



OPEN Comparative analysis of shallow and hybrid deep learning models for predicting the cooling efficiency of nanofluid-cooled photovoltaic panel across multiple materials

Yunus Özdemir¹, Neşe Budak Ziyadanoğulları¹, Enes Bakış^{2,3✉} & Emrullah Acar⁴

Active cooling can mitigate temperature-induced performance losses in photovoltaic (PV) modules, and nanofluids are a promising coolant option. This study develops data-driven models to predict the cooling efficiency of an actively cooled PV panel using seven working fluids: water and Al₂O₃/TiO₂ nanofluids at 0.01%, 0.1%, and 1 vol%. For each fluid, outdoor measurements were collected over six hours at 30-min intervals (13 observations), including inlet/outlet temperatures, electrical variables of cooled and reference panels, and ambient conditions. Shallow regression models (Bayesian Ridge, SVR-RBF, Random Forest) were evaluated using leave-one-out cross-validation, and a hybrid deep learning (CNN+LSTM) model was also tested using k-fold cross-validation. Bayesian Ridge achieved the most consistent performance across materials (RMSE ≈ 0.35–0.66; R² ≈ 0.88–0.98). The hybrid CNN+LSTM reached RMSE as low as 0.28 with R² up to 0.98. SHAP-based interpretability analysis indicates that ambient temperature, irradiance, and the cooled-panel electrical variables are among the most influential predictors. These results show that lightweight ML models can reliably estimate PV cooling performance and reduce repetitive experimentation.

Keywords Photovoltaic panel, Nanofluid-based cooling, Machine learning, Hybrid deep learning, Cooling efficiency

Abbreviations

G	Solar irradiance (W/m ²)
T _{amb}	Ambient temperature (°C)
T _s	PV surface temperature (°C)
T _{in}	Inlet fluid temperature (°C)
T _{out}	Outlet fluid temperature (°C)
V	Output voltage (V)
I	Output current (A)
v _w	Wind speed (m/s)
η	Photovoltaic efficiency (%)
φ	Nanoparticle volume concentration (%)
R ²	Coefficient of determination (-)
R	Correlation coefficient (-)
RMSE	Root mean square error (-)
MSE	Mean squared error (-)
MAE	Mean absolute error (-)
ρ	Fluid density (kg/m ³)
c _p	Specific heat capacity (J/kg K)

¹Department of Mechanical Engineering, Faculty of Engineering and Architecture, Batman University, Batman 72000, Turkey. ²Department of Electrical and Electronics Engineering, Faculty of Engineering, Piri Reis University, Istanbul 34940, Turkey. ³Department of Electrical and Electronics Engineering, Postgraduate Education Institute, Batman University, Batman 72000, Turkey. ⁴Department of Electrical and Electronics Engineering, Faculty of Engineering and Architecture, Batman University, Batman 72000, Turkey. ✉email: ebakis@pirireis.edu.tr

PV	Photovoltaic (-)
PV/T	Photovoltaic/thermal (-)
PCM	Phase change material (-)
ML	Machine learning (-)
DL	Deep learning (-)
RF	Random forest (-)
BR	Bayesian Ridge Regression (-)
GPR	Gaussian Process Regression (-)
SVR	Support vector regression (-)
CNN	Convolutional neural network (-)
LSTM	Long short-term memory (-)
LOOCV	Leave-one-out cross-validation (-)
CV	Cross-validation (-)
SHAP	Shapley additive explanations (-)
GW	Gigawatt (W)
TW	Terawatt (W)
Al ₂ O ₃	Aluminium oxide (-)
TiO ₂	Titanium dioxide (-)
ZnO	Zinc oxide (-)

Introduction

Energy is, in its most general sense, defined as the capacity to perform work. In recent years, renewable energy sources have gained increasing prominence in the generation of electrical power. Resources such as wind, solar, hydropower and geothermal energy now play a critical role in electricity production. Among these, solar energy has been particularly favored due to its ease of installation and relatively low operation and maintenance costs. According to data reported by SolarPower Europe, 597 GW of photovoltaic (PV) capacity was installed in 2024, and the cumulative installed PV capacity worldwide has reached approximately 2.2 TW¹.

Despite the widespread deployment of PV panels and the continuous increase in installed capacity each year, there remains a major drawback: the power conversion efficiency of PV cells decreases with increasing cell temperature. It is commonly reported that for every 1 °C rise in temperature, the panel efficiency drops by approximately 0.5%². Therefore, enhancing the long-term efficiency and extending the service life of PV panels has become a critical research issue. In order to mitigate this efficiency degradation, both passive and active cooling strategies have been investigated in the literature. Active cooling systems employ auxiliary components such as fans, pumps and similar devices to drive the coolant. Although they generally provide higher cooling performance compared to passive techniques, they require additional power-consuming components, which make the system more complex and increase both the initial investment and maintenance costs³. In passive cooling systems, on the other hand, no power-consuming device is used and heat transfer is mainly achieved by natural convection. Phase change materials (PCMs), heat sinks and fins are commonly employed as passive cooling elements⁴. In active cooling applications, water or air is most frequently used as the working fluid. However, in recent years, nanofluids—obtained by dispersing nanoparticles into base fluids such as water or ethylene glycol—have been increasingly utilized in order to enhance the heat transfer performance.

In photovoltaic (PV) panel cooling studies, several parameters have a pronounced influence on the cooling performance, including the design of the cooling system, the type of working fluid, the volumetric flow rate and, when a nanofluid is employed, the nanoparticle concentration (volume/mass fraction). In addition, the ambient and operating conditions under which the experiments are conducted significantly affect the overall performance. For instance, in an air-cooling configuration using a fan, the electrical output power of the cooled panel was reported to increase by 10.7% compared with the uncooled reference panel⁵. Selected studies from the literature addressing these aspects are summarised below:

Mojtaba et al.⁶ investigated the effect of spraying water onto the front surface of a PV panel on its electrical performance. In their study, the influence of several geometric and operating parameters, such as nozzle inclination angle, nozzle–panel distance and the number of nozzles, was examined. The results indicated that positioning the nozzles closer to the panel surface had a beneficial effect and that a nozzle angle of 15° yielded the maximum enhancement in efficiency, leading to an overall increase in panel efficiency of 25.86% under the optimum operating conditions. In another configuration⁷ where a microchannel was integrated on the rear side of the PV panel, a water flow rate of 3 L·min⁻¹ resulted in a 14% increase in output power and a 3% increase in electrical efficiency. In a further study⁸ in which both water and a nanofluid were employed as coolants, the panel temperature at noon was reduced by 10 °C when using water, whereas an Al₂O₃–water nanofluid achieved a reduction of 20 °C. In a separate work utilizing Al₂O₃ nanofluid⁹ it was reported that employing 0.9 wt% Al₂O₃ led to a 21% enhancement in electrical efficiency. Mohammad et al. used different nanofluids¹⁰ in their experiments and observed that, while the electrical efficiency of the uncooled panel was 12.73%, the corresponding values for water, TiO₂, ZnO and Al₂O₃ coolants were 13.41%, 13.63%, 13.59% and 13.44%, respectively.

In active cooling systems, it is necessary to repeat the experiments for each value of the relevant operating parameters. Moreover, to obtain reliable data for each working fluid, measurements must be recorded continuously throughout the day, which renders the experimental process both labor-intensive and time-consuming. In addition, variations in ambient conditions on different days (such as air temperature, humidity, wind speed and cloud cover) hinder the establishment of a consistent correlation between successive experiments.

To overcome these limitations, recent studies on active PV cooling have increasingly employed machine learning (ML) and deep learning (DL) techniques for performance prediction. By using such data-driven models,

the need to repeat experiments on a daily basis can be substantially reduced, leading to a significant saving in experimental time and effort. Furthermore, these approaches enable the influence of any given parameter variation on system performance to be evaluated within seconds. Some of the studies conducted in this field in the literature are as follows;

Cao et al.¹¹ investigated the electrical performance of photovoltaic/thermal collectors cooled with various nanofluids using machine-learning-based modeling and evolutionary optimization. They employed performance metrics such as R^2 , RMSE, and MSE to assess prediction accuracy, reporting high agreement between model outputs and experimental data. Their results showed that nanofluid cooling substantially enhanced electrical efficiency and that the ML models offered highly reliable prediction capability. Safae et al.¹² analyzed the thermal and electrical power generation of a photo-thermal system by integrating multiple machine-learning models. Model performance was evaluated using R^2 , RMSE, and MAE, achieving values close to unity for R^2 and low error scores. Their findings demonstrated that the hybrid ML framework accurately captured the operational behavior of the system and effectively predicted energy outputs. Margoum et al.¹³ explored hybrid nanofluid-based solar collectors and implemented advanced artificial intelligence models to optimize both thermal and electrical efficiencies. They assessed predictive accuracy using R^2 , MSE, and RMSE, obtaining exceptionally high scores, including an R^2 value approaching 0.99, indicating near-perfect prediction. Their results confirmed that AI-based modeling can significantly enhance the design and optimization of hybrid collector systems. Jakhar et al.¹⁴ developed machine-learning predictive models to evaluate a photovoltaic/thermal collector incorporating nanofluids and geothermal cooling. Although specific numerical error metrics were not prominently reported, the authors validated their models against experimental observations and confirmed strong predictive reliability. The study concluded that combining geothermal cooling with nanofluids enhances system stability and that ML modeling provides valuable insights for optimal system configuration. Diwania et al.¹⁵ investigated thermo-electrical performance optimization of hybrid photovoltaic/thermal systems using machine-learning techniques. They primarily relied on the R^2 metric to evaluate model accuracy, reporting strong coefficients of determination for both thermal and electrical predictions. Their results highlighted the capability of ML algorithms to identify optimal operating conditions and improve hPVT system efficiency. Alqaed et al.¹⁶ modeled nanofluid flow within a solar thermal panel containing phase-change materials using a machine-learning-based framework. Their models were validated using RMSE, MAE, and R^2 , and the low error values indicated that the ML algorithms accurately captured the nonlinear and transient heat-transfer behavior. Their findings showed that integrating PCM with nanofluids effectively improved thermal regulation and enhanced system performance. Khudhur et al.¹⁷ assessed the performance of nanofluid-enhanced PV-T solar collectors using machine-learning techniques. Although detailed numerical metrics were not broadly disclosed, their analysis showed that the ML models accurately predicted system outputs and confirmed the efficiency gains associated with nanofluid usage. Their results supported the applicability of ML for performance evaluation in PV-T configurations. Abdulrahman¹⁸ examined the cooling of photovoltaic panels using heat pipes coupled with nanofluids through a hybrid machine-learning methodology. He evaluated predictive performance using RMSE, reporting an error of approximately 3.95 W, which reflects strong model precision. His findings confirmed that the combined cooling approach significantly reduced panel temperature and that the ML framework effectively assessed system performance. Thermal management remains one of the primary factors limiting the electrical efficiency and long-term reliability of photovoltaic (PV) systems. Recent experimental and thermal analysis studies have demonstrated that nanofluid-based cooling techniques can significantly enhance heat transfer performance due to improved thermophysical properties compared to conventional fluids. However, the cooling effectiveness of nanofluids is highly dependent on material type, nanoparticle composition, and operating conditions, leading to complex nonlinear thermal behavior. As highlighted in the literature, conventional empirical or physics-based models often struggle to accurately capture these interactions, particularly when multiple cooling materials are involved, thereby motivating the adoption of data-driven modelling approaches for reliable performance prediction¹⁹. Sharaby et al. (2024) provided an extensive state-of-the-art review regarding the application of various nanofluids in PV/T systems, highlighting the theoretical potential of metallic and metal-oxide nanoparticles. However, while their work offers a robust theoretical foundation, it lacks a specific predictive framework that utilizes deep learning to forecast real-time system behaviour. Our study addresses this by transitioning from theoretical review to a practical, data-driven forecasting model using a hybrid CNN+LSTM architecture²⁰. Sharshir et al. (2025) conducted a critical analysis of degradation mechanisms and stability challenges in perovskite solar cells, emphasizing the importance of thermal stability for long-term efficiency. While their research identifies the critical nature of thermal management, it focuses primarily on material degradation rather than active cooling optimization through computational intelligence. Our work supplements this by demonstrating how active nanofluid cooling, modelled with high accuracy ($R^2=0.981$), can be optimized to maintain thermal stability under fluctuating outdoor conditions²¹. Sharaby et al. (2025) investigated the performance of fixed and sun-tracking PV systems integrated with spray cooling, providing valuable experimental data on mechanical cooling methods. Despite these advancements, their study does not explore the comparative benefits of different nanofluid concentrations or the use of temporal deep learning for predictive maintenance. In contrast, our manuscript benchmarks seven different configurations and utilizes SHAP analysis to interpret the importance of climatic inputs on cooling efficiency²². Sharaby et al. (2025) focused on the implementation of hybrid nanofluid cooling through a 3E (Energy, Exergy, and Environmental) analysis approach, achieving significant performance improvements. While their 3E analysis is comprehensive for system assessment, it does not incorporate hybrid temporal models like CNN+LSTM to handle the stochastic nature of experimental time-series data. Our research fills this gap by achieving low error rates (RMSE=0.28 and MAE=0.19) while specifically targeting the time-dependent thermal interactions of Al_2O_3 and TiO_2 nanofluids²³.

