International Journal of Low-Carbon Technologies, 2026, **21**, 1–25, ctag108
https://doi.org/10.1093/ijlct/ctag108
Published: 1 September 2026
Original Article

OXFORD

Comprehensive techno-economic analysis of phase change material-stabilized building-integrated photovoltaic thermal façades in residential buildings: a comparative machine learning approach for hot-arid climates toward less carbon emission

Khaled Almazam[1], Eman Nasrallah[2], Faizah Mohammed Bashir[3], Abubakar Sadiq Mahmoud[4], Yakubu Aminu Dodo[1], Yohannes Mehari Andiye[5,*]

[1]Architectural Engineering Department, College of Engineering, Najran University, Najran 66426, Kingdom of Saudi Arabia
[2]Department of Interior Design, College of Design and Applied Art, Taif University, Taif 26571, Kingdom of Saudi Arabia
[3]Department of Decoration and Interior Design Engineering, College of Engineering, University of Ha'il, Ha'il 55427, Kingdom of Saudi Arabia
[4]Department of Facilities and Construction Project Management, International College of Engineering and Management, Oman, Affiliated with the University of Lancashire, UK, Muscat 111, Oman
[5]Faculty of Hydraulic and Water Resources Engineering, Arba Minch University, Ethiopia
*Corresponding author. Faculty of Hydraulic and Water Resources Engineering, Arba Minch University, Ethiopia. E-mail: yohannes.mehari@amu.edu.et.

Abstract

Building-integrated photovoltaic-thermal (BIPVT) façades can help hot-arid cities use limited envelope area for both electricity and domestic hot water production, but their assessment is often weakened by uncertain climate inputs and weak operational forecasting. To address these gaps, this study developed a simulation-based framework for a phase change material-stabilized BIPVT façade in Tabuk, Saudi Arabia. The framework combines satellite/reanalysis climate inputs, station-validated temperature and wind data, NASA Prediction of Worldwide Energy Resources–Copernicus Atmosphere Monitoring Service irradiance screening, clear-sky index checks, elastic net regression and Extreme Gradient Boosting (XGBoost) day-ahead forecasting, layer-resolved thermoelectrical modeling, EnergyPlus cross-validation, and life-cycle economic analysis. XGBoost gave the best irradiance forecast against the screened satellite irradiance series, with root mean squared error of 41.2 W·m^{-2} and daylight mean absolute percentage error of 8.6%. Increasing water flow from 0.02 to 0.10 kg·s^{-1} lowered photovoltaic temperature from ~52°C to 33°C and raised electrical efficiency from ~9% to nearly 16%. EnergyPlus deviations remained within ~5%, and east/west façades reached payback in 9–10 years.

Keywords BIPVT façade, phase change material, climate-data screening, solar forecasting, XGBoost, techno-economic analysis

Nomenclature

Symbol	Description	Unit
Latin symbols		
A	Collector or active BIPVT area	m^2
A_c	Tube contact area	m^2
c	Weibull scale parameter	–
C	Total cost	USD
C_{inst}	Installation or capital cost	USD
$C_{O\&M}$	Operation and maintenance cost	USD·yr^{-1}
C_p	Specific heat capacity	J·kg^{-1}·K^{-1}
d	Discount rate	–
D	Set of daylight hours	–
E_{el}	Electrical energy output	Wh
f_k	Regression tree function in XGBoost	–
F_{obs}	Cumulative distribution function of observed wind speed	–
F_{sat}	Cumulative distribution function of satellite/reanalysis wind speed	–
FV	Future value or accumulated invested savings	USD
GHI	Global horizontal irradiance	W·m^{-2}
GHI_{cs}	Clear-sky global horizontal irradiance	W·m^{-2}
GHI'	Screened satellite-derived global horizontal irradiance	W·m^{-2}
$h_{g\text{-}a}$	Convective heat-transfer coefficient between glass and ambient air	W·m^{-2}·K^{-1}

Received 19 October 2025. **Revised** 30 June 2026. **Accepted:** 13 July 2026

I	Incident solar irradiance	$W{\cdot}m^{-2}$
J	Number of leaves in an XGBoost regression tree	–
K	Number of regression trees in the XGBoost ensemble	–
k	Weibull shape parameter	–
k^*	Clear-sky index	–
$k^{*\prime}$	Screened clear-sky index	–
$l(y_i, \hat{y}_i)$	Prediction loss function	–
L_{PCM}	Latent heat of PCM	$J{\cdot}kg^{-1}$
L_{XGB}	XGBoost objective/loss function	–
$\dot{m}$	Mass flow rate	$kg{\cdot}s^{-1}$
MAE	Mean absolute error	Variable dependent
$MAPE_{day}$	Daylight-only mean absolute percentage error	%
MSE	Mean squared error	Variable-dependent squared units
MSS	Mean standard deviation scaling method	–
n	Number of samples	–
N_d	Number of daylight samples	–
P_{el}	Electrical power output	W
PP	Payback period	yr
$\dot{Q}$	Heat-transfer rate	W
$\dot{Q}_u$	Useful thermal-energy rate transferred to the water loop	W
RMSE	Root mean squared error	Variable dependent
R^2	Coefficient of determination	–
T_{Ab}	Absorber plate temperature	°C or K
T_{air}	Ambient air temperature	°C or K
T'_{air}	Corrected ambient air temperature	°C
T_f	Fluid temperature	°C or K
$T_{f,av}$	Average fluid temperature	K
T_g	Glass-cover temperature	°C
T_{in}	Inlet fluid temperature	°C
T_{ins}	Insulation temperature	°C
T_m	PCM melting temperature or melting range	°C
T_{out}	Outlet fluid temperature	°C
T_{PV}	PV cell temperature	°C
T_{ref}	Reference temperature under STC	°C
T_s	Effective sun temperature	K
T_t	Tube surface temperature	°C
U	Overall heat-transfer coefficient	$W{\cdot}m^{-2}{\cdot}K^{-1}$
$U_{Ab\text{-}ins}$	Conductance between absorber and insulation	$W{\cdot}m^{-2}{\cdot}K^{-1}$
$U_{Ab\text{-}t}$	Conductance between absorber and tube	$W{\cdot}m^{-2}{\cdot}K^{-1}$
$U_{PV\text{-}Ab}$	Conductance between PV layer and absorber	$W{\cdot}m^{-2}{\cdot}K^{-1}$
$U_{PV\text{-}g}$	Conductance between PV layer and glass	$W{\cdot}m^{-2}{\cdot}K^{-1}$
$U_{t\text{-}fluid}$	Conductance between tube and working fluid	$W{\cdot}m^{-2}{\cdot}K^{-1}$
$U_{t\text{-}ins}$	Conductance between tube and insulation	$W{\cdot}m^{-2}{\cdot}K^{-1}$

v_w	Wind speed	$m{\cdot}s^{-1}$
v'_w	Corrected wind speed	$m{\cdot}s^{-1}$
w_j	Leaf-node weight in XGBoost	–
x_t	Feature vector at time t	–
y	Target variable	Variable dependent
$\dot{E}_{x,in}$	Input exergy rate	W
$\dot{E}_{x,out}$	Output exergy rate	W
$\hat{y}$	Predicted value	Variable dependent
Greek symbols		
$\Omega(f_k)$	XGBoost tree-complexity penalty	–
α	Elastic-net mixing parameter	–
α_{PV}	Absorptance of PV laminate	–
α_g	Absorptance of glass	–
β	PV temperature coefficient	K^{-1}
β_{0h}	Intercept term in ENR model for forecast horizon h	Variable dependent
β_h	Regression coefficient vector in ENR for forecast horizon h	Variable dependent
γ	XGBoost leaf-complexity penalty	–
δ	Layer thickness	m
η_{STC}	PV efficiency at standard test conditions	–
η_T	Total first-law efficiency	–
η_{el}	Electrical efficiency	–
η_{ex}	Exergy efficiency	–
η_{th}	Thermal efficiency	–
$\eta_{th,net}$	Net thermal efficiency	–
λ	Thermal conductivity	$W{\cdot}m^{-1}{\cdot}K^{-1}$
λ_{EN}	Elastic-net regularization strength	–
λ_{XGB}	L2 regularization parameter in XGBoost	–
μ	Mean value	Variable dependent
μ_{obs}	Mean of observed station data	Variable dependent
μ_{sat}	Mean of satellite/reanalysis data	Variable dependent
ρ	Density	$kg{\cdot}m^{-3}$
σ	Standard deviation	Variable dependent
σ_{obs}	Standard deviation of observed station data	Variable dependent
σ_{sat}	Standard deviation of satellite/reanalysis data	Variable dependent
τ_g	Glass transmissivity	–
Subscripts		
0	Intercept/reference index where used	–
a	Ambient air	–
Ab	Absorber plate	–
av	Average value	–
cs	Clear-sky condition	–
d	Daylight sample/hour	–
el	Electrical output	–
f	Fluid or working medium	–
g	Glass cover	–
h	Forecast horizon	–

