thermal energy system design. *Computer Methods in Applied Mechanics and Engineering*, 412, 115991. doi: 10.1016/j.cma.2023.115991
- [35] Chilambarasan, L., Thangarasu, V., & Ramasamy, P. (2024). Solar flat plate collector's heat transfer enhancement using grooved tube configuration with alumina nanofluids: Prediction of outcomes through artificial neural network modelling. *Energy*, 289, 129953. doi: 10.1016/j.energy.2023.129953
- [36] Yadav, A.K., Yadav, V., Kumar, A., Kumar, R., Lee, D., & Singh, T. (2024). Novel feature selection based ANN for optimal solar panels tilt angles prediction in microgrid. *Case Studies in Thermal Engineering*. 61, 104853. doi: 10.1016/j.csite.2024.104853
- [37] Ganeshkumar, P., Sivalingam, V., Prasanth, B.M., Kumar, K.S., Kim, S.C., Vanaraj, R., Alodhayb, A.N., & Pandiaraj, S. (2024).

- Predicting thermal performance in solar air heaters with V-corrugated, shot-blasted absorber plate, and black pebble-based sensible heat storage: A machine learning approach. *Journal of Energy Storage*, 100, 113450. doi: 10.1016/j.est.2024.113450
- [38] Thakur, A., Kumar, R., Lee, D., Kumar, S., Kumar, P., Asif, M., & Ağbulut, Ü. (2024). Improved performance evaluation of a dual passage solar thermal collector using staggered turbulators. *Thermal Science and Engineering Progress*, 53, 102668. doi: 10.1016/j.tsep.2024.102668
- [39] Rajan, A., & Reddy, K.S. (2024). Integrated optical-thermal model and deep learning technique to estimate the performance of a conical cavity receiver coupled solar parabolic dish collector. *Energy Conversion and Management*, 301, 118052. doi: 10.1016/j.enconman.2023.118052
- [40] Bingbin, G., Dianwei Q., Zhou, J., Qian, N., & Zhang, L. (2024). Interpretable machine learning study of a collector based on combined twisted-tape and wavy-tape inserts. *Case Studies in Thermal Engineering*, 63, 105236. doi: 10.1016/j.csite.2024.105236
- [41] Gupta, A., Das, B., Arslan, E., Das, M., Kosan, M., & Can, O.F. (2024). Artificial neural networks based computational and experimental evaluation of thermal and drying performance of partially covered PVT solar dryer. *Process Safety and Environmental Protection*, 183, 1170–1185. doi: 10.1016/j.psep.2024.01.068
- [42] Alawi, O.A., Kamar, H.M., Salih, S.Q., Abba, S.I., Ahmed, W., Homod, R.Z., Jamei, M., Shafik, S.S., & Yaseen, Z.M. (2024). Development of optimized machine learning models for predicting flat plate solar collectors thermal efficiency associated with Al₂O₃-water nanofluids. *Engineering Applications of Artificial Intelligence*, 133, Part B, 108158. doi:10.1016/j.engappai.2024.108158.
- [43] Alade, I.O., Rahman, M.A.A., & Saleh, T.A. (2019). Predicting the specific heat capacity of alumina/ethylene glycol nanofluids using support vector regression model optimized with Bayesian algorithm. *Solar Energy*, 183, 74–82. doi: 10.1016/j.solener.2019.02.060
- [44] Alade, I.O., Rahman, M.A.A., & Saleh, T.A. (2019). Modeling and prediction of the specific heat capacity of Al₂O₃/water nanofluids using hybrid genetic algorithm/support vector regression model. *Nano-Structures and Nano-Objects*, 17, 103–111. doi:10.1016/j.nanoso.2018.12.001
- [45] Meybodi, M.K., Naseri, S., Shokrollahi, A., & Daryasafar, A. (2015). Prediction of viscosity of water-based Al₂O₃, TiO₂, SiO₂, and CuO nanofluids using a reliable approach. *Chemometrics and Intelligent Laboratory Systems*, 149(5), 60–69. doi: 10.1016/j.chemolab.2015.10.001
- [46] Baghban, A., Jalali, A., Shafiee, M., Ahmadi, M.H., & Chau, K. (2019). Developing an ANFIS-based swarm concept model for estimating the relative viscosity of nanofluids. *Engineering Applications of Computational Fluid Mechanics*, 13(1), 26–39. doi: 10.1080/19942060.2018.1542345
- [47] Wu, S., Wang, C., & Tang, R. (2022). Optical efficiency and performance optimization of a two-stage secondary reflection hyperbolic solar concentrator using machine learning. *Renewable Energy*, 188, 437–449. doi: 10.1016/j.renene.2022.01.117
- [48] Alhamayani, A. (2024). Performance analysis and machine learning algorithms of parabolic trough solar collectors using Al₂O₃-MWCNT as a hybrid nanofluid. *Case Studies in Thermal Engineering*, 57, 104321. doi: 10.1016/j.csite.2024.104321
- [49] Said, Z., Sharma, P., Sundar, L.S., Li, C., Tran, D.C., Pham, N.D.K., & Nguyen, X.P. (2022). Improving the thermal efficiency of a solar flat plate collector using MWCNT-Fe₃O₄/water hybrid nanofluids and ensemble machine learning. *Case Studies in Thermal Engineering*, 40, 102448. doi: 10.1016/j.csite.2022.102448
- [50] Al-Rashed, A.A.A.A., Alnaqi, A.A., & Alsarraf, J. (2021). Numerical investigation and neural network modeling of the performance of a dual-fluid parabolic trough solar collector containing non-Newtonian water-CMC/Al₂O₃ nanofluid. *Sustainable Energy Technologies and Assessments*, 48, 101555. doi: 10.1016/j.seta.2021.101555