In recent years, machine learning techniques have been increasingly employed for predicting the thermal and electrical performance of photovoltaic and energy systems^{24,25}. Several studies have reported promising

predictive accuracy using data-driven approaches; however, most of these investigations focus on a single cooling configuration or a limited set of input features and learning algorithms²⁴. More recent works have extended these models to hybrid or advanced machine learning structures, yet they often neglect material diversity and rely on simplified cooling scenarios¹². Furthermore, studies employing deep learning typically do not integrate temporal–spatial feature learning in conjunction with systematic material-level comparisons²⁶. Consequently, the combined effects of cooling material variability and hybrid deep learning architectures remain insufficiently explored. In contrast, the present study systematically evaluates multiple cooling materials and integrates shallow machine learning models with a CNN+LSTM hybrid deep learning architecture, thereby addressing critical methodological and application-oriented gaps identified in recent literature.

The main contributions of the proposed study and the limitations in the literature are presented in Table 1.

As critically synthesized in Table 1, while previous works have significantly advanced PV/T modelling, gaps remain regarding multi-material benchmarking and the application of temporal hybrid architectures. To address these limitations, this study evaluates seven distinct cooling configurations—including pure water and two metal-oxide nanofluids (Al_2O_3 and TiO_2) at three volume concentrations (0.01%, 0.1%, and 1.0%)—using a hybrid CNN+LSTM framework under real-world outdoor conditions.

The remainder of this paper is structured as follows: Section II provides a comprehensive description of the materials used, the experimental methodology, and the computational framework, including the development of the hybrid CNN+LSTM model, SHAP (Shapley Additive Explanations) analysis, and the cross-validation procedures. Section III presents the Results and Discussion, focusing on the thermal performance of the configurations and the comparative benchmarking of the models. Finally, Section IV concludes the study by summarizing the key findings and highlighting their practical implications for solar energy systems.

Materials and methods

Materials

In this study, two identical PV panels with 50 W power were used. The technical specifications of the modules are presented in Table 2. One of the modules was cooled from the rear side using copper tubes with an inner diameter of 3 mm, whereas no cooling arrangement was applied to the other module, which served as the reference. In order to enhance the cooling performance, cylindrical fins with a diameter of 1.25 mm were attached to the tubes through which the coolant flows. The rear-side cooling loop consisted of copper tubes with cylindrical fins to enhance heat transfer. The cylindrical fins (diameter 1.25 mm) were soldered to the copper tubes, ensuring durable attachment and good thermal contact between the extended surfaces and the tube wall. The assembled tube network was then mounted on the rear surface of the PV module and fixed using silicone adhesive at nine discrete locations distributed along the tube layout. Discrete (point-wise) bonding was intentionally adopted instead of continuous bonding to limit the formation of a continuous adhesive layer, which could otherwise introduce additional contact thermal resistance between the copper tube and the PV backsheet. This mounting strategy ensured mechanical stability while preserving effective conductive heat transfer from the PV rear surface to the coolant. The cooling tubes and their placement on the rear surface of the PV module are shown in Fig. 1. The list of materials is also shown in Table 4.

The experimental study was conducted at Batman University, West Raman Campus. According to the Köppen–Geiger climate classification, Batman has a Csa climate, characterized by mild winters and extremely hot, dry summers. In line with these conditions, the experiments were performed between 28 June 2022 and 6 July 2022. Moreover, given Batman's latitude of 37.88° and longitude of 41.12° , the PV panels were installed facing south at a tilt angle of 31° .

Although the two PV modules are of the same brand and model, minor deviations in current and voltage may occur due to manufacturing tolerances. To eliminate these discrepancies, both panels were first operated without any cooling, and their current and voltage values were recorded. A correction factor was then determined to equalize the electrical output of the two panels, and this factor was subsequently applied during the experiments.

As cooling fluids, water and Al_2O_3 - and TiO_2 -based nanofluids with volumetric concentrations of 0.01%, 0.1% and 1% were employed. The nanoparticle concentration was specified on a volumetric basis in order to avoid the influence of density differences between the nanoparticles and to ensure an approximately equal number of particles in a given dispersion volume. In other words, the “particle population” per unit volume is kept comparable, which enables a more reliable assessment of the associated heat transfer enhancement and efficiency improvement. Three volumetric concentrations (0.01%, 0.1% and 1%) were selected to span a practically relevant range for PV cooling while sampling concentration effects efficiently on a logarithmic scale. Many PV/PVT nanofluid-cooling studies operate at sub-1% loadings (e.g., 0.1–0.5%) and already report measurable performance enhancement in this region. Moreover, increasing nanoparticle loading simultaneously increases viscosity, which raises pressure drop and pumping power and can offset heat-transfer gains. Higher particle loadings also intensify dispersion-stability challenges; aggregation/sedimentation can degrade heat transfer and may cause fouling or blockage in small-diameter flow passages. For these thermo-hydraulic and stability reasons, 1 vol% was adopted as an upper practical bound in the present closed-loop PV cooling system. We note that the selected levels were not assumed to represent a global optimum; rather, they define a realistic operational range (≤ 1 vol%) over which the cooling and electrical-performance impacts can be reliably quantified and modeled.

In the experimental setup, fluid circulation was provided by a pump, and the volumetric flow rate was kept constant at $0.78 \text{ L}\cdot\text{min}^{-1}$ throughout the tests. The volumetric flow rate was fixed at $0.78 \text{ L}\cdot\text{min}^{-1}$ to maintain a consistent operating point for comparing different coolants while achieving a favorable heat-transfer-to-pumping-power trade-off. With a tube inner diameter of 3 mm, this flow rate corresponds to a mean velocity of approximately $1.84 \text{ m}\cdot\text{s}^{-1}$ and a Reynolds number of $\text{Re} \approx 6.2 \times 10^3$ for water at $\sim 25^\circ\text{C}$, which is above commonly cited thresholds for turbulent internal flow (often ~ 4000 in some sources, with transition frequently discussed in the ~ 2300 – 4000 range). This choice therefore promotes turbulence-driven convective heat transfer enhancement.

Study (ref)	Cooling approach	Working fluid/nanofluid type	Operating conditions	Inputs/targets	ML/DL method and validation	Best metrics	Literature limitations addressed
19	Solar adsorption cooling	Al ₂ O ₃ , TiO ₂ - water	Outdoor experimental	Amb. Temp, Rad/thermal COP	Experimental/analytical	R ² >0.90 (Exp. correlation)	Lack of data-driven predictive modeling for performance forecasting
27	PV/T cooling (review)	Various (SiC, Al ₂ O ₃ , CuO)	Review-based	N/A	Review study	N/A (review study)	No direct experimental modeling or Deep Learning (DL) application
11	PVT collector cooling	Water (conventional)	Indoor simulated	Inlet Temp, Flow/Tout	ML + genetic algorithm	R ² =0.999, RMSE = 0.012	Focuses on conventional fluids; lacks material and concentration diversity
12	Photo-thermal system	Water/glycol	Indoor/outdoor mix	Solar flux, Flow/Temp gain	Integrated ML (cross-validation)	R ² =0.998, MSE = 0.004	Nanofluid cooling and varying material concentrations not considered
15	PVT + geothermal	Water (conventional)	Outdoor experimental	Amb. Temp, wind/efficiency	ML predictive modeling	R ² =0.985, MAE = 0.45	Did not utilize hybrid deep learning architectures or nanofluids
17	hPVT system	Water/Al ₂ O ₃	Outdoor experimental	Radiation, wind/power yield	Gaussian process regression	R ² =0.992, RMSE = 0.15	Lack of hybrid DL (CNN/LSTM) for temporal feature extraction
18	Heat-pipe PV cooling	Al ₂ O ₃ - water	Outdoor experimental	Rad, Amb. Temp/Tpanel	Hybrid ANN-PSO (train/test)	RMSE = 3.95 W	Limited to a single nanofluid; utilizes shallow hybrid ML models
28	PV/T cooling	ZnO - water	Outdoor experimental	Solar Rad, Flow/Tpanel	Experimental analysis	Error < 5% (uncertainty)	Lacks predictive algorithms and Machine Learning (ML) integration
29	PV/T cooling	SiO ₂ - water	Outdoor experimental	Solar Rad, Temp/efficiency	ML Regression (K-fold)	R ² =0.989, MSE = 0.021	Hybrid temporal deep learning (CNN+LSTM) not explored.
Proposed study	Active PV cooling	Water + Al ₂ O ₃ & TiO ₂	Real outdoor experimental	Rad, Tamb, Wind, V, I/efficiency	Hybrid CNN+LSTM, Bayesian Ridge, Random Forest	R ² =0.981, RMSE = 0.28, MAE = 0.19	Addresses: (i) Multi-material/concentration comparison; (ii) Hybrid DL (CNN+LSTM) architecture; (iii) High-fidelity outdoor benchmarking

Table 1. Key contributions of the proposed study and literature limitations.

However, substantially increasing flow rate in small-diameter tubing would markedly increase pressure drop and auxiliary pumping demand, and this penalty becomes more significant for nanofluids due to higher effective viscosity. Hence, 0.78 L·min⁻¹ was selected to enter (near-)turbulent conditions without incurring excessive friction losses. A schematic representation of the experimental setup is shown in Fig. 2.

The current and voltage values of the panels were measured and logged via the resistive loading mechanism, and the influence of cooling on electrical power output was assessed by comparing the power generated by the cooled and reference panels. Simultaneously, panel efficiencies were determined as a function of the measured solar irradiance and panel surface area, allowing the impact of cooling on electrical efficiency to be evaluated.

Type	Monocrystalline
Model	SP-50 M
Max power (P_{max})	50 W _p (\pm %5)
Max voltage (V_{mp})	20.70 V
Max current (I_{mp})	2.42 A
Open circuit voltage (V_{oc})	23.80 V
Short circuit current (I_{sc})	2.54 A
Nominal efficiency	%17
Temperature coefficient of power	-0.5 (\pm 0.05) %/°C
Max system voltage	1000 V
Weight	3,8 kg
Dimensions	679 × 433 × 20 mm

Table 2. Specification of PV panel.

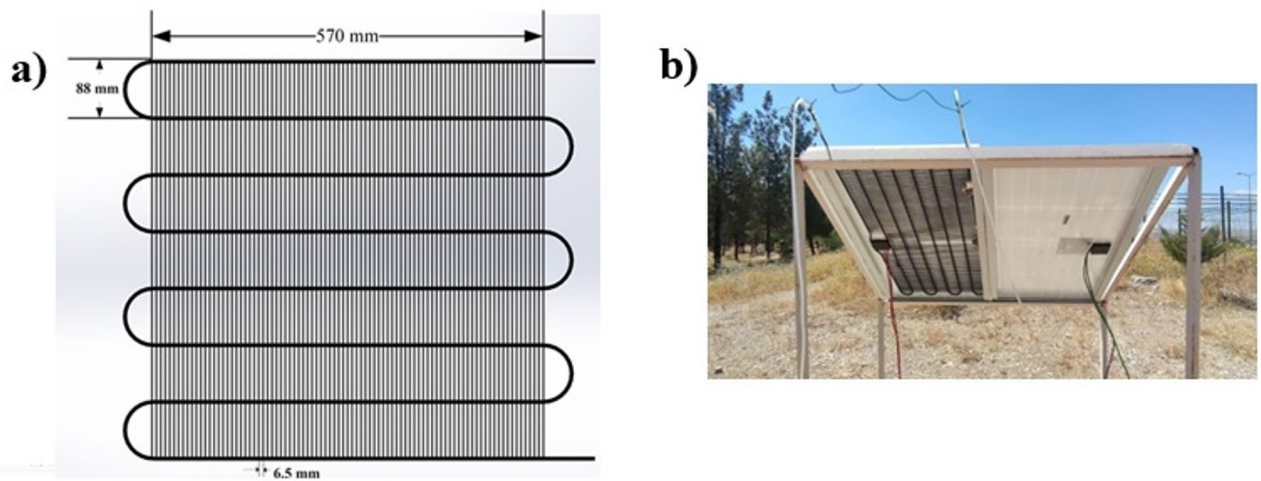


Fig. 1. (a) Schematic view of the pipe, (b) the assembly state of the panel.