in	Inlet condition	–
ins	Insulation layer	–
j	Leaf index in XGBoost	–
k	Tree index in XGBoost	–
m	Melting condition	–
net	Net value	–
obs	Observation or ground-measured value	–
out	Outlet condition	–
PV	Photovoltaic laminate/cell	–
ref	Reference condition	–
sat	Satellite/reanalysis-derived value	–
STC	Standard test condition	–
t	Tube surface or time index, depending on context	–
th	Thermal output	–
w	Wind	–
x	Exergy	–
XGB	XGBoost-related parameter	–
Superscripts		
∗	Clear-sky index notation	–
ˆ	Predicted or estimated value	–
˙	Rate quantity per unit time	–
′	Corrected, screened, or postprocessed variable	–
Abbreviations		
AC	Alternating current	–
BIPV	Building-integrated photovoltaic	–
BIPV/T	Building-integrated photovoltaic/thermal	–
BIPVT	Building-integrated photovoltaic-thermal	–
CAMS	Copernicus Atmosphere Monitoring Service	–
CAPEX	Capital expenditure	–
CDF	Cumulative distribution function	–
CSI	Clear-sky index	–
DC	Direct current	–
DHI	Diffuse horizontal irradiance	–
DHW	Domestic hot water	–
DNI	Direct normal irradiance	–
DSF	Double-skin façade	–
DT	Decision tree	–
EN	Elastic net	–
ENR	Elastic net regression	–
EPW	EnergyPlus weather file	–
ERA5	Fifth-generation ECMWF atmospheric reanalysis	–
EVA	Ethylene-vinyl acetate	–
GHG	Greenhouse gas	–
GHI	Global horizontal irradiance	–
ICAO	International Civil Aviation Organization	–
LCC	Life-cycle cost	–
LCCA	Life-cycle cost analysis	–
LCOE	Levelized cost of electricity	–
LSTM	Long short-term memory	–
MAE	Mean absolute error	–
MAPE	Mean absolute percentage error	–

ML	Machine learning	–
MSE	Mean squared error	–
MSS	Mean standard deviation scaling	–
NASA POWER	NASA Prediction of Worldwide Energy Resources	–
O&M	Operation and maintenance	–
PCM	Phase change material	–
PP	Payback period	–
PV	Photovoltaic	–
PV/T	Photovoltaic/thermal	–
PVT	Photovoltaic-thermal	–
RF	Random forest	–
RMSE	Root mean squared error	–
ROI	Return on investment	–
SI	International System of Units	–
STC	Standard test conditions	–
SVR	Support vector regression	–
TRNSYS	Transient System Simulation Tool	–
UPV	Uniform present value	–
WMO	World Meteorological Organization	–
XGBoost	Extreme Gradient Boosting	–

1. Introduction

In recent years, solar energy technologies have attracted considerable attention as sustainable solutions for reducing fossil fuel dependency and mitigating environmental impacts [1]. Recent advances in prediction and intelligent energy management methods have also contributed to improving the efficiency and operational performance of renewable and electrified energy systems [2]. The building sector accounts for a large share of final energy consumption and carbon dioxide emissions [3–5]. Therefore, green buildings and active energy shells have been proposed as effective strategies in the energy transition. At the same time, combining solar energy into the building envelope can supply both electricity and heat while helping to moderate façade thermal loads [6–8].

The evolution of building solar technologies began with building-integrated photovoltaic-thermal (BIPVT) (electricity only) and thermal collectors (heat only) and reached PVT/hybrid photovoltaic-thermal (hPVT), which generates both electricity and heat by recycling the heat behind the cell [9–11]. The next step is BIPVT, in which the PVT system is combined directly into the building envelope, particularly the façade. In addition to the simultaneous generation of electricity and heat, this integration provides architectural-thermal benefits such as shading and reducing the radiation load of the shell, which is of double value in hot and dry climates. In this regard, Yu *et al.* [12] presented a comprehensive review of façade-based BIPVT systems and classified them into seven main categories according to the use of recovered thermal energy: PV cooling by air, space heating, ventilation, water heating, PV-phase change material (PCM) systems, BIPVT coupled with heat pumps, and PV thermoelectric walls. Their review showed that façade-based BIPVT systems can simultaneously increase solar energy utilization, improve PV thermal management, and reduce building heating/cooling loads. However, the reported performance varied substantially among designs because electrical output, thermal recovery, and building-load reduction depend on the working fluid, façade configuration, thermal storage strategy, and climatic condition. The review also emphasized that future work should move beyond isolated performance assessment toward climate-responsive design, standardized performance evaluation, and integrated techno-economic analysis.

More recently, Şirin *et al.* [13] reviewed BIPV/T systems for green buildings and emphasized that these systems can use building façades to overcome urban space limitations while simultaneously producing electrical and thermal energy. Their review classified BIPV/T configurations, discussed performance-improvement strategies, and highlighted that façade-integrated systems can contribute to net-zero and energy-efficient buildings when technical design is combined with appropriate economic and policy conditions. However, the review also indicated that many studies still focus on system classification and component-level performance rather than integrated climate-specific techno-economic workflows. Mangherini *et al.* [14] recently investigated the combination of paraffin PCM into BIPV windows based on semitransparent luminescent solar concentrator technology. Using finite-element simulations in COMSOL Multiphysics, they optimized PCM latent heat capacity and melting temperature to improve thermal regulation of PV cells embedded in the window frame. Their results showed that a PCM melting temperature of 36°C combined with higher latent heat reduced the nominal operating cell temperature from 42°C to 36°C and increased the overall PV-window power output by 11.80%. This study confirms the recent interest in PCM-based thermal regulation for building-integrated PV envelopes, but it focused on semitransparent PV windows rather than an active water-based BIPVT façade with domestic hot water (DHW) recovery and operational forecasting. Čurpek and Čekon [15] further examined the climate response of a BIPV façade enhanced with latent PCM-based thermal energy storage. Their study focused on how PCM thermophysical properties and latent heat utilization interact with incident solar radiation and outdoor air temperature during the diurnal cycle. The results demonstrated that the effectiveness of PCM in façade-integrated PV applications is strongly climate dependent and controlled by the matching between PCM phase-change temperature, solar gains, and ambient temperature variation. This finding is directly relevant to hot-arid contexts, where PCM selection must be linked to local irradiance and temperature profiles rather than treated as a generic material choice.