- [51] Bouzgou, H., & Gueymard, C.A. (2017). Minimum redundancy – Maximum relevance with extreme learning machines for global solar radiation forecasting: Toward an optimized dimensionality reduction for solar time series. *Solar Energy*, 158, 595–609. doi: 10.1016/j.solener.2017.10.035
- [52] Fuxi, S., Sina, N., Sajadi, S.M., Mahmoud, M.Z., Abdelrahman, A., & Aybar, H.Ş. (2022). Artificial neural network modeling to examine spring turbulators influence on parabolic solar collector effectiveness with hybrid nanofluids. *Engineering Analysis with Boundary Elements*, 143, 442–456. doi: 10.1016/j.enganabound.2022.06.026
- [53] Zayed, M.E., Kabeel, A.E., Shboul, B., Ashraf, W.M., Ghazy, M., Irshad, K., Rehman, S., & Zayed, A.A.A. (2023). Performance augmentation and machine learning-based modeling of wavy corrugated solar air collector embedded with thermal energy storage: Support vector machine combined with Monte Carlo simulation. *Journal of Energy Storage*, 74, 109533. doi: 10.1016/j.est.2023.109533
- [54] Badji, A., Benseddik, A., Boukhelifa, A., Bensaha, H., Erregani, R.M., Bendriss, A., Bouhoun, S., Nettari, C.H., Kaouane, M., & Lalmi, D. (2023). Solar air heater with underground latent heat storage system for greenhouse heating: Performance analysis and machine learning prediction. *Journal of Energy Storage*, 74, Part B, 109548. doi: 10.1016/j.est.2023.109548
- [55] Mausam, K., Singh, S., Ghosh, S.K., & Singh, R.P. (2024). Thermal performance modelling of solar flat plate parallel tube collector using ANN. *Energy*, 303, 131940. doi: 10.1016/j.energy.2024.131940
- [56] Mustafa, J., Alqaed, S., & Sharifpur, M. (2022). Numerical study on performance of double-fluid parabolic trough solar collector occupied with hybrid non-Newtonian nanofluids: Investigation of effects of helical absorber tube using deep learning. *Engineering Analysis with Boundary Elements*, 140, 562–580. doi: 10.1016/j.enganabound.2022.04.033
- [57] ASHRAE Standard 93. (2003). *Method of testing to determine the thermal performance of solar collectors*. American Society of Heating, Refrigeration and Air Conditioning Engineers, 30329, Atlanta, USA.
- [58] Taler, D., & Taler, J. (2017). Simple heat transfer correlations for turbulent tube flow. *E3S Web of Conferences*, 13, 02008. doi: 10.1051/e3sconf/20171302008
- [59] Ghritlahre, H.K., Chandrakar, P., & Ahmad, A. (2020). Application of ANN model to predict the performance of solar air heater using relevant input parameters. *Sustainable Energy Technologies and Assessments*, 40, 100764. doi: 10.1016/j.seta.2020.100764
- [60] Kumar, G.H., Purvi, C., & Ashfaq, A. (2020). Solar air heater performance prediction using artificial neural network technique with relevant input variables. *Archives of Thermodynamics*, 41(3), 255–282. doi: 10.24425/ather.2020.134579
- [61] Ghritlahre, H.K., & Prasad, R.K. (2019). Modelling of back propagation neural network to predict the thermal performance of porous bed solar air heater. *Archives of Thermodynamics*, 40(4), 103–128. doi: 10.24425/ather.2019.131430
- [62] Ghritlahre, H.K., & Verma, M. (2020). Solar air heaters performance prediction using multi-layer perceptron neural network–

- A systematic review. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 41(1), 7682–7699. doi: 10.1080/15567036.2021.1923869
- [63] Ghritlahre, H.K., Chandrakar, P., & Ahmad, A. (2021). A Comprehensive Review on Performance Prediction of Solar Air Heaters Using Artificial Neural Network. *Annals of Data Science*, 8, 405–449. doi: 10.1007/s40745-019-00236-1
- [64] Tschopp, D., Ohnewein, P., Stelzer, R., Feierl, L., Hamilton-Jones, M., Moser, M., & Holter, C. (2023). One year of high-precision operational data including measurement uncertainties from a large-scale solar thermal collector array with flat plate collectors, located in Graz, Austria. *Data in Brief*, 48, 109224. doi: 10.1016/j.dib.2023.109224
- [65] Falope, T.O., Lao, L., & Hanak, D. (2024). A Three-Step Weather Data Approach in Solar Energy Prediction Using Machine Learning. *Renewable Energy Focus*, 50, 100615. doi: 10.1016/j.ref.2024.100615
- [66] Pereira, F., & Silva, C. (2024). Machine learning for monitoring and classification in inverters from solar photovoltaic energy plants. *Solar Compass*, 9, 100066. doi: 10.1016/j.solcom.2023.100066
- [67] Kumar, B.V., Kanna, P.R., Manikandan, G., Taler, D., Taler, J., Sobota, T., & Nowak-Ocłoń, M. (2023). Investigation of thermo-hydraulic performances of artificial ribs mounted in a rectangular duct. *Energies*, 16(11), 4404. doi: 10.3390/en16114404
- [68] Holman, J.P. (2004). *Experimental methods for engineers* (7 th ed.). Tata McGraw-Hill Publishing Company Limited, New Delhi.