The electrical power of the panel was calculated using Eq. (1), and the electrical efficiency was obtained from Eq. (2):

$$Q_e = V \times I \tag{1}$$

Q_e : The resulting electrical power (Watt).

V: Voltage (Volt).

I: Current (Amper)

$$\eta_e = \frac{Q_e}{G \times A} \tag{2}$$

η_e : Electrical efficiency.

G: Irradiation (W/m²).

A: Panel surface area (m²).

The overall uncertainty R is evaluated as a function of the independent variables $x_1, x_2, x_3, \dots, x_n$, where $w_1, w_2, w_3, \dots, w_n$ denote the uncertainties associated with these variables. The accuracies of the measurement devices used in this study are listed in Table 3. Using Eq. (3), the combined experimental uncertainty was found to be less than 2%, indicating a high level of measurement accuracy. Additionally, the list of materials used in the study is presented in Table 4.

$$WR = \sqrt{\left(\frac{\partial R}{\partial x_1} w_1\right)^2 + \left(\frac{\partial R}{\partial x_2} w_2\right)^2 + \dots + \left(\frac{\partial R}{\partial x_n} w_n\right)^2} \tag{3}$$

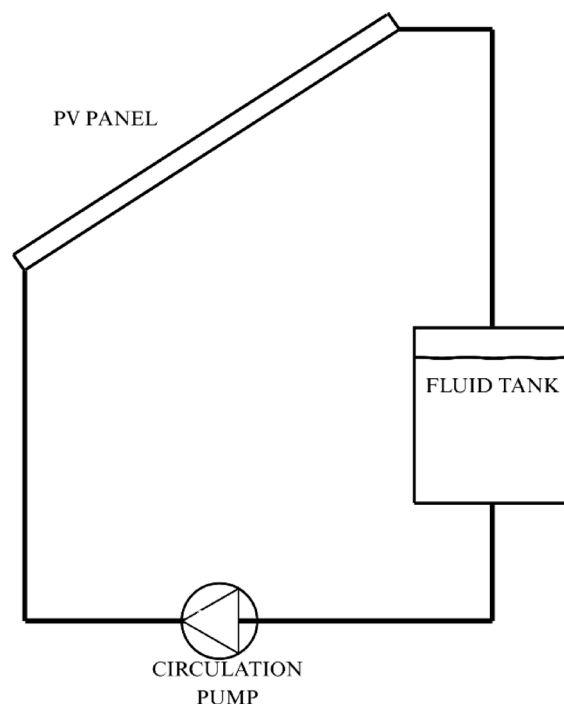


Fig. 2. Schematic representation of the experimental setup.

Measurement parameter	Device - model	Accuracy
Surface temperature	Wellhise HT 9815	$\pm 0,2$ °C
Solar radiation	Cem DT-1307	± 10 W/m ²
Ambient temperature	UNI-T UT363S	± 2 °C
Wind speed	UNI-T UT363S	± 0.5 m/s
Voltmeter	Fluke 179	$\pm \%0.09$ V
Ammeter	Fluke 179	$\pm \%1$ A

Table 3. The measurement devices and their accuracies.

Material1	Water
Material2	Al ₂ O ₃ (0.01%)
Material3	Al ₂ O ₃ (0.1%)
Material4	Al ₂ O ₃ (1%)
Material5	TiO ₂ (0.01%)
Material6	TiO ₂ (0.1%)
Material7	TiO ₂ (1%)

Table 4. List of materials.

Methods

In this study, we evaluated and compared the performance of several conventional machine learning models^{30–32} and a hybrid deep learning model^{33,34} to predict the cooling efficiency of different materials under controlled experimental conditions. The dataset was obtained from a cooling tank setup in which seven distinct materials were tested for six hours with measurements taken every thirty minutes. Each material produced thirteen observations. The recorded variables included tank temperature, pipe inlet and outlet temperatures, current and voltage values for both cooled and normal panels, as well as ambient conditions such as wind speed, solar radiation, and ambient temperature. In addition, surface temperatures were measured from five specific regions (top right, top left, middle, bottom right, and bottom left) for both cooled and normal samples. The dependent variable in the study was the measured cooling efficiency for each material.

The experimental setup and raw data used in this study were originally collected and published in our previous work³⁵. In the current research, these data are not reused for reporting experimental outcomes but are employed solely for developing and evaluating new machine learning and hybrid deep learning models. The novelty of the present work lies in the proposed data-driven prediction framework, comparative algorithmic analysis, and SHAP-based interpretability assessment.

Before model training, all features were standardized using the StandardScaler method to ensure that differences in measurement scales did not bias model learning. Three regression algorithms were implemented: Bayesian Ridge Regression, Support Vector Regression (SVR) with an RBF kernel, and a Random Forest Regressor with limited tree depth to prevent overfitting, considering the relatively small dataset size. Furthermore, an ensemble model was created by taking the average of the two models with the lowest RMSE values, aiming to improve the overall prediction stability and robustness.

Model performance was evaluated using Leave-One-Out Cross-Validation (LOOCV) for the regression-based models. In this method, each observation was once treated as a test sample while the remaining data were used for training. The model's predictive ability was then quantified by calculating the Root Mean Square Error (RMSE) and the Coefficient of Determination (R^2) for each iteration, and the results were averaged.

For the deep learning approach, a hybrid CNN+LSTM architecture was designed to model the temporal behavior of the system. The CNN layers were responsible for extracting short-term spatial patterns from the sequential data windows, while the LSTM layers captured the underlying temporal dependencies between observations. The model was trained using the Adam optimizer with the Mean Squared Error (MSE) as the loss function, while performance evaluation was conducted using the Root Mean Square Error (RMSE) and the coefficient of determination (R^2). To avoid overfitting, early stopping and learning rate reduction strategies were applied during training. The CNN+LSTM model was evaluated through K-Fold Cross-Validation (up to five folds), which allowed a reliable assessment of its performance even with a limited dataset.

To gain insights into model interpretability, SHAP (SHapley Additive exPlanations) analysis was conducted after training. This approach quantified the contribution of each input feature to the predicted efficiency values, helping to identify which physical parameters had the greatest influence on the cooling performance.

All experiments and analyses were carried out in the Python environment (Jupyter Notebook) using open-source libraries such as pandas, NumPy, scikit-learn, TensorFlow, and SHAP. The imported libraries and Their Purposes were presented in Table 5. Comprehensive summary used all models was demonstrated in Table 6. The flow chart of proposed study is demonstrated in Fig. 3.

Bayesian ridge regression

Bayesian Ridge Regression is a linear regression technique that incorporates Bayesian inference to automatically determine the optimal level of regularization for the model coefficients. This approach helps prevent overfitting, especially in datasets with correlated features or limited samples, by penalizing large coefficients through a probabilistic prior. In this study, BRR was used to model the relationship between cooling system parameters and the resulting efficiency, providing predictions. Its probabilistic framework makes it particularly suitable for small experimental datasets, where conventional linear regression may fail to generalize³⁶.

Support vector regression

Support Vector Regression is a flexible, nonlinear regression method based on the principles of support vector machines. It seeks to find a function that deviates from observed targets by no more than a specified margin while maintaining model complexity. SVR is robust to outliers and capable of capturing nonlinear patterns in the data. In this study, SVR was applied to predict cooling efficiency from the input features, allowing the model

Library/module	Purpose
os	Operating system utilities (e.g., file path handling)
random	Python random number generation, for reproducibility
numpy	Numerical computing, array manipulation, reshaping sequences
pandas	Reading Excel files and handling tabular data (DataFrames)
matplotlib.pyplot	Visualization of true vs. predicted values and performance plots
sklearn.preprocessing.StandardScaler	Standardizes features and target for better model training. Prevents scale issues
sklearn.model_selection.KFold	K-Fold Cross-Validation: splits data into k folds for model evaluation. Reduces overfitting bias
sklearn.metrics.mean_squared_error	Computes MSE, used for RMSE calculation (performance metric)
sklearn.metrics.r2_score	Computes R^2 score, measures proportion of variance explained by the model
tensorflow	Deep learning framework used for defining and training neural networks
tensorflow.keras.layers	Provides core neural network layers like Conv1D, LSTM, Dense, Dropout, MaxPooling1D
tensorflow.keras.models	Functions for building, compiling, and managing models (Model, Sequential)
tensorflow.keras.regularizers	Provides L2 regularization to prevent overfitting
tensorflow.keras.callbacks.EarlyStopping	Stops training when validation loss stops improving, prevents overfitting
tensorflow.keras.callbacks.ReduceLROnPlateau	Reduces learning rate if the validation loss plateaus, improving convergence
tensorflow.keras.backend	Provides backend functions like K.clear_session() to reset models and release memory between folds

Table 5. Imported libraries and their purposes.

Model	Type/description	Hyperparameters (used in this study)	Evaluation method	Notes/key points
Bayesian Ridge Regression (BRR)	Probabilistic linear regression with Bayesian L2 regularization	n_iter = 300, alpha_1 = 1e-6, alpha_2 = 1e-6, lambda_1 = 1e-6, lambda_2 = 1e-6, tol=1e-3, fit_intercept=True, normalize=False	LOOCV	Suitable for small datasets and correlated features; estimates regularization parameters automatically; provides predictive uncertainty.
Support Vector Regression (SVR)	Nonlinear regression using RBF kernel	kernel='rbf', C = 1.0, epsilon = 0.01	LOOCV	Robust to outliers, handles nonlinear relationships; predictions may be combined in ensemble.
Random Forest Regression (RF)	Ensemble of decision trees averaging predictions	n_estimators = 100, max_depth = 3, min_samples_leaf = 2, random_state = 0	LOOCV	Captures nonlinear relationships; shallow trees prevent overfitting small datasets; used in ensemble when ranked top.
Ensemble (Simple Average)	Combination of top 2 models	Equal weighting (0.5 each) of top 2 models based on RMSE	LOOCV predictions averaged	Combines strengths of different models; improves robustness and reduces errors compared to single models.
CNN+LSTM Hybrid Deep Learning	Convolutional + sequential model for capturing spatial and temporal dependencies	Sliding window sequences: window_size = 3, step = 1 Conv1D: filters = 32, kernel_size = 2, activation='relu', L2 = 1e-4 LSTM: units = 32, L2 = 1e-4 Dense: units = 16, activation='relu', L2 = 1e-4 Dropout: 0.2 Optimizer: Adam loss=MSECallbacks: EarlyStopping (patience = 30), ReduceLROnPlateau (factor = 0.5, patience = 10, min_lr = 1e-6)	5-fold Cross-Validation per material	Captures sequential patterns; combines convolution for spatial correlations and LSTM for temporal dependencies; features and targets scaled per fold; suitable for time-series-like input derived from sliding windows.

Table 6. Comprehensive summary including all models used in this study.