However, a reliable techno-economic assessment for BIPVT requires accurate hourly climate inputs (irradiance, temperature, wind) [16, 17]. In many regions (including cities in northern Saudi Arabia), access to continuous and high-quality station data is limited (low dispersion, temporal gaps, or lack of radiation measurements) [18, 19]. Therefore, satellite/reanalysis data are considered an essential alternative/complement, but their use requires validation against ground observations and bias correction to control input uncertainties [20]. In addition, realistic operation and short-term planning of BIPVT systems require day-ahead hourly forecasts of solar irradiance and temperature [21]. Machine learning (ML) plays a key role both in forecasting hourly time series and in postprocessing satellite data to reduce bias [22]. Instead of relying on a single model, a comparison framework [such as long short-term memory (LSTM), support vector regression (SVR), and XGBoost] allows for the selection of the optimal model based on real-world performance. In this regard, Imam *et al.* [23] evaluated satellite/reanalysis-driven, ML forecasting as an operational pathway to generate day-ahead hourly drivers suitable for BIPVT modeling. Using historical NASA POWER data for a representative Northern-Saudi site, they developed and compared ML regressors, including SVR, random forest (RF), artificial neural network, decision tree (DT), elastic net (EN), and linear regression to predict global horizontal irradiance (GHI) and ambient temperature. To control input uncertainties, satellite series were first validated against available ground observations and bias corrected prior to ML training. Model performance was assessed using the coefficient of determination (R^2), root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) on time-ordered splits to reflect real-world deployment. Their results revealed clear quantitative differences among the tested algorithms. The DT model achieved the best performance, with $R^2 = 1.0$, MSE = 0.0, RMSE = 0.0, MAPE = 0.0%, and MAE = 0.0, whereas EN showed the weakest performance, with $R^2 = 0.8396$, MSE = 0.4389, RMSE = 0.6549, MAPE = 9.66%, and MAE = 0.5534. Although feature-selection methods produced limited improvements for most models, they substantially enhanced the SVR model. These numerical differences demonstrate that the selection of an ML model can strongly affect GHI forecasting performance and support the use of a comparative forecasting framework rather than relying on a single algorithm. By delivering validated, day-ahead hourly irradiance and temperature, the proposed approach provided the operational inputs required by BIPVT performance and techno-economic analyses, thereby addressing the data-scarcity constraint common to many Saudi locales and similar hot-arid regions.

To improve the reliability of the simulation results, out-of-model validation is also important. Comparing the BIPVT physical model with a standard tool, such as EnergyPlus, helps confirm that the results are reproducible and not dependent on a single implementation [24–26]. For this purpose, Vuong *et al.* [27] developed a BIPV/T object (BIPVTSYSTEM) in EnergyPlus and benchmarked it against the well-established TRNSYS Type 567 formulation. The EnergyPlus implementation followed the same governing equations with targeted adjustments to thermal–fluid relations, enabling application to both façade- and roof-integrated BIPV/T assemblies and large-scale community modeling. A full-year comparative simulation highlighted areas of close agreement as well as systematic discrepancies attributable to differences in weather data interpolation, sky-temperature computation, and electrical submodels in the two tools. These results provided a transparent, tool-agnostic reference for BIPV/T simulation and showed how cross-tool validation can diagnose modeling sensitivities and strengthen the reproducibility of BIPV/T performance predictions. It directly addresses the need for identification in the above paragraph.

This study was conducted in Tabuk, Saudi Arabia, a region characterized by high solar radiation, a hot and dry climate, considerable DHW demand, and extensive urban façade surfaces. National solar-resource assessments indicate that Saudi Arabia receives high daily GHI values, typically around 5.7–6.7 $kWh \cdot m^{-2} \cdot day^{-1}$, with inland regions generally outperforming coastal locations; within this context, the Tabuk region has been reported as a favorable northwestern site for PV applications, especially during the high-load summer season [28, 29]. In this regard, Alqaed [30] analyzed façade strategies across three Saudi cities—Jeddah, Abha, and Tabuk—representing distinct climate regimes, with a particular emphasis on Tabuk's hot-dry conditions and extensive urban façades. Using DesignBuilder to resolve façade heat-transfer processes, the study compared a simple façade, a double-skin façade (DSF), and a PCM-filled DSF over representative months. The results showed that the energy effect of façade design was strongly climate dependent. In Jeddah, DSF with PCM reduced the required energy by 11.5% during cold months and 5.6% during warm months compared with a simple façade. In Abha, a simple DSF reduced energy demand by 20% during cold months but increased it by 14.5% during warm months, indicating that the same façade strategy may not be equally effective across Saudi climates. For Tabuk, which is the focus of the present study, a simple DSF and a DSF with PCM reduced heating energy by 18% and 40%, respectively; however, the simple DSF increased cooling energy by 5%, while the DSF with PCM reduced it by 25% compared with a simple façade. These quantitative findings confirm that PCM-assisted façade systems can provide measurable energy benefits in Saudi climates, particularly in Tabuk. However, Alqaed's study focused on passive façade configurations and building heating/cooling demand, without quantifying the coupled electrical output, thermal recovery, ML-based operational forecasting, EnergyPlus cross-validation, and life-cycle economic performance of an active BIPVT façade system.

A closer active system example was presented by González Gallero *et al.* [31] who developed and analyzed a novel façade-based BIPVT system with PCM for DHW production. The proposed modular system used PV glass for electricity generation, PCM for heat storage, and a water-PCM exchanger for DHW production. Their design and verification were supported by numerical modeling, including computational fluid dynamics-based assessment, and the system was shown to meet DHW requirements under typical usage conditions. This work is closely related to the present study because it combines façade integration, PCM storage, electricity generation, and DHW production; however, it did not address hot-arid Saudi boundary conditions, ML-based day-ahead forecasting, or life-cycle economic comparison of façade orientations. In a Saudi context, Alsagri and Alrobaian [32] analyzed and predicted the performance of a BIPV/T system with and without PCM under the weather conditions of Sharurah, Saudi Arabia. Their study coupled TRNSYS, MATLAB, and Python, and used RF prediction for future performance assessment. The reported results showed that the PCM layer reduced PV cell temperature by up to 25.47°C, corresponding to a 37.87% reduction, while thermal efficiency ranged from 29.65% to 59.40%. RF prediction also achieved high accuracy, with R^2 values around 96%–98%. This study is particularly relevant because it combines PCM-assisted BIPV/T and ML prediction in Saudi Arabia; however, it focused on Sharurah and did not evaluate façade orientation, satellite-data screening, EnergyPlus cross-validation, or life-cycle economic performance for Tabuk.

Although previous studies have separately examined façade-integrated BIPVT systems, PCM-assisted façade concepts, ML-based solar forecasting, and technoeconomic assessment, a unified and operation-oriented framework for hot-arid Saudi residential buildings is still missing. This gap is particularly important for Tabuk, where high solar irradiance, elevated ambient temperature, DHW demand, and extensive façade surfaces coexist.

The novelty of this study lies in developing an integrated, operation-oriented assessment framework for PCM-stabilized BIPVT façades under hot-arid Saudi conditions. Unlike many previous BIPVT studies that treated façade combination, PCM thermal buffering, solar forecasting, model validation, and economic feasibility as separate issues, this work links these components within a single workflow. Specifically, the study advances previous research by combining locally validated temperature and wind inputs, physically screened satellite irradiance products, comparative day-ahead ML forecasting using elastic net regression (ENR) and XGBoost, a layer-resolved PCM-BIPVT physical model, EnergyPlus-based cross-validation, and life-cycle technoeconomic assessment for different façade orientations in Tabuk.

Accordingly, this study addresses the following research question: can an ML-assisted, PCM-stabilized BIPVT façade provide technically reliable and economically feasible performance for residential buildings under data-limited hot-arid conditions? To answer this question, the specific objectives of the study are (i) to prepare a quality-controlled climate dataset for Tabuk by validating temperature and wind speed against local station records and screening satellite-derived irradiance using NASA POWER–CAMS consistency and clear-sky index (CSI) checks; (ii) to compare ENR and XGBoost for day-ahead forecasting of the main climate drivers required for BIPVT operation; (iii) to couple the forecasted climate series with a layer-resolved PCM-stabilized BIPVT façade model and quantify its electrical, thermal, total, and exergy performance; (iv) to evaluate the role of the PCM layer in moderating PV temperature and stabilizing thermal output under hot-arid conditions; (v) to cross-validate the in-house BIPVT simulation results with EnergyPlus; and (vi) to assess the life-cycle technoeconomic feasibility of different façade orientations using return-on-investment and payback-period indicators.