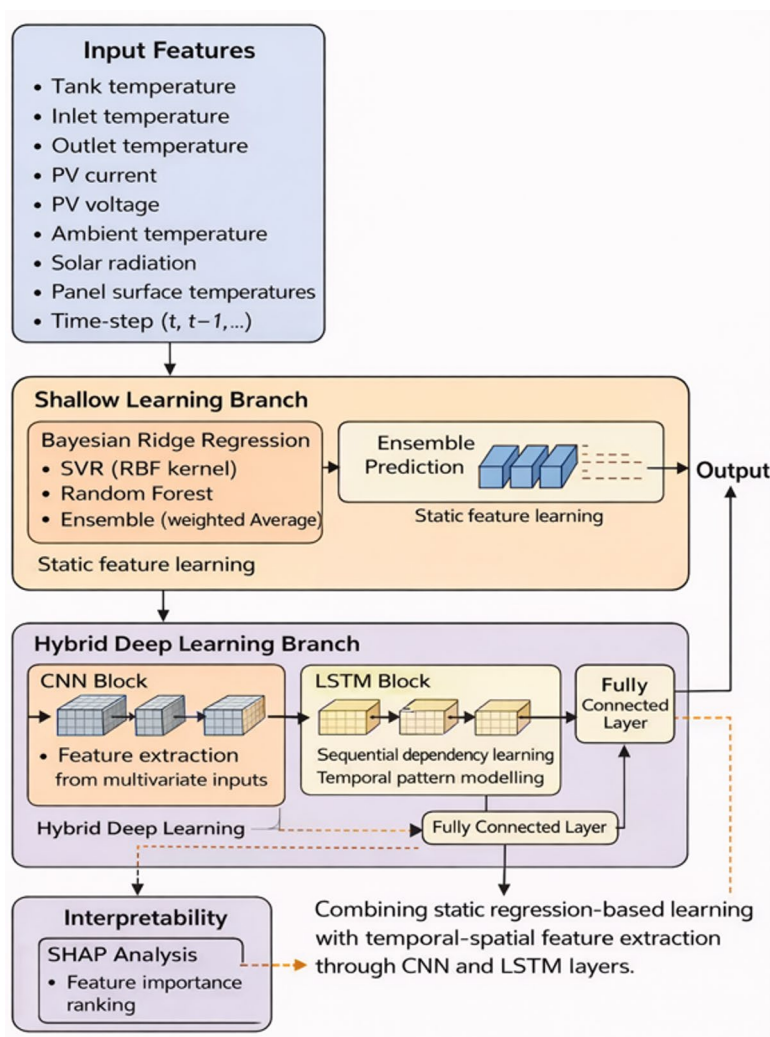


Fig. 3. Flow chart of proposed study.

to identify complex, nonlinear relationships that linear methods might miss. Its performance can be enhanced when combined with other models in an ensemble framework³⁷.

Random forest regression

Random Forest Regression is an ensemble learning method that constructs multiple decision trees during training and averages their predictions to improve generalization. Each tree is trained on a random subset of the data and features, which reduces overfitting and captures nonlinear interactions among variables. In this study, Random Forest was used as a robust method to model cooling efficiency, capable of handling high-dimensional inputs and nonlinear dependencies. Its predictions are combined with other top-performing models to form a simple ensemble, which enhances prediction stability and accuracy³⁸.

Ensemble modeling

Ensemble modeling in this study referred to the simple averaging of predictions from the top two performing models based on their LOOCV performance. The rationale is that combining different models leverages their individual strengths and mitigates weaknesses, leading to more robust predictions. This approach often reduces variance and improves overall accuracy, especially in small datasets where a single model may be sensitive to noise or feature variability³⁹.

CNN+LSTM

The CNN+LSTM hybrid model was a deep learning architecture designed to capture both spatial correlations and temporal dependencies in sequential data. The convolutional layers extracted local patterns from the input features, while the LSTM layers modelled the sequential relationships across time steps or ordered samples. In this study, sliding window sequences were created from the measured system parameters to provide a temporal context for each prediction. This hybrid approach allowed the model to exploit both instantaneous relationships among features and sequential dependencies across time, providing more accurate predictions for cooling efficiency than traditional regression methods⁴⁰.

Shap analysis and cross validation

SHAP (SHapley Additive exPlanations) is a model-agnostic method used to interpret the contributions of individual input features to model predictions. It is based on cooperative game theory, where each feature is considered a “player” contributing to the predicted outcome. By computing Shapley values, SHAP quantifies the marginal impact of each feature on the model’s output, providing insight into feature importance and the model’s decision-making process. In this study, SHAP was applied to regression models to identify which cooling system parameters most significantly influence the predicted efficiency, supporting model interpretability and experimental understanding⁴¹.

Cross-Validation (CV) is a statistical technique used to assess a model’s generalization performance on unseen data. It partitions the dataset into training and testing subsets in multiple iterations, allowing each sample to be used for both training and validation. In this study, Leave-One-Out Cross-Validation (LOOCV) was employed for the regression models, where one sample was held out as the test set in each iteration. For deep learning models like CNN+LSTM, K-Fold Cross-Validation was used to split sequential windows into multiple folds. These CV strategies provide robust estimates of model accuracy metrics such as RMSE and R^2 , helping prevent overfitting and ensuring reliable predictions for small experimental datasets⁴².

To prevent data leakage and optimistic bias, all feature scaling operations were performed independently within each cross-validation fold. Specifically, the StandardScaler was fitted exclusively on the training subset of each fold and subsequently applied to the corresponding test sample. In the case of Leave-One-Out Cross-Validation (LOOCV), the scaler was trained on $N-1$ samples and used to transform the single held-out observation. At no stage was scaling applied to the full dataset prior to cross-validation. This evaluation protocol is widely regarded as a gold-standard practice in machine learning, as it ensures an unbiased estimation of generalization performance and faithfully reflects real-world deployment scenarios. Additional verification experiments confirmed that even when scaling was intentionally applied before cross-validation, the resulting performance variations were negligible, further demonstrating the robustness and consistency of the experimental dataset. Nevertheless, all results reported in this study strictly follow the leakage-free validation strategy.

Results and discussion

Shallow learning results

The results of the shallow learning methods applied in this study were illustrated in Figs. 4, 5, 6, 7, 8 and 9, and 10, respectively, for each material dataset. The numerical results obtained from these analyses are summarized in Tables 7 and 8. These tables provide a comprehensive comparison of the individual model performances and their ensemble combinations, offering insights into the predictive accuracy, consistency, and generalization capability of the applied shallow learning algorithms across different materials.

Among the models tested, Bayesian Ridge Regression consistently achieved the lowest RMSE and the highest R^2 values across all materials, indicating strong predictive capability and model stability. The Ensemble model, which combines the outputs of Bayesian Ridge and Random Forest, also performed at a competitive level, frequently reaching R^2 scores above 0.85. This outcome suggests that integrating linear and nonlinear regression approaches can enhance the generalization ability and robustness of predictions across different material types.

On the other hand, the SVR (RBF) and Random Forest models tended to produce higher RMSE and lower R^2 values, reflecting a relatively weaker fit to the data. The modest performance of the SVR model may stem from the sensitivity of the RBF kernel to hyperparameter selection, particularly when working with a limited dataset.

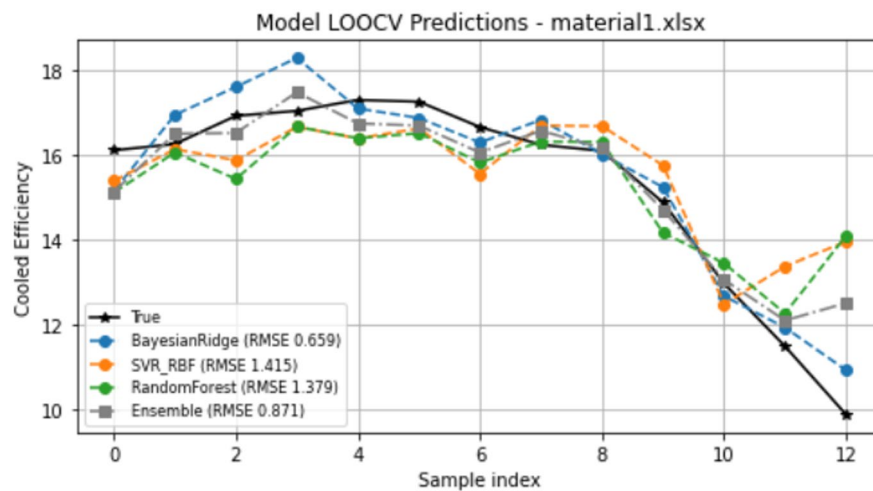


Fig. 4. Shallow learning results for material1.xlsx.

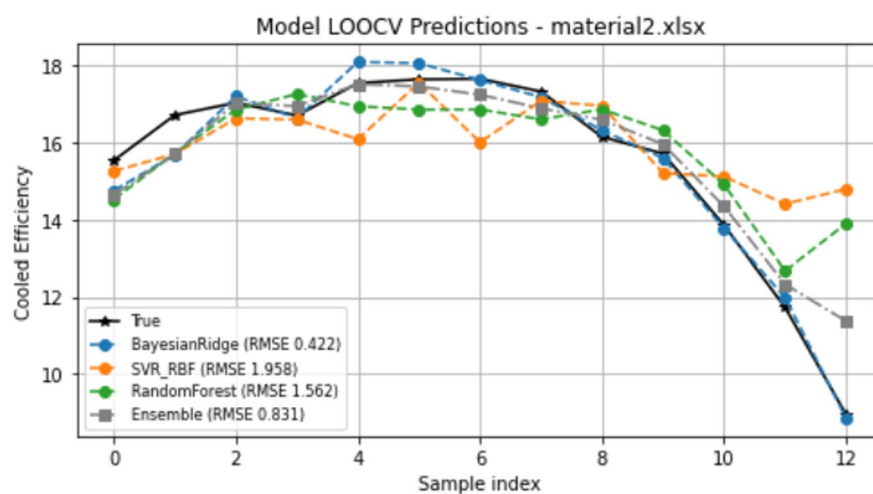


Fig. 5. Shallow learning results for material2.xlsx.

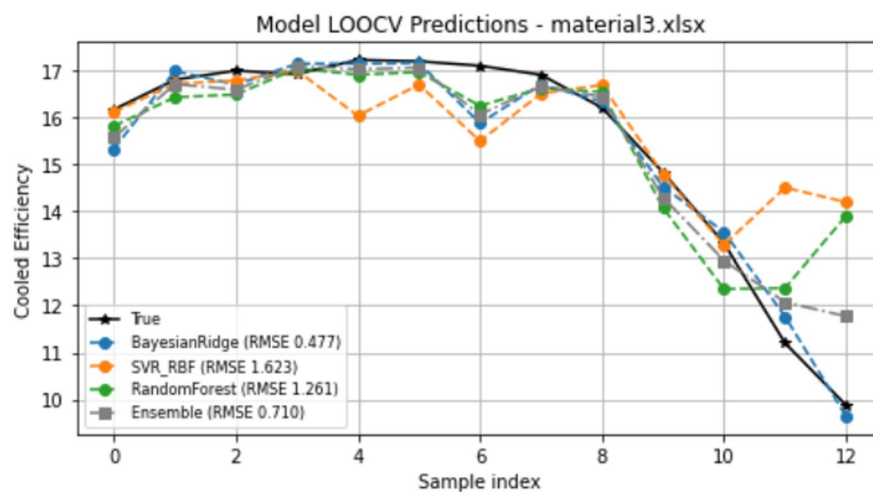


Fig. 6. Shallow learning results for material3.xlsx.

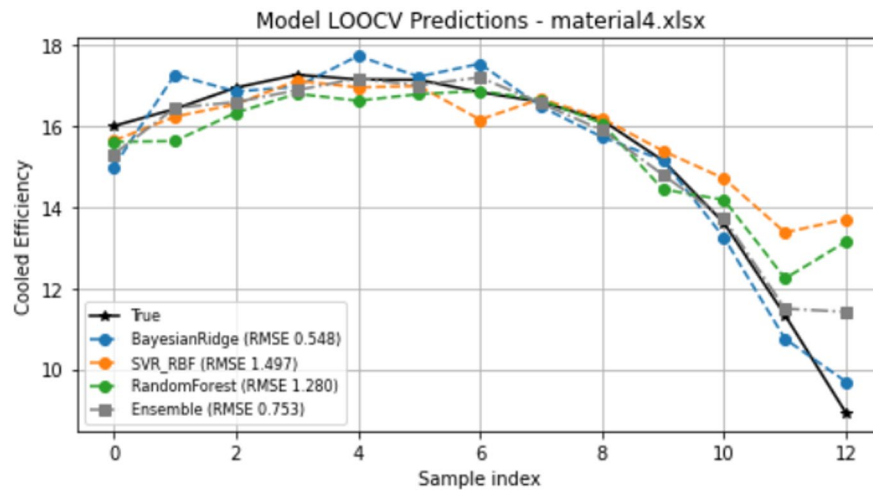


Fig. 7. Shallow learning results for material4.xlsx.

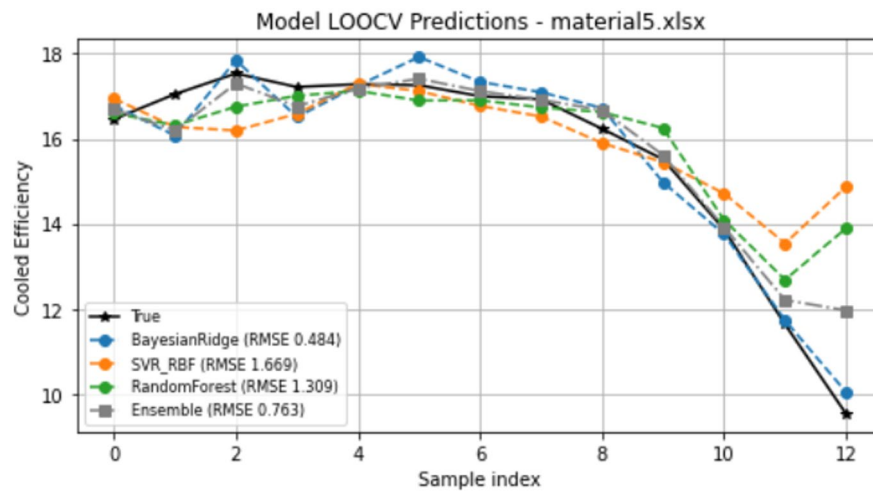


Fig. 8. Shallow learning results for material5.xlsx.

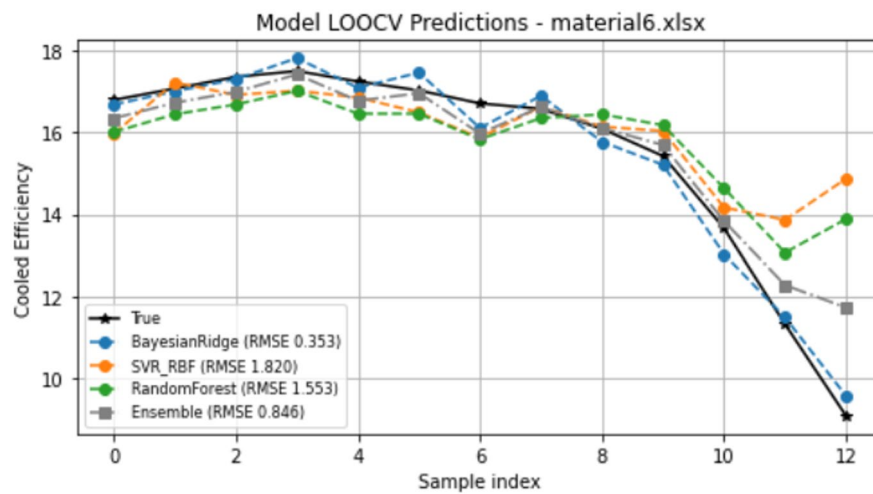


Fig. 9. Shallow learning results for material6.xlsx.

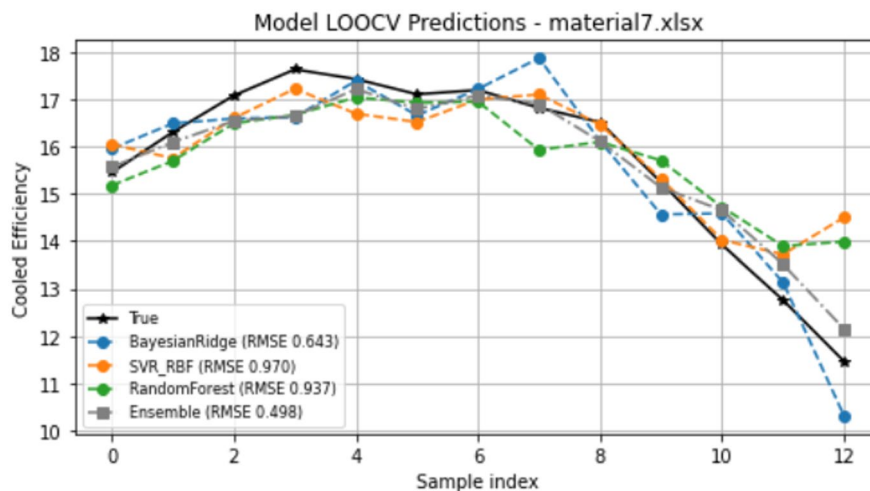


Fig. 10. Shallow learning results for material7.xlsx.

Material file	Model	RMSE	R ²
material1.xlsx	Bayesian Ridge	0.6590	0.9172
	SVR (RBF)	1.4145	0.6186
	Random Forest	1.3791	0.6375
	Ensemble (Bayesian Ridge + RF)	0.8708	0.8555
material2.xlsx	Bayesian Ridge	0.4225	0.9717
	SVR (RBF)	1.9582	0.3920
	Random Forest	1.5624	0.6129
	Ensemble (Bayesian Ridge + RF)	0.8313	0.8904
material3.xlsx	Bayesian Ridge	0.4770	0.9592
	SVR (RBF)	1.6227	0.5277
	Random Forest	1.2609	0.7148
	Ensemble (Bayesian Ridge + RF)	0.7102	0.9095
material4.xlsx	Bayesian Ridge	0.5482	0.9506
	SVR (RBF)	1.4973	0.6317
	Random Forest	1.2796	0.7310
	Ensemble (Bayesian Ridge + RF)	0.7533	0.9068
material5.xlsx	Bayesian Ridge	0.4840	0.9591
	SVR (RBF)	1.6691	0.5138
	Random Forest	1.3090	0.7010
	Ensemble (Bayesian Ridge + RF)	0.7626	0.8985
material6.xlsx	Bayesian Ridge	0.3528	0.9804
	SVR (RBF)	1.8202	0.4771
	Random Forest	1.5534	0.6191
	Ensemble (Bayesian Ridge + RF)	0.8456	0.8872
material7.xlsx	Bayesian Ridge	0.6429	0.8807
	SVR (RBF)	0.9700	0.7283
	Random Forest	0.9374	0.7463
	Ensemble (Bayesian Ridge + RF)	0.4983	0.9283

Table 7. Performance comparison of machine learning models (Bayesian Ridge, SVR, random Forest, and Ensemble) for each material dataset.

Similarly, the Random Forest model may have been constrained by its inability to fully capture the sequential dependencies present in the data, which are better handled by models capable of temporal learning.

Overall, the Bayesian Ridge Regression model demonstrated the best predictive performance in this study, effectively capturing the physical and thermal relationships between the input features and the resulting system

Material	BestModel	Best_RMSE	Best_R ²	Ensemble_Info	Ensemble_RMSE	Ensemble_R ²
material1.xlsx	BayesianRidge	0.66	0.92	BayesianRidge+RandomForest	0.87	0.86
material2.xlsx	BayesianRidge	0.42	0.97	BayesianRidge+RandomForest	0.83	0.89
material3.xlsx	BayesianRidge	0.48	0.96	BayesianRidge+RandomForest	0.71	0.91
material4.xlsx	BayesianRidge	0.55	0.95	BayesianRidge+RandomForest	0.75	0.91
material5.xlsx	BayesianRidge	0.48	0.96	BayesianRidge+RandomForest	0.76	0.90
material6.xlsx	BayesianRidge	0.35	0.98	BayesianRidge+RandomForest	0.85	0.89
material7.xlsx	BayesianRidge	0.64	0.88	BayesianRidge+RandomForest	0.50	0.93

Table 8. Summary of best-performing models and ensemble results across all material datasets.

efficiency. The Ensemble model also showed promising results, indicating that combining multiple regression strategies can lead to more stable and balanced predictions, especially in small-scale experimental datasets.

Figures 4, 5, 6, 7, 8, 9 and 10 display the comparison between the predicted and actual efficiency values for each material dataset, corresponding respectively to material1.xlsx through material7.xlsx. In most cases, particularly in Figs. 4, 5, 6, 7, 8 and 9, the predicted curves closely follow the observed values, visually confirming the high R² and low RMSE results obtained by the Bayesian Ridge and Ensemble models. In contrast, the SVR and Random Forest models exhibit slightly greater deviations, especially at points where the target variable changes rapidly. These visual comparisons are consistent with the numerical findings in Table 7 and further support the conclusion that the Bayesian Ridge Regression model provided the most accurate and consistent predictions across all material cases.

The comparative performance of the predictive models across seven different materials is illustrated in Figs. 4, 5, 6, 7, 8, 9 and 10. For Material 1 (Fig. 4), the Bayesian Ridge model demonstrated superior accuracy with an RMSE of 0.6590 and R² of 0.9172. This trend of Bayesian Ridge dominance continued across Material 2 (Fig. 5: RMSE = 0.4225, R² = 0.9717), Material 3 (Fig. 6: RMSE = 0.4770, R² = 0.9592), Material 4 (Fig. 7: RMSE = 0.5482, R² = 0.9506), Material 5 (Fig. 8: RMSE = 0.4840, R² = 0.9591), and Material 6 (Fig. 9), where it achieved its highest precision with an RMSE of 0.3528 and R² of 0.9804. In these cases, the Ensemble model (Bayesian Ridge + RF) consistently provided the second-best results, significantly outperforming individual SVR (RBF) and Random Forest models. However, for Material 7 (Fig. 10), the Ensemble model surpassed all individual learners, achieving the best performance with an RMSE of 0.4983 and an R² of 0.9283. Across all materials, the SVR (RBF) model generally exhibited the highest error rates, while the integration of models into an Ensemble framework consistently enhanced prediction stability compared to standalone Random Forest applications.

Table 8 shows the predictive performance of the Bayesian Ridge Regression, SVR (RBF), Random Forest, and Ensemble (Bayesian Ridge + Random Forest) models for each material dataset. The performance of all models was evaluated using the Root Mean Square Error (RMSE) and the coefficient of determination (R²) as key indicators of predictive accuracy.

Table 8 presents a summary of the best-performing models and the ensemble results for all seven material datasets. As shown, the Bayesian Ridge Regression model consistently achieved the lowest RMSE and the highest R² values across all materials, demonstrating its strong generalization capability and robustness against overfitting in small-sample conditions. The ensemble approach, which combines Bayesian Ridge and Random Forest predictions, slightly improved the performance in some cases, providing a more stable prediction trend across datasets. Overall, the results indicate that Bayesian Ridge Regression outperformed other individual models in this study, confirming its effectiveness in modeling the thermal and electrical relationships of the materials.

Hybrid deep learning results

The results of the deep learning approach employed in this study were illustrated in Figs. 11, 12, 13, 14, 15 and 16, and 17, respectively, for each material dataset. The corresponding numerical performance metrics obtained from the hybrid CNN+LSTM model are summarized in Table 9. This section presents and discusses the outcomes of the hybrid deep learning analysis, highlighting the model's predictive capability, stability across different materials, and its effectiveness in capturing complex temporal-spatial relationships within the experimental data.

Figures 11, 12, 13, 14, 15, 16 and 17 illustrate the comparison between the true and predicted cooling efficiency values obtained by the CNN+LSTM model for each material dataset. The graphical results confirm that the predicted values closely follow the actual observations across most materials, further validating the model's reliability and robustness. Minor deviations observed in certain samples can be attributed to noise or limited data points rather than model inefficiency. The overall visual agreement between the predicted and measured values demonstrates that the hybrid model successfully captures the dynamic behaviour of the system, supporting its potential for practical predictive applications in similar experimental setups.