The paper is structured as follows: Section 2 explains the study scope, data, ML framework, BIPVT physical model and scenarios, economic evaluation methodology and validation; the Section 3 discusses the validation findings, technical performance, economic comparison of scenarios, and sensitivity analysis; and finally, Section 4 provides a summary and future directions.

2. Methodology

2.1. Case study and utilized data

The case study is Tabuk, Saudi Arabia (28.38° N 36.56° E; Fig. 1). Tabuk is located in the northwestern inland part of Saudi Arabia and represents a hot-arid climate with high solar availability, hot summers, and relatively mild winters. Previous national-scale solar-resource assessments have shown that daily GHI across Saudi Arabia is generally high, commonly ranging from ∼5.7 to 6.7 kWh·m^{-2}·day^{-1}, with inland locations often receiving higher solar resources than coastal areas. In comparative PV assessments across Saudi sites, inland northern and northwestern locations, including the Tabuk region and nearby Timaa/Tayma area, have been identified as favorable for solar-power applications, particularly during the high-load summer period. Therefore, Tabuk provides a relevant case for evaluating façade-integrated solar technologies under conditions of strong solar input, elevated ambient temperature, and substantial cooling/DHW demand. Its urban fabric also offers extensive façade area suitable for BIPVT

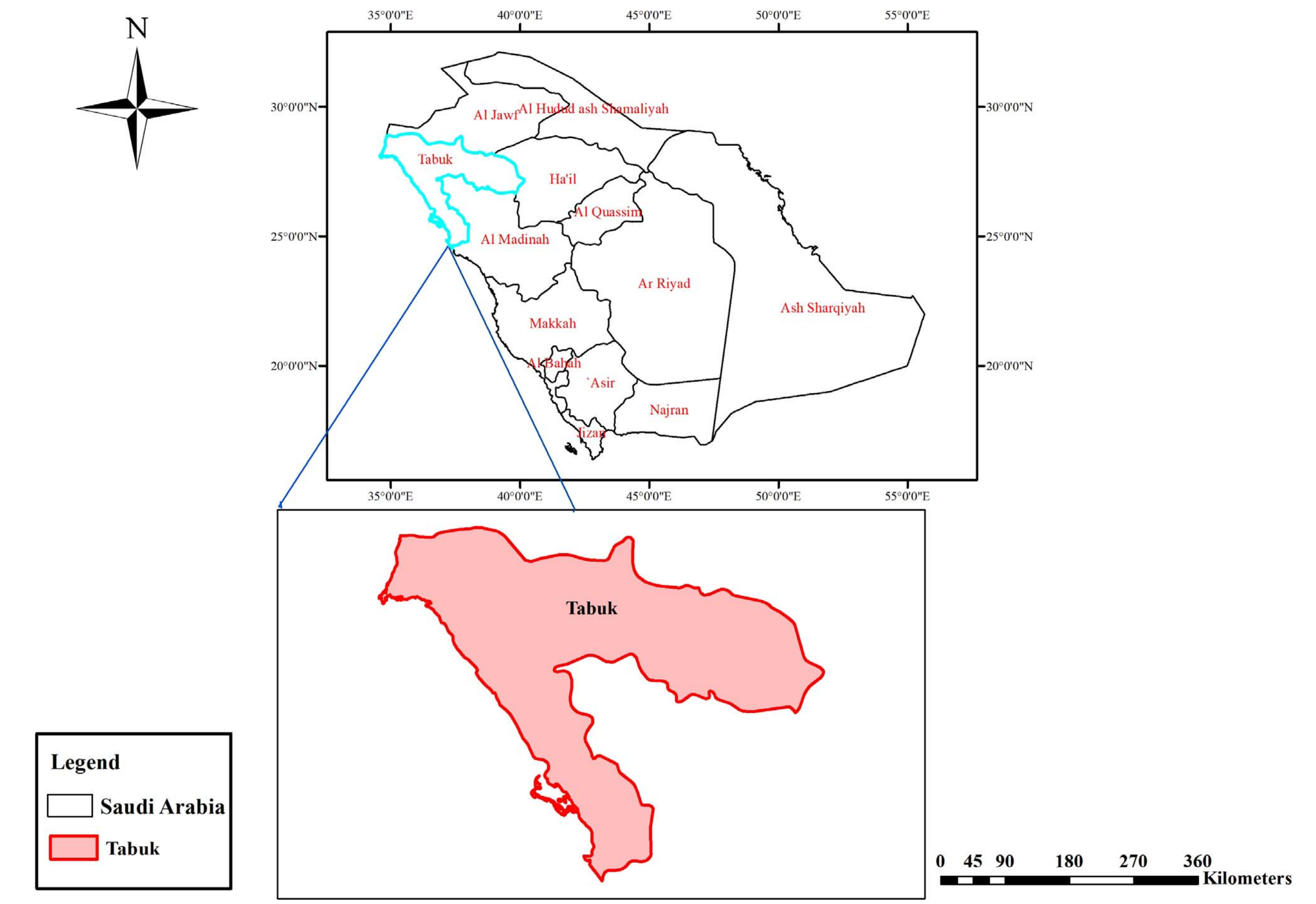

Figure 1 Map of the study area, Tabuk, Saudi Arabia.

installation, making the city an appropriate testbed for assessing the combined electrical, thermal, and economic performance of PCM-stabilized BIPVT façades [33–35].

Three hourly climatic drivers were required for reliable BIPVT simulation and forecasting:

(i) GHI [W·m^{-2}], (ii) ambient temperature (T_{air}) [°C], and (iii) 10-m wind speed (v_w) [m·s^{-1}]. Table 1 summarizes the satellite/reanalysis datasets, while the ground-station metadata are described in the text to improve reproducibility.

Satellite/reanalysis products were used as the primary climate-data source. Ground-based observations were used for validation where reliable station records were available. In particular, ambient temperature and 10-m wind speed were checked against observations from the Tabuk Regional Airport meteorological station, also known as Prince Sultan bin Abdulaziz International Airport, identified by WMO station code 40375 and ICAO code OETB. The station is located near Tabuk at ∼28.37° N 36.60–36.62° E, with an elevation of ∼778 m above sea level. The validation period was limited to the temporal overlap between the station records and the satellite/reanalysis products used in this study.

Before validation, the station records were converted to a consistent hourly timestamp and checked for duplicated records, missing values, physically unrealistic values, and isolated spikes. Temperature records outside a physically plausible range for Tabuk were removed, and wind-speed values below 0 m·s^{-1} or unrealistically high isolated values were flagged as invalid. Missing or invalid station records were not infilled for validation; instead, only time steps with simultaneous station and satellite/reanalysis values were retained for metric calculation and bias adjustment. When multiple station records were available within the same hour, they were averaged to match the hourly resolution of ERA5 and NASA POWER. All timestamps were aligned to local time and then matched with the corresponding satellite/reanalysis records using the nearest hourly time step. This ensured that validation and bias-adjustment statistics were calculated from temporally consistent station–satellite pairs.

This study used a 1-year hourly climate dataset covering 1 January 2024 to 31 December 2024. The same 2024 dataset was used for climate characterization, data screening/bias adjustment, ML model development, BIPVT simulation, and EnergyPlus cross-validation. Therefore, the climatological summaries and figures are consistently reported for 2024.

Figure 2 presents monthly summaries of the three drivers used in this study. Figure 2a shows that average daily shortwave energy peaks from late spring to midsummer (May–August), reaching its highest value around mid-June and its lowest levels in winter (∼8.8 kWh·m^{-2}·day^{-1}) and minima in winter (December–February, ∼4–5 kWh·m^{-2}·day^{-1}). This pronounced seasonal cycle confirms sufficient input energy for both electricity generation and DHW preheating, and motivates assessing façade orientations (vertical vs inclined) under peak radiation months. Figure 2b illustrates the monthly temperature distributions, showing hot summers and mild winters (typical daytime highs of 35°C–42°C) and mild winters (mean temperatures ≈ 10°C–18°C). Elevated

Table 1 Overview of datasets used.