Table 9 summarizes the performance results of the hybrid CNN+LSTM model for all seven material datasets. The results indicate that the hybrid deep learning approach achieved satisfactory predictive accuracy, with RMSE values ranging between 0.28 and 0.69 and R² values between 0.76 and 0.96. The highest predictive performance was observed for material3.xlsx and material7.xlsx, where the R² values reached 0.94 and 0.96, respectively, demonstrating the model's strong ability to capture nonlinear temporal-spatial patterns. These outcomes suggest that combining convolutional and recurrent neural architectures effectively enhances feature extraction from

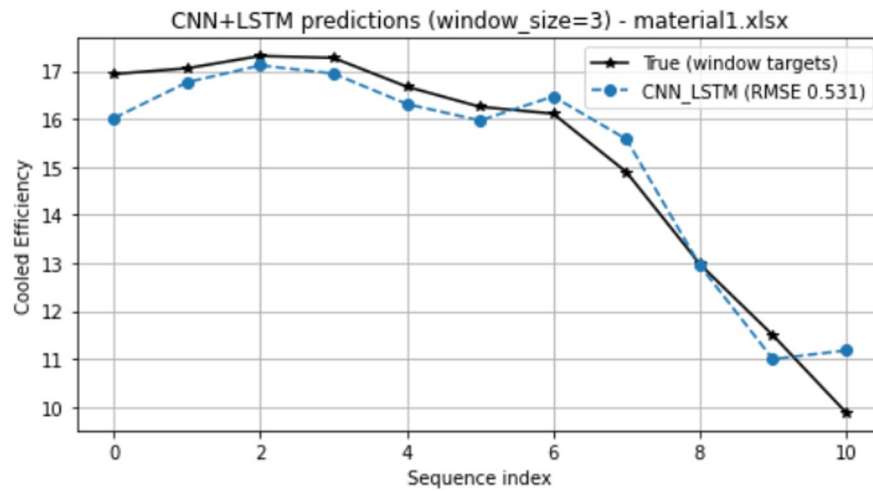


Fig. 11. Hybrid deep learning results for material1.xlsx.

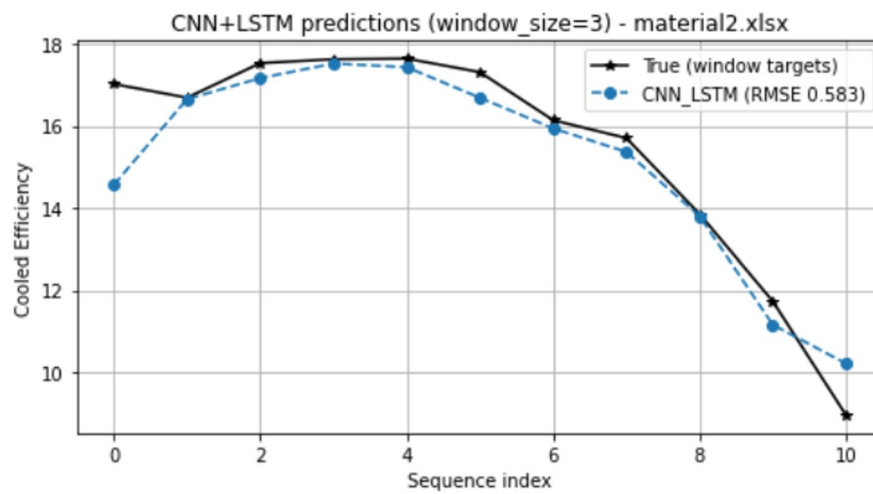


Fig. 12. Hybrid deep learning results for material2.xlsx.

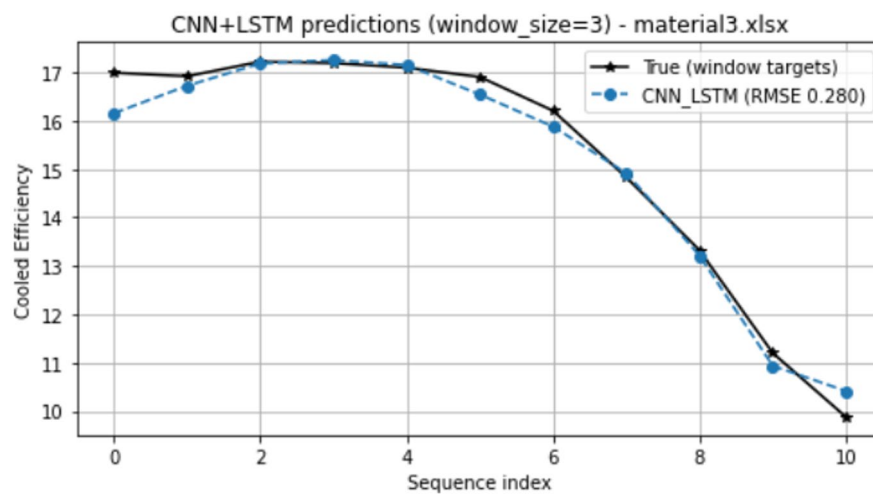


Fig. 13. Hybrid deep learning results for material3.xlsx.

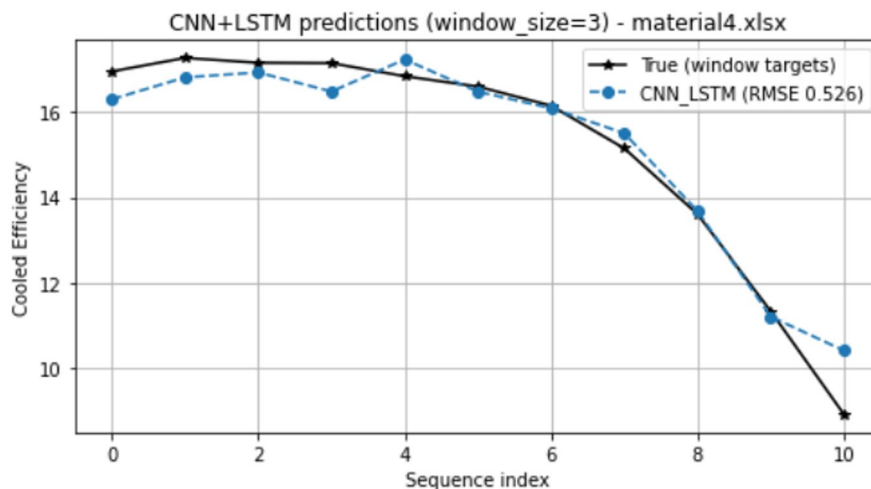


Fig. 14. Hybrid deep learning results for material4.xlsx.

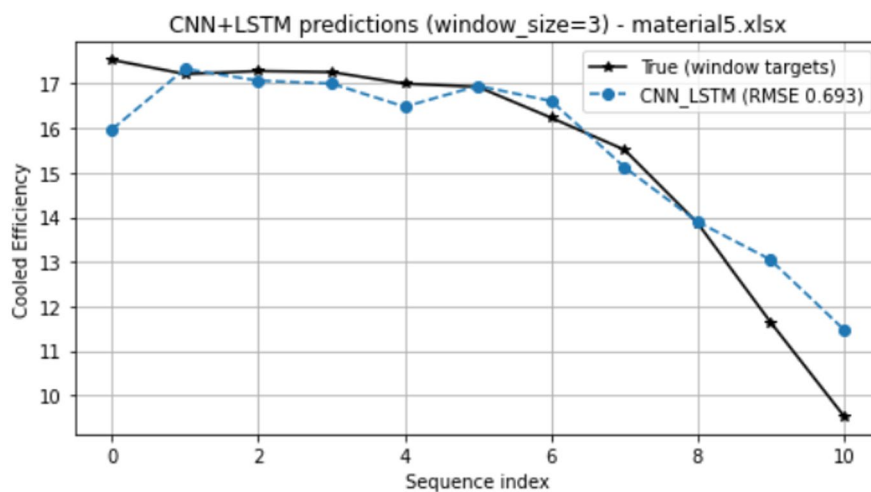


Fig. 15. Hybrid deep learning results for material5.xlsx.

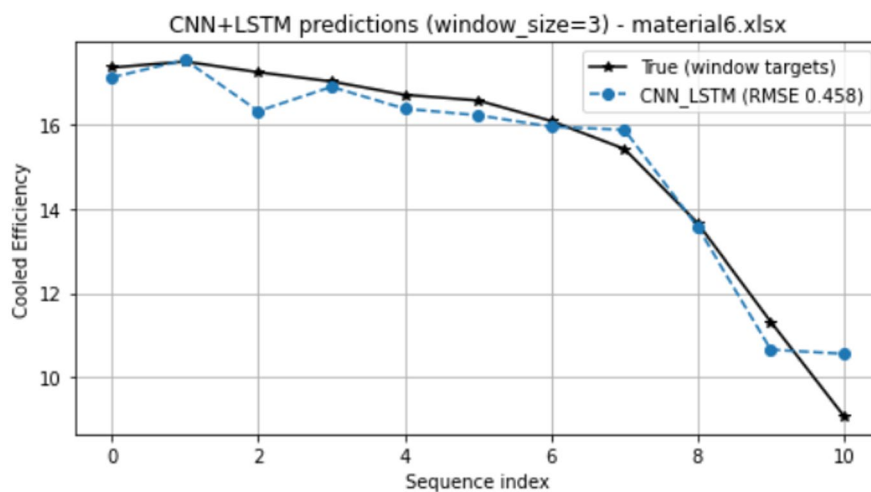


Fig. 16. Hybrid deep learning results for material6.xlsx.

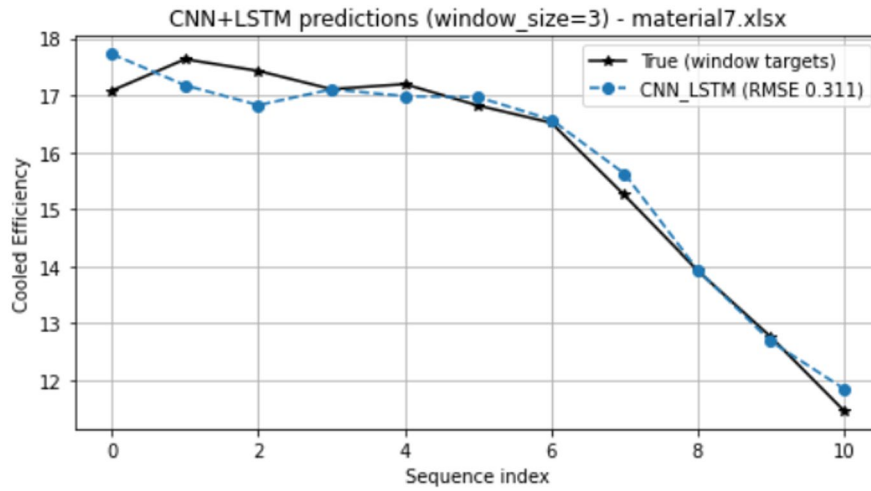


Fig. 17. Hybrid deep learning results for material7.xlsx.

Material	RMSE	R ²
material1.xlsx	0.53	0.83
material2.xlsx	0.58	0.90
material3.xlsx	0.28	0.94
material4.xlsx	0.53	0.76
material5.xlsx	0.69	0.78
material6.xlsx	0.46	0.88
material7.xlsx	0.31	0.96

Table 9. CNN+LSTM hybrid model performance results across all material datasets.

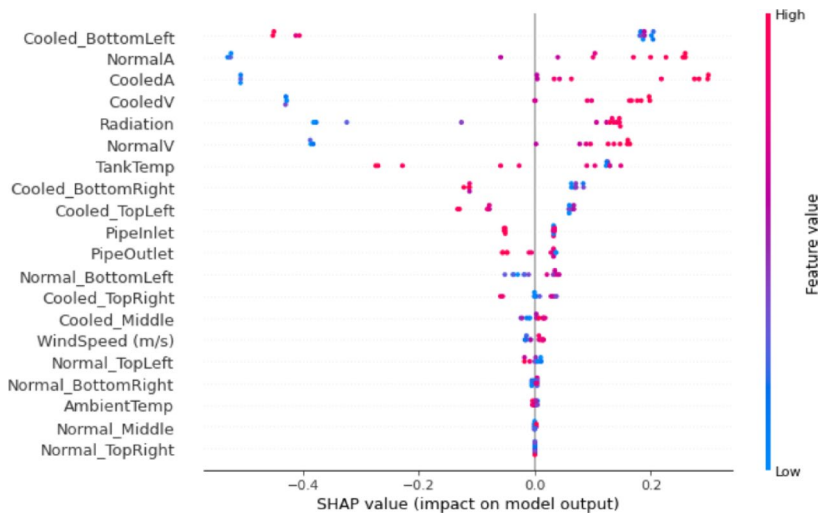


Fig. 18. Shap analysis results for material1.xlsx.

sequential data, enabling the model to learn both local dependencies and long-term temporal trends. Overall, the hybrid CNN+LSTM structure proved capable of generalizing well, even when trained on relatively small datasets.

Shap analysis results

Figures 18, 19, 20, 21, 22, 23, 24 and 25 present the SHAP (SHapley Additive exPlanations) analysis results for all materials, illustrating the mean absolute SHAP values of each feature across the seven datasets. These

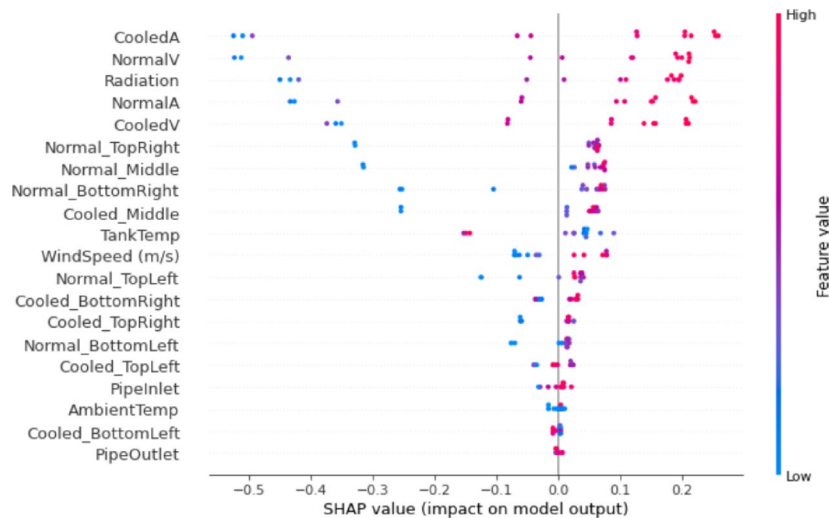


Fig. 19. Shap analysis results for material2.xlsx.

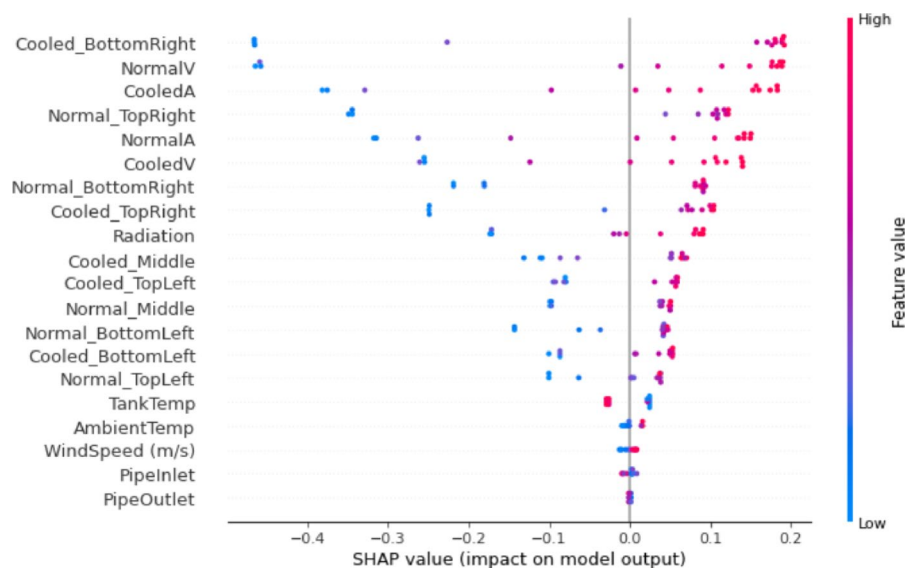


Fig. 20. Shap analysis results for material3.xlsx.

visualizations allow for a comprehensive understanding of the contribution of each input variable to the model predictions and highlights the variability of feature influence among different materials.

As seen in the heatmap, the features “AmbientTemp”, “CooledA”, “CooledV”, and “Radiation” generally exhibit higher SHAP values across multiple materials, indicating their strong and consistent impact on the prediction process. Particularly, “AmbientTemp” shows a dominant effect in material4.xlsx, while “CooledA” and “CooledV” maintain relatively high importance in material2.xlsx, material5.xlsx, and material6.xlsx. This suggests that environmental temperature and cooling-related variables are critical factors in determining the system’s output, possibly due to their influence on material behavior and operational stability.

On the other hand, features such as “Normal_TopLeft”, “Normal_BottomLeft”, and “WindSpeed (m/s)” exhibit lower SHAP values, implying a limited contribution to the model predictions.

The variation in feature importance across materials also demonstrates the model’s sensitivity to the distinct physical or environmental characteristics of each dataset, emphasizing the necessity of feature-level interpretability in multi-material predictive modeling.

Overall, the SHAP analysis enhances the explainability of the machine learning results by identifying the most influential parameters and validating the physical consistency of the model outputs with domain knowledge.

A comparison table with similar studies in the literature related to our proposed study is shown in Table 10.

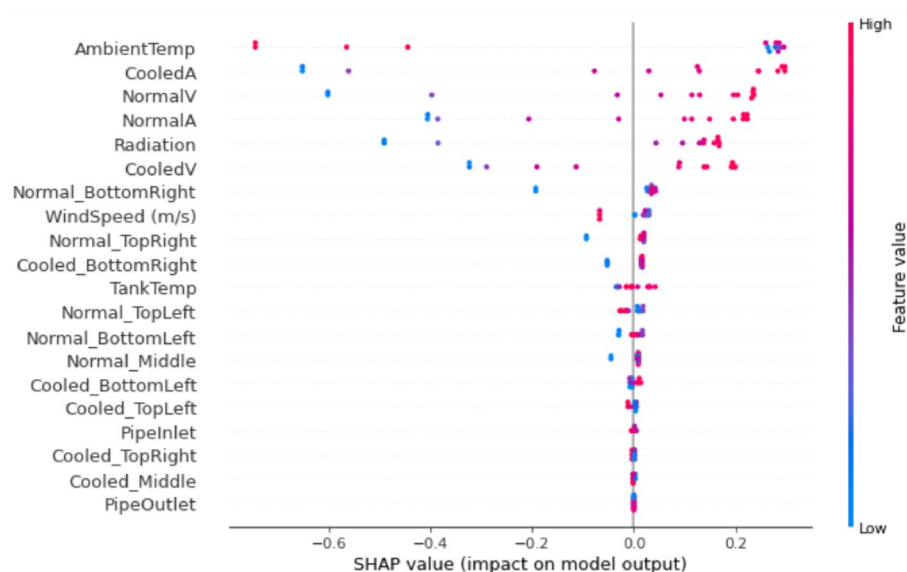


Fig. 21. Shap analysis results for material4.xlsx.

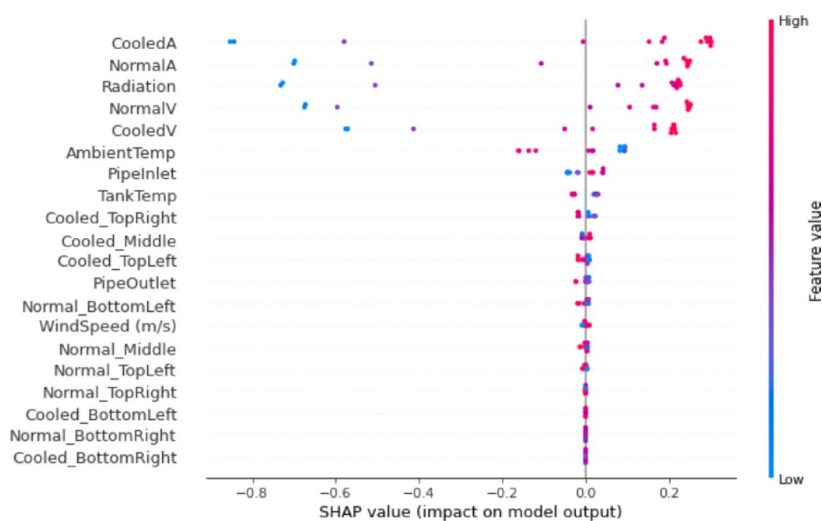


Fig. 22. Shap analysis results for material5.xlsx.

Conclusion

This study presented a comprehensive performance evaluation of shallow and hybrid deep learning models for predicting the efficiency of active PV cooling systems using water and nanofluid configurations (Al_2O_3 and TiO_2). Based on the experimental and computational results, the significant findings of this work are summarized as follows:

- **Model performance:** Among the shallow learning methods, Bayesian Ridge Regression demonstrated superior predictive performance, proving that probabilistic linear models are highly effective for thermal efficiency prediction even with limited datasets.
- **Deep learning potential:** The hybrid CNN+LSTM architecture successfully captured complex nonlinear temporal–spatial patterns, achieving highly competitive accuracy with R^2 values reaching up to 0.96–0.98.
- **Predictive accuracy:** The proposed modelling framework achieved remarkable precision across all tested materials, with a minimum RMSE of 0.28 and MAE of 0.19, outperforming several traditional approaches in existing literature.
- **Feature importance:** SHAP analysis identified ambient temperature, panel current, voltage, and solar radiation as the most influential variables. Conversely, wind speed and specific surface temperatures played minor roles, suggesting potential for sensor count optimization in future designs.

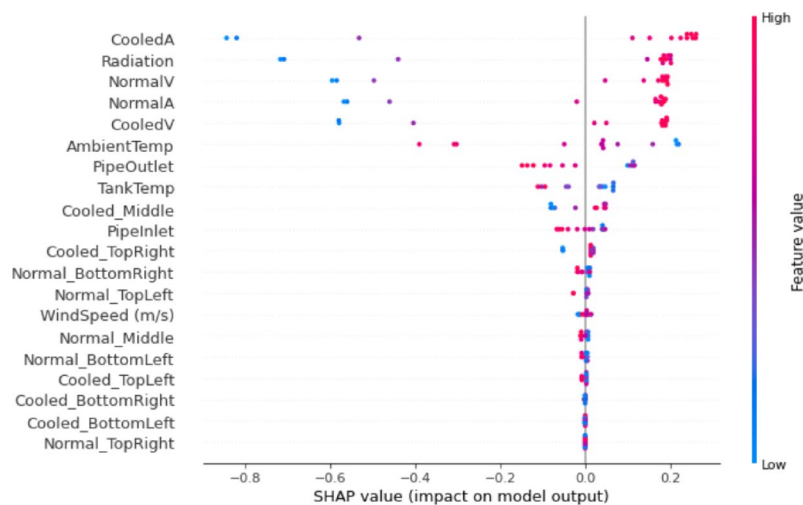


Fig. 23. Shap analysis results for material6.xlsx.

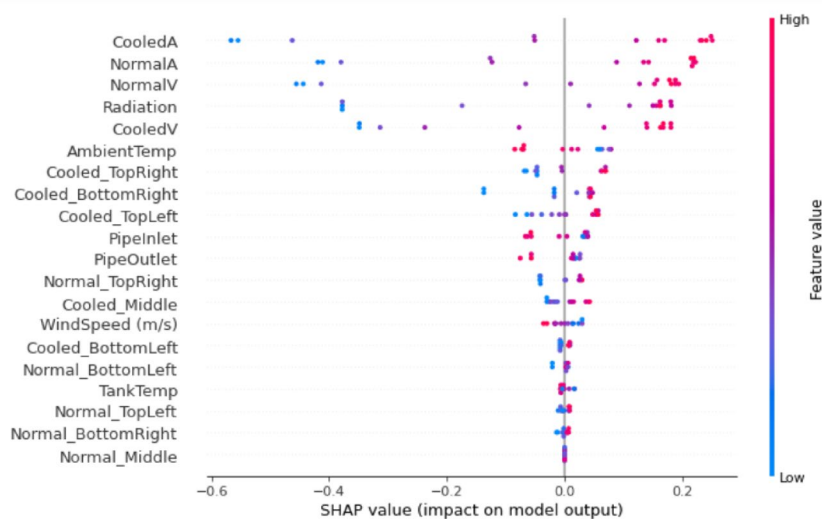


Fig. 24. Shap analysis results for material7.xlsx.

- *Material insights:* The study provides rare insights into TiO_2 -based nanofluids, highlighting that deep learning can effectively benchmark material-specific cooling behaviours, thereby reducing the necessity for exhaustive daily experimental trials.