Variable	Dataset source	Provider/link	Temporal resolution	Spatial resolution	Notes
GHI ($W{\cdot}m^{-2}$)	NASA POWER	https://power.larc.nasa.gov	Hourly	$0.5^{\circ} \times 0.5^{\circ}$ (∼55 km)	Long-term satellite-derived solar radiation
GHI, DNI, DHI	CAMS radiation service	https://atmosphere.copernicus.eu	Hourly/15 min	$0.1^{\circ} \times 0.1^{\circ}$ (∼10 km)	Higher accuracy for solar irradiance
T_{air} (°C)	ERA5 reanalysis	https://cds.climate.copernicus.eu	Hourly	$0.25^{\circ} \times 0.25^{\circ}$ (∼30 km)	Widely used for temperature/wind fields
Wind speed ($m{\cdot}s^{-1}$)	ERA5 reanalysis	https://cds.climate.copernicus.eu	Hourly	$0.25^{\circ} \times 0.25^{\circ}$ (∼30 km)	10-m wind speed used for convective coefficients

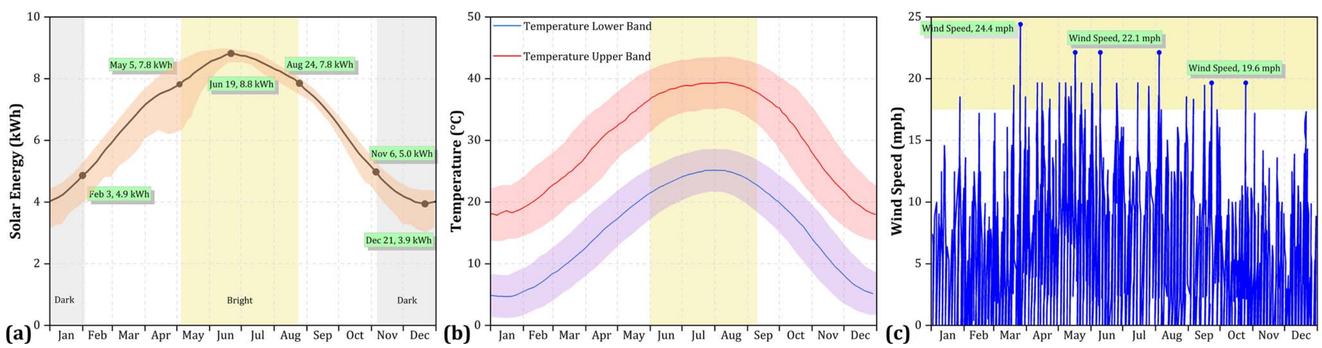

Figure 2 Monthly summaries of (a) average daily incident shortwave solar energy, (b) distributions of ambient temperature, and (c) distributions of 10-m wind speed in Tabuk during 2024.

summer cell temperatures justify active cooling and the PCM scenario to stabilize PV cell temperature and sustain electrical efficiency during peak irradiance. Figure 2c shows that wind speeds are generally moderate, with greater variability in spring and slightly calmer conditions from late summer to autumn. Since the external convective coefficient in the façade model depends on v_w, these wind patterns affect heat removal from the PV back surface. Therefore, the sensitivity analysis includes the observed monthly variability.

The hourly satellite/reanalysis series were temporally matched with the available ground-station observations. Ambient temperature and wind speed were validated against the Tabuk Regional Airport station records for the overlapping period, while the irradiance products were evaluated using interproduct consistency between NASA POWER and CAMS and physical plausibility checks based on the CSI. Therefore, station-based validation was applied only to meteorological variables available from the ground station, whereas GHI was treated as a screened satellite-derived input rather than a ground-validated variable. This procedure improved the transparency and local consistency of the climate inputs used in day-ahead forecasting and technoeconomic evaluation.

2.2. Climate-data validation, screening, and bias adjustment

The quality-controlled and temporally aligned station–satellite pairs described in Section 2.1 were used for station-based bias adjustment of ambient temperature and wind speed. Since continuous ground-based GHI observations were not available from the Tabuk Regional Airport station, irradiance preprocessing was limited to satellite-product consistency checks and CSI-based physical screening.

The correction and screening method was selected according to the physical and statistical behavior of each climatic variable and the availability of ground observations. Solar irradiance has a strong deterministic diurnal and seasonal structure controlled by solar geometry; therefore, the CSI was used to normalize GHI by the theoretical clear-sky irradiance and to identify physically unrealistic satellite values. Station-based quantile mapping was not applied to GHI because continuous local GHI measurements were unavailable. Ambient temperature is a smoother variable whose satellite/reanalysis bias is commonly dominated by shifts in mean and variance; therefore, mean standard deviation scaling (MSS) was used to adjust its first two statistical moments against station observations. Wind speed is nonnegative, skewed, and commonly represented by Weibull-type distributions; therefore, Weibull-based quantile mapping was used to correct both the magnitude and the distributional shape of the wind-speed series.

The calibration period for all validation, screening, and adjustment procedures was the 1-year modeling period from 1 January 2024 to 31 December 2024. For station-based adjustment, the matched station–reanalysis pairs for ambient temperature and wind speed were grouped by month and hour of day, resulting in 12×24 conditional bins. This binning was used to preserve the strong seasonal and diurnal structure of Tabuk's climate. For temperature, the MSS correction was calibrated separately within each month–hour bin using the mean and standard deviation of the paired station and reanalysis values. For wind speed, the Weibull-based quantile mapping was calibrated within the same month–hour bins. All valid paired samples within each bin were ranked and used to estimate empirical probabilities before

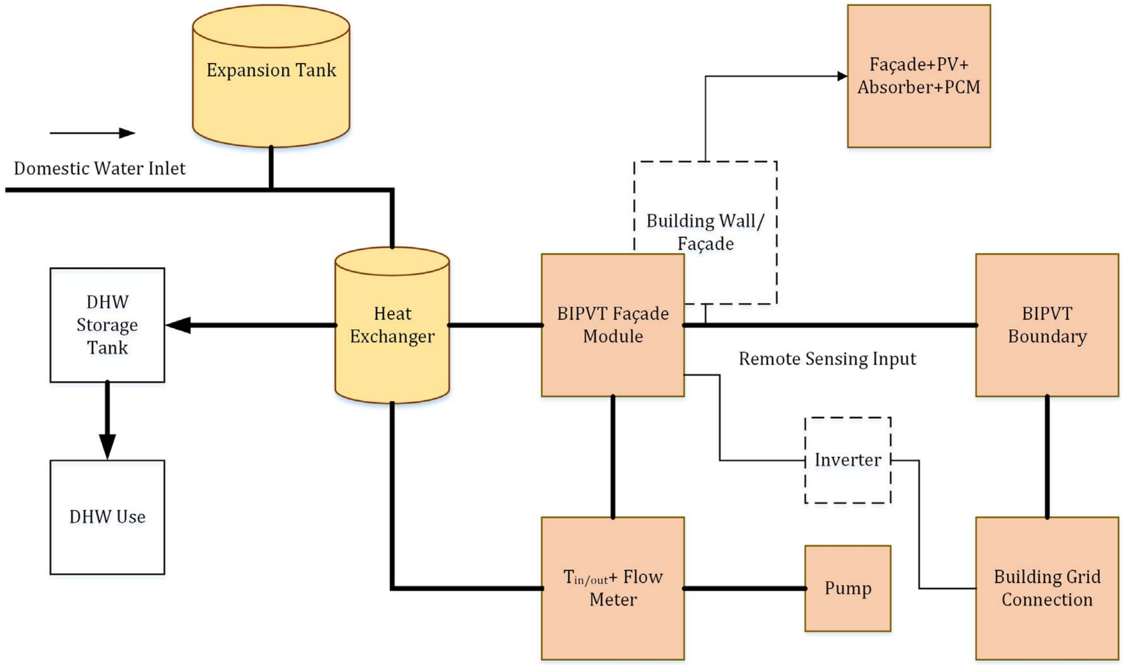

Figure 3 Schematic diagram of the façade-integrated BIPVT system with water loop for DHW, PCM stabilization layer, and grid-tied inverter connection.

mapping through the fitted Weibull cumulative distributions; therefore, no fixed number of quantiles was imposed, and the effective number of quantile points depended on the number of valid paired records in each bin. A separate cross-validation step was not applied to the bias-adjustment procedure because this step was used only to construct the quality-controlled reference climate series for the 2024 case-study simulation. To avoid temporal inconsistency, correction parameters were derived only from matched station–reanalysis pairs within the 2024 modeling period, while predictive generalization was evaluated separately in the ML forecasting stage using chronological hold-out testing.

For *GHI*, the CSI approach was used to normalize the irradiance series and to identify physically unrealistic values [36, 37]. A clear-sky irradiance GHI_{cs} was first estimated using the Ineichen–Perez clear-sky model. The CSI k^* is defined in Eq. (1):

$$k^* = \frac{GHI}{GHI_{cs}} \tag{1}$$

where *GHI* denotes the measured or satellite-derived GHI [W m^{-2}], GHI_{cs} is the clear-sky irradiance estimated from the clear-sky model [W·m^{-2}], and k^* is the dimensionless CSI.

Because continuous ground-based *GHI* observations were not available from the Tabuk Regional Airport station, station-based quantile mapping was not applied to irradiance. Instead, NASA POWER and CAMS irradiance products were compared during the study period, and physically unrealistic values were screened using CSI constraints. Values with negative irradiance, nighttime nonzero artifacts, or unrealistically high CSI values were removed or clipped. The screened irradiance series was then reconstructed as Eq. (2) [38, 39]:

$$GHI' = k^{*'} \times GHI_{cs} \tag{2}$$

where $k^{*'}$ denotes the screened CSI after physical plausibility filtering and interproduct consistency checking, and $GHI^{'}$ is the resulting quality-controlled irradiance series used for forecasting and BIPVT simulation. For ambient temperature ($\underline{T}_{air}$), an MSS method was applied in each month and hour-of-day bin [40, 41]. The correction is expressed as follows:

$$T' = \frac{T - \mu_{sat}}{\sigma_{sat}} \sigma_{obs} + \mu_{obs} \tag{3}$$

where T is the raw satellite-derived temperature [°C], μ_{sat} and σ_{sat} are the mean and standard deviation of the satellite data, μ_{obs} and σ_{obs} are the mean and standard deviation of the ground observations, and $T^{'}$ is the corrected temperature [°C].

For wind speed (v_W), quantile mapping was also employed, as wind data typically follow a Weibull distribution [42]. Thus, a parametric mapping of the Weibull shape (k) and scale (c) parameters was applied (Eq. 4):

$$v_w' = F_{obs}^{-1}\left(F_{sat}\left(v_w\right)\right) \tag{4}$$

where v_W is the raw satellite-derived wind speed [m·s^{-1}], F_{sat} and F_{obs} are cumulative distribution functions of satellite and observed wind speeds, respectively, and v'_w is the corrected wind speed [m·s^{-1}].

2.3. Façade-integrated building-integrated photovoltaic-thermal system

The schematic configuration of the BIPVT system considered in this study is illustrated in Fig. 3.

In contrast to conventional hybrid PVT arrangements, where the module is usually a stand-alone collector connected to auxiliary components, the proposed BIPVT module was directly integrated into the building envelope. This architectural integration enabled simultaneous production of electricity and thermal energy while reducing the radiative and convective heat load on the building façade, which was of particular relevance in hot–arid climates such as Tabuk. The system is designed to meet both the electricity demand of the building (through grid connection) and the DHW demand (through a hydronic loop), thereby ensuring dual-use of the captured solar resource.

In the proposed configuration, the BIPVT façade module consists of a glass cover, PV cells laminated onto a substrate, an absorber plate with embedded tubes, and a back insulation layer attached to the building wall. An optional thin layer of PCM is placed between the absorber and the insulation in order to stabilize the thermal response during peak irradiance, reduce the temperature rise of PV cells, and enhance both electrical efficiency and the usability of the thermal output. A forced circulation water loop is employed as the working fluid baseline, which is more realistic for building applications in Saudi Arabia than nanofluids considered in other studies. Water enters from the domestic inlet, passes through the pump and flow–temperature sensors, circulates inside the absorber tubes of the BIPVT module, and transfers heat to the heat exchanger, which charges a DHW storage tank. The hot water is then delivered to the building load. An expansion tank is included to maintain system pressure stability. On the electrical side, the PV module output is directed to a maximum power point tracking inverter, which ensures efficient DC–AC conversion and integration with the building grid connection. Unlike previous BIPVT configurations that employed a charge controller and battery, this grid-tied pathway better reflects practical implementation in residential buildings in Tabuk. The boundary conditions of the simulation are driven by quality-controlled satellite irradiance (GHI') and station-adjusted ambient temperature (T'_{air}) and wind speed (v'_{w}). The external boundary condition of the BIPVT façade was defined using the hourly quality-controlled climate series for Tabuk. Solar irradiance was applied to the exterior glass surface, ambient temperature was used as the outdoor air temperature, and wind speed was used to calculate the external convective heat-transfer coefficient. The indoor side of the wall was represented by a fixed residential indoor set-point temperature of 24°C, which is commonly adopted for cooling-dominated residential simulations in hot climates. Heat transfer through the back concrete wall was modeled as one-dimensional conduction between the BIPVT rear insulation and the indoor zone. Long-wave radiation exchange with the sky and surroundings was represented through an effective sky-temperature approach, while short-wave radiation was applied as incident solar irradiance after glass transmission and absorption losses. Thermal contact between adjacent solid layers was assumed to be ideal, except where the effective conductance terms in the layer-resolved equations account for material thickness and conductivity. The PCM layer was modeled as a paraffin-class material because paraffin PCMs are widely used in building-integrated solar thermal and PV/T applications due to their chemical stability, negligible supercooling, commercial availability, and suitable latent heat capacity. The selected melting range of 45°C–55°C was chosen to overlap with the expected PV back-surface and absorber temperatures under peak solar conditions in Tabuk. This temperature interval allows the PCM to absorb heat during midday overheating, when PV temperature normally rises above the efficient operating range, and to release stored heat later when irradiance decreases. The adopted PCM properties, including latent heat of $\sim$180–220 $kJ{\cdot}kg^{-1}$, density of $\sim$800 $kg{\cdot}m^{-3}$, and thermal conductivity of $\sim$0.20 $W{\cdot}m^{-1}{\cdot}K^{-1}$, are consistent with typical paraffin-based PCM values reported for solar façade and PV/T thermal-regulation applications.

In the following, the main simplifying assumptions of the BIPVT model are mentioned as follows:

I. Heat transfer was treated as one-dimensional across the façade layers;
II. Each layer was assumed to have uniform thermophysical properties;
III. Thermal contact resistances between adjacent layers were neglected and represented implicitly through effective conductance terms;
IV. The water flow rate was assumed constant during each simulation run;
V. Dust accumulation, long-term PV degradation, PCM aging, and moisture effects were not included in the hourly thermal model;
VI. The indoor temperature was fixed at 24°C;
VII. Long-wave radiation exchange was represented using an effective sky-temperature formulation; and
VIII. The PCM phase change was represented by an effective heat-capacity method over the melting interval rather than by an explicit moving-boundary phase-change model.

PCM latent heat storage was included using an effective heat-capacity approach. Within the melting interval, the sensible heat capacity of the PCM was increased by an equivalent latent-heat term so that the PCM could absorb heat during phase transition and release it when irradiance decreased. Therefore, the PCM was not treated as a purely static insulation layer. Instead, its temperature-dependent heat storage was included in the absorber–insulation energy balance. This formulation captures the main thermal-buffering effect of the PCM while avoiding the computational complexity of explicitly tracking a moving solid–liquid interface. The energy absorbed by the PCM during melting reduces the rate of heat accumulation behind the PV laminate, thereby limiting midday PV temperature rise and improving electrical efficiency.

The thermal-electrical performance of the system is governed by energy balances across its layers. The detailed heat-transfer mechanism of the façade-integrated BIPVT module is illustrated in Fig. 4.

Figure 4 illustrates the direction of energy transfer and the main thermal nodes considered in the layer-resolved BIPVT model. The incoming solar irradiance first reaches the glass cover, where a small portion is reflected or absorbed and the remaining radiation is transmitted toward the PV laminate. Part of the absorbed solar energy in the PV layer is converted into electricity, while the remaining fraction is transformed into heat and transferred toward the absorber plate. The absorber conducts heat to the bonded fluid tube, where the circulating water removes useful thermal energy for DHW production. The PCM layer, positioned behind the absorber, acts as a thermal buffer by absorbing excess heat during high-irradiance periods and releasing it when irradiance decreases. The back insulation reduces heat losses toward the building wall, while the wall layer represents the actual façade coupling with the indoor environment. Therefore, Fig. 4 is not only a material-layer diagram but also a representation of the coupled radiative, conductive, convective,

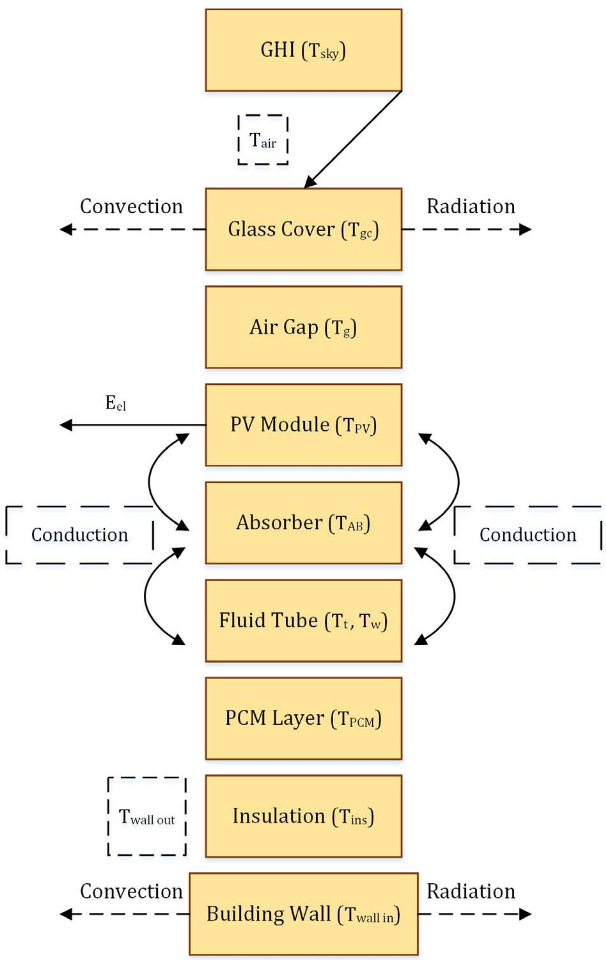

Figure 4 Layered thermal-electrical schematic of the façade-integrated BIPVT module, showing solar input, electrical extraction, conductive heat transfer through the solid layers, convective heat removal by the water tube, PCM latent heat storage, and thermal coupling with the building wall.

electrical, and latent-heat pathways used in the energy-balance model.

Unlike conventional stand-alone BIPVT systems, where the thermal unit is separated from the building envelope, the present configuration explicitly models the BIPVT as part of the façade. The layered structure includes (i) a glass cover, which transmits most of the incident solar irradiance while reflecting a small fraction; (ii) an optional air gap that reduces convective losses and enhances insulation; (iii) a PV laminate, responsible for electrical generation; (iv) a metallic absorber plate with embedded fluid tubes that extracts the heat accumulated behind the PV; (v) a water stream as the baseline working fluid, chosen due to its availability, high specific heat capacity, and practical compatibility with domestic systems in Saudi Arabia; (vi) a thin PCM layer inserted between absorber and insulation, providing latent heat storage to smooth out temperature fluctuations and to mitigate PV overheating during peak irradiance; (vii) a back insulation layer that minimizes unwanted heat losses; and (viii) the building wall itself, which is thermally coupled to the BIPVT unit and represents the realistic architectural boundary of façade integration. This configuration was selected for several reasons. First, the addition of the PCM layer improves thermal inertia, allowing the PV cells to operate at lower average temperatures, thereby enhancing electrical efficiency and extending module lifetime. Second, the explicit coupling with the building wall captures the cobenefits of façade integration, including shading and reduced transmission of heat into the indoor space, which is especially valuable in the hot-arid climate of Tabuk. Third, the adoption of water as the working fluid, instead of nanofluids considered in previous BIPVT studies, reflects a more practical and economically viable solution for large-scale residential applications. Finally, the layered formulation permits rigorous accounting of radiative, convective, and conductive heat transfers, which are crucial for reliable technoeconomic assessment. The model resolves radiative, convective, and conductive exchanges between layers, as well as the outputs of electricity (E_{el}) and hot water (heat to water). The additional PCM layer provides latent heat storage, while the building wall explicitly represents façade integration. To implement the model, the thermophysical properties of all layers were specified as summarized in Table 2.

Material properties were chosen to represent realistic building-integrated systems under Saudi conditions. The glass cover is sized at 4 mm with $\lambda \approx 1$ W·m^{-1}·K^{-1} and $\tau_g \approx 0.90$ to reflect common low-iron glazing on BIPV modules; this gives robust transmittance in Tabuk's high-irradiance conditions while maintaining mechanical integrity. The PV laminate properties ($\delta \approx 3$ mm, $\lambda \approx 0.2$ W·m^{-1}·K^{-1}, $c_p \approx 678$ J·kg^{-1}·K^{-1}) capture the composite nature of crystalline-Si cells encapsulated in EVA with a polymeric backsheet; using $\eta_{STC} \approx 0.18$ and $\beta \approx 0.004$ K^{-1} provides a conservative electrical baseline for hot climates where cell temperature is the dominant derating factor. The absorber plate is modeled in aluminum ($\lambda \approx 205$ W·m^{-1}·K^{-1}, δ 1 mm) to secure high in-plane thermal spreading with low mass, which helps keep PV cell temperatures down under peak GHI while enabling efficient heat pickup by the tubes. Copper is retained for the bonded serpentine tubes ($\lambda \approx 385$ W·m^{-1}·K^{-1}, $\delta \approx 1$ mm) to minimize conduction resistance from absorber to fluid; the combination Al-plate/Cu-tube is standard in compact PVT collectors and aligns with manufacturability. Water is chosen as the working fluid because its high heat capacity ($\approx$ 4180 J·kg^{-1}·K^{-1}) delivers strong thermal buffering at low cost and without the maintenance and stability issues of nanofluids; in the implementation, $c_p(T)$ can vary with temperature, but the tabulated value at 25°C is a clear reference. The PCM layer is set to 10 mm of a paraffin-based material with $\lambda \approx 0.2$ W·m^{-1}·K^{-1}, density $\approx$ 800 kg·m^{-3}, sensible $c_p \approx 2000$ J·kg^{-1}·K^{-1}, latent heat around 180–220 J·kg^{-1}·K^{-1}, and a melting interval of 45°C–55°C; this temperature window is deliberately matched to DHW preheat needs and to typical PV back-surface operating temperatures in Tabuk, so the PCM absorbs transients near noon, caps PV temperatures, and later releases heat when irradiance drops—flattening the thermal profile and improving effective electrical yield. Behind the active stack, 20 mm of closed-cell polyurethane foam ($\lambda \approx 0.03$ W·m^{-1}·K^{-1}) is used to suppress back losses and to limit parasitic heat flow into the wall, which is represented as a 150-mm concrete layer ($\lambda \approx 1.4$ W·m^{-1}·K^{-1}, $c_p \approx 880$ J·kg^{-1}·K^{-11}) typical of Saudi residential façades; including the wall explicitly is essential in BIPVT because it governs the coupling between the collector and the building's thermal mass and thus the façade heat-gain penalty/benefit. These parameter choices balance realism (market-available materials), climatic suitability

Table 2 Thermophysical properties.

Layer/material	Thickness δ (m)	Conductivity λ ($W{\cdot}m^{-1}{\cdot}K^{-1}$)	Density ρ ($kg{\cdot}m^{-3}$)	Heat capacity c_p ($J{\cdot}kg^{-1}{\cdot}K^{-1}$)	Special property/note
Glass cover (low iron)	0.004	1.05	2500	750	Transmissivity $\tau_g \approx 0.90$
PV laminate (Si + EVA/backsheet)	0.003	0.20	2330	678	$\eta_{STC} \approx 0.18$; $\beta \approx 0.004\ K^{-1}$
Absorber plate (Al)	0.001	205	2700	897	High in-plane conductivity
Fluid tube (C_u)	0.001	385	8900	385	Serpentine, bonded to absorber
Water (working fluid)	–	–	998 (at 25°C)	4180	Baseline; temperature-dependent c_p in code
PCM (paraffin class)	0.010	0.20	800	2000	Latent heat $L \approx 180$–$220\ kJ{\cdot}kg^{-1}$; $T_m \approx 45°C$–$55°C$
Back insulation (PU foam)	0.020	0.030	30	1400	Closed cell, moisture protected
Building wall (normal-weight concrete)	0.150	1.40	2300	880	Represents façade coupling

(high-irradiance, hot-arid Tabuk), and model identifiability; they also provide sensible defaults for sensitivity analysis on δ, λ, and PCM melting range to quantify design trade-offs in the technoeconomic results.

For the glass cover, the absorbed portion of incident irradiance is balanced by convective heat loss to ambient air and conduction to the PV surface:

$$I\alpha_g = h_{g-a}\left(T_g - T_{air}\right) + U_{PV-g}\left(T_g - T_{PV}\right) \tag{5}$$

where I is the incident solar irradiance [$W{\cdot}m^{-2}$], α_g is the glass absorptance, T_g is the glass temperature, T_{air} is the ambient temperature, and h_{g-a} is the external convective coefficient estimated as a function of wind speed ($h_{g-a} \approx 3v_w + 2.8$, $W{\cdot}m^{-2}{\cdot}K^{-1}$). The term U_{PV-g} denotes the conductive/convective coupling between the glass and the PV laminate.

For the PV layer, the portion of irradiance transmitted through the glass and absorbed by the PV cells is partitioned into electrical output and thermal transfer to adjacent layers (Eq. 6):

$$I\tau_g\alpha_{PV}\left(1-\eta_{el}\right) = U_{PV-g}\left(T_{PV}-T_g\right) + U_{PV-Ab}\left(T_{PV}-T_{Ab}\right) + U_{PV-t}\left(T_{PV}-T_t\right) \tag{6}$$

where τ_g is the glass transmissivity, α_{PV} is the absorptance of the PV laminate, T_{PV} is the PV cell temperature, η_{el} is the electrical efficiency, T_{Ab} is the absorber temperature, and T_t is the tube surface temperature.

The absorber-tube assembly exchanges heat with the PV laminate, the insulation, and the working fluid inside the tubes:

$$U_{PV-Ab}\left(T_{PV}-T_{Ab}\right) = U_{Ab-t}\left(T_{Ab}-T_t\right) + U_{Ab-ins}\left(T_{Ab}-T_{ins}\right) \tag{7}$$

$$U_{Ab-t}\left(T_{Ab}-T_t\right) + U_{PV-t}\left(T_{PV}-T_t\right) = U_{t-fluid}\left(T_t-T_f\right) + U_{t-ins}\left(T_t-T_{ins}\right) \tag{8}$$

where T_{ins} is the insulation temperature, T_f is the fluid bulk temperature, and the U terms represent effective thermal conductances derived from material thicknesses and thermal conductivities.

The fluid energy balance within the tubes is expressed as follows:

$$U_{t-fluid}A_c\left(T_t-T_f\right) = \dot{m}c_p\frac{dT_f}{dx} \tag{9}$$

where A_c is the tube contact area, $\dot{m}$ is the mass flow rate [$kg{\cdot}s^{-1}$], and c_p is the specific heat of water [$J{\cdot}kg^{-1}{\cdot}K^{-1}$]. This equation governs the temperature rise of the fluid as it passes through the module.

At the back insulation and wall interface, the balance reads (Eq. 10):

$$U_{Ab-ins}\left(T_{Ab}-T_{ins}\right) + U_{t-ins}\left(T_t-T_{ins}\right) = U_{ins-a}\left(T_{ins}-T_{air}\right) \tag{10}$$

which in the case of BIPVT is extended to include conductive coupling with the building wall and possible heat transfer to the indoor zone, highlighting the architectural integration effect.

The electrical efficiency of the PV cells is temperature dependent as follows:

$$\eta_{el} = \eta_{STC}\left[1-\beta\left(T_{PV}-T_{ref}\right)\right] \tag{11}$$

where η_{STC} is the reference efficiency under standard test conditions, β is the temperature coefficient [K^{-1}], and $T_{ref} = 25°C$.

The thermal efficiency is calculated as follows:

$$\eta_{th} = \frac{\dot{m}c_p\left(T_{out}-T_{in}\right)}{AI} \tag{12}$$

where T_{in} and T_{out} are the inlet and outlet fluid temperatures, and A is the collector area. A net version that accounts for electrical extraction is also considered (Eq. 13):

$$\eta_{th,net} = \frac{\dot{m}c_p\left(T_{out}-T_{in}\right)}{AI-P_{el}}, P_{el} = \eta_{el}AI \tag{13}$$

The total first-law efficiency is defined as the sum of thermal and electrical efficiencies (Eq. 14):

$$\eta_T = \eta_{th} + \eta_{el} \tag{14}$$

Finally, the exergy efficiency was used to account for the quality of the electrical and thermal energy streams. The input solar exergy was calculated using the Petela formulation, while the useful output exergy was defined as the sum of electrical exergy and the thermal exergy of the heated water stream. The exergy efficiency is therefore expressed as Eq. (15):

$$\eta_{ex} = \frac{\dot{E}_{x,out}}{\dot{E}_{x,in}} = \frac{P_{el} + \dot{Q}_u\left(1 - \frac{T_{air}}{T_{f,av}}\right)}{AI\left[1 - \frac{4}{3}\left(\frac{T_{\mathrm{air}}}{T_s}\right) + \frac{1}{3}\left(\frac{T_{\mathrm{air}}}{T_s}\right)^4\right]} \tag{15}$$

where η_{ex} is the exergy efficiency, $\dot{E}_{x,in}$ is the solar exergy input, and $\dot{E}_{x,out}$ is the useful exergy output. P_{el} is the electrical power output, which is considered pure work and therefore fully convertible to exergy. $\dot{Q}_u$ is the useful thermal energy transferred to the water loop, and $T_{f,av}$ is the average fluid temperature, calculated as $(T_{in} + T_{out})/2$. T_{air} is the ambient temperature in Kelvin, T_s is the effective sun temperature, taken as $\sim$5777 K, A is the collector area, and I is the incident solar irradiance. The term in brackets represents the Petela factor, which converts incident solar radiation into available solar exergy. This formulation allows the model to distinguish between high-quality electrical output and lower-grade thermal output [43].

2.4. Machine learning framework

Reliable day-ahead hourly inputs are essential for operational BIPVT simulation and technoeconomic appraisal. To balance interpretability with predictive power, two complementary regression models