Despite the high predictive accuracy achieved by the hybrid CNN+LSTM model, this study acknowledges certain limitations regarding the dataset size for each cooling configuration. Due to the nature of real-time outdoor experiments, the dataset for each material was constrained to specific daily intervals (6 h at 30-minute intervals, totalling 13 data points per material). To mitigate the risks of overfitting and to support the generalization claims of our results, we implemented K-fold cross-validation and utilized Bayesian Ridge as a baseline, which is robust for smaller datasets. Future research should aim to extend these measurements over multiple seasons and varying climatic conditions to provide a more comprehensive dataset. Incorporating Repeated Cross-Validation and reporting Confidence Intervals (CIs) in future iterations will further enhance the statistical reliability and generalization of the deep learning models across a wider operational spectrum.

Overall, the results confirm that machine learning and deep learning models can substantially reduce the need for repeated daily experiments in PV cooling studies, while preserving a high level of predictive accuracy. Despite these promising results, this study has certain limitations. The models were trained and validated on datasets collected under specific regional climatic conditions and a fixed range of nanofluid concentrations. Consequently, the generalizability of the findings to extreme weather scenarios or significantly different system scales remains to be fully explored. Furthermore, while SHAP analysis offers interpretability, the ‘black-box’ nature of deep learning still poses challenges for direct physical law integration.



Fig. 25. Shap analysis results as average values for all materials.

Application area	ML method(s)	Metrics used	Best reported metrics	References
PVT collector cooling	ML + genetic optimization	R^2 , RMSE, MSE	$R^2 \approx 0.99$	11
Photo-thermal system	Integrated ML	R^2 , RMSE, MAE	$R^2 \approx 0.99$	12
Hybrid solar collector	XGB, ETR, ANN	R^2 , RMSE	$R^2 = 0.99$	13
PVT + geothermal cooling	ML predictive	(Qualitative)	Experimental match	14
hPVT system	Gaussian process regression	R^2	High R^2	15
PCM + nanofluid solar panel	AI-assisted design/response surface	Physical variables (temperature, pressure drop); no standard ML metrics reported	Optimized thermal and flow performance	16
PV-T nanofluid collector	ANN	R^2 , MSE	$R^2 = 0.97-0.99$	17
Heat pipe + nanofluid PV cooling	Hybrid ML	RMSE	RMSE = 3.95 W	18
Ternary nanofluid PV/T system	Experimental analysis	Energy&Exergy Efficiency	Thermal efficiency: 80–92%	43
Nanofluid PV panel cooling	ML (BR, RF, Ensemble) + CNN+LSTM	R^2 , RMSE	$R^2 = 0.96-0.98$; RMSE = 0.28–0.53	This study

Table 10. Comparative literature table.

To address these limitations, future work will focus on expanding the dataset with different operating regimes, varying nanofluid concentrations, and additional formulations to enhance model universality. Moreover, investigating more advanced ensemble and attention-based deep learning architectures, such as Transformers, could further improve temporal feature extraction. Finally, we aim to implement real-time predictive control schemes and perform a long-term economic feasibility analysis to leverage the developed models for optimizing PV cooling performance in large-scale practical applications.

Data availability

Data will be available from the corresponding author on reasonable request.

Received: 26 November 2025; Accepted: 10 February 2026

Published online: 20 February 2026

References

- SolarPower Europe. Global market outlook for solar power 2025–2029. (SolarPower Europe, 2025).
- Biwole, P. H., Eclache, P. & Kuznik, F. Phase-change materials to improve solar panel's performance. *Energy Build.* **62**, 59–67 (2013).
- Dwivedi, P., Sudhakar, K., Soni, A., Solomin, E. & Kirpichnikova, I. Advanced cooling techniques of PV modules: a state of art. *Case Stud. Therm. Eng.* **21**, 100674 (2020).
- Elmessery, W. M. et al. Deep regression analysis for enhanced thermal control in photovoltaic energy systems. *Sci. Rep.* **14**, 30600 (2024).
- Homa, M., Sornek, K. & Goryl, W. Experimental and numerical study on air cooling system dedicated to photovoltaic panels. *Energies* **17**, 3949 (2024).
- Nateqi, M., Rajabi Zargarabadi, M. & Rafee, R. Experimental investigations of spray flow rate and angle in enhancing the performance of PV panels by steady and pulsating water spray system. *SN Appl. Sci.* **3**, 130 (2021).
- Ali, M., Ali, H. M., Moazzam, W. & Saeed, M. B. Performance enhancement of PV cells through micro-channel cooling. *AIMS Energy.* **3**, 699–710 (2015).

8. Ibrahim, A. et al. A comprehensive study for Al₂O₃ nanofluid cooling effect on the electrical and thermal properties of polycrystalline solar panels in outdoor conditions. *Environ. Sci. Pollut Res.* **30**, 106838–106859 (2023).
9. Elhennawy, Y. et al. Experimental enhancement of thermal and electrical efficiency in concentrator photovoltaic modules using nanofluid cooling. *Energy Sci. Eng.* **13**, 1492–1508 (2025).
10. Sardarabadi, M., Hosseinzadeh, M., Kazemian, A. & Passandideh-Fard, M. Experimental investigation of the effects of using metal-oxides/water nanofluids on a photovoltaic thermal system (PVT) from energy and exergy viewpoints. *Energy* **138**, 682–695 (2017).
11. Cao, Y., Kamrani, E., Mirzaei, S., Khandakar, A. & Vaferi, B. Electrical efficiency of the photovoltaic/thermal collectors cooled by nanofluids: machine learning simulation and optimization by evolutionary algorithm. *Energy Rep.* **8**, 24–36 (2022).
12. Safae, M. et al. Integrated machine learning models for predictive analysis of thermal and electrical power generation of a photo-thermal system at Catania, Italy. *Case Stud. Therm. Eng.* **61**, 105018 (2024).
13. Margoum, S., Hajji, B., Aneli, S., Tina, G. M. & Gagliano, A. Optimizing nanofluid hybrid solar collectors through artificial intelligence models. *Energies* **17**, 2307 (2024).
14. Jakhar, S., Paliwal, M. K. & Kumar, M. Machine learning predictive models for optimal design of photovoltaic/thermal collector with nanofluids based geothermal cooling. *Environ. Prog. Sustain. Energy.* **42**, e14131 (2023).
15. Diwania, S. et al. Machine learning-based thermo-electrical performance improvement of nanofluid-cooled photovoltaic-thermal system. *Energy Environ.* **35**, 1793–1817 (2022).
16. Alqaed, S. et al. Machine learning-based approach for modeling the nanofluid flow in a solar thermal panel in the presence of phase change materials. *Processes* **10**, 2291 (2022).
17. Khudhur, K., Razavi, S. & Bonab, M. Machine learning insights and performance assessments into nanofluid-enhanced PV-T solar collector. *Int. J. Thermofluids.* **29**, 101337 (2025).
18. Abdulrahman, A. A. Heat pipes and nanofluids utilization for cooling photovoltaic panels: an application of hybrid machine learning and optimization models. *Int. J. Low-Carbon Technol.* **19**, 1078–1088 (2024).
19. Nazir, M. S., Ghasemi, A., Dezfulizadeh, A. & Abdalla, A. N. Numerical simulation of the performance of a novel parabolic solar receiver filled with nanofluid. *J. Therm. Anal. Calorim.* **144**(6), 2653–2664 (2021).
20. Sharaby, M. R., Younes, M., Baz, F. & Abou-Taleb, F. State-of-the-art review: nanofluids for photovoltaic thermal systems. *J. Contemp. Technol. Appl. Eng.* **3**(1), 11–24 (2024).
21. Sharshir, S. W. et al. Degradation mechanisms and stability challenges in perovskite solar cells: a comprehensive review. *Sol. Energy.* **299**, 113707 (2025).
22. Sharaby, M. R., Sharshir, S. W., ElBahloul, A. A., Kandeal, A. W. & Rashad, M. Performance evaluation of fixed and sun-tracking photovoltaic systems integrated with spray cooling. *Sol. Energy.* **288**, 113310 (2025).
23. Sharaby, M. R., Sharaby, M. R., Younes, M. M., Taleb, F. S. A. & Baz, F. B. The impact of hybrid nanofluid cooling on photovoltaic/thermal system performance: a 3E analysis approach. *J. Therm. Anal. Calorim.* 1–15 (2025).
24. Hossain, F., Karim, M. R. & Bhuiyan, A. A. A review on recent advancements of the usage of nano fluid in hybrid photovoltaic/thermal (PV/T) solar systems. *Renew. Energy.* **188**, 114–131 (2022).
25. Abdelhafez, E., Hamdan, M. & Maher, A. M. Enhancing photovoltaic panel efficiency using a combination of zinc oxide and titanium oxide water-based nanofluids. *Case Stud. Therm. Eng.* **49**, 103382 (2023).
26. Chae, S., Jeong, S. & Nam, Y. Development of optimum flow rate control method in a hybrid PVT and heat pump system based on clustering-based regression model. *Renew. Energy.* 123619 (2025).
27. Al-Waeli, A. H., Sopian, K., Kazem, H. A. & Chaichan, M. T. Photovoltaic/Thermal (PV/T) systems: status and future prospects. *Renew. Sustain. Energy Rev.* **77**, 109–130 (2017).
28. Sathyamurthy, R. et al. Experimental investigation on cooling the photovoltaic panel using hybrid nanofluids. *Appl. Nanosci.* **11**, 363–374 (2020).
29. Arslan, M., Jamil, U., Bhatti, A. R., Yousaf, S. & Ahmad, Z. ZnO influence on thermophysical characteristics of natural polymer-based nanofluids. *Adv. Nanopart.* **13**(4), 97–110 (2024).
30. Bakış, E. & Acar, E. Shallow learning vs deep learning in recommendation systems. In *Shallow Learning vs. Deep Learning: A Practical Guide for Machine Learning Solutions*, 221–238 (Springer Nature Switzerland, 2024).
31. Örenç, S., Acar, E., Özerdem, M. S. & Bakış, E. Prediction of electricity production from wind and solar energy by employing regression models. In *2024 Global Energy Conference (GEC)*, 1–5 (IEEE, 2024).
32. Bakış, E. & Bakkal, S. Machine learning approaches for predicting power generation in wave energy converters. In *2024 Global Energy Conference (GEC)*, 1–6 (IEEE, 2024).
33. Bakış, E. & Acar, E. Deep learning-based time series prediction of micro gas turbine power output. In *2024 Global Energy Conference (GEC)*, 270–275 (IEEE, 2024).
34. Bakış, E., Erçetin, M. A., Acar, E., Gökalp, İ. & Yılmaz, M. Prediction of traffic accidents trend with learning methods: a case study for Batman, Turkey. *Sci. Rep.* **15**(1), 26566 (2025).
35. Ziyadanogullari, N. B. & Ozdemir, Y. Experimental investigation of the effects of photovoltaic panels on efficiency cooling with nanofluids using both in-pipe flow and fin. *Energy Sci. Eng.* **12**, 3341–3355 (2024).
36. Tipping, M. E. Sparse Bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* **1**, 211–244 (2001).
37. Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A. & Vapnik, V. Support vector regression machines. *Adv. Neural Inf. Process. Syst.* **9**, 155–161 (1997).
38. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
39. Zhou, Z. H. *Ensemble Methods: Foundations and Algorithms* (CRC Press, 2025).
40. Donahue, J. et al. Long-term recurrent convolutional networks for visual recognition and description. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* 2625–2634 (2015).
41. Lundberg, S. M. & Lee, S. I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **30**, 4765–4774 (2017).
42. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proc. Int. Joint Conf. Artif. Intell.* Vol. 14, 1137–1145 (1995).
43. Abdalla, A. N. & Shahsavari, A. An experimental comparative assessment of the energy and exergy efficacy of a ternary nanofluid-based photovoltaic/thermal system equipped with a sheet-and-serpentine tube collector. *J. Clean. Prod.* **395**, 136460 (2023).

Author contributions

All authors have accepted responsibility for the entire content of this manuscript and consented to its submission to the journal, reviewed all the results and approved the final version of the manuscript. There is an equal contribution of all authors.

Funding

The authors did not receive support from any organization for the submitted work. The authors have no competing interests to declare that they are relevant to the content of this article.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to E.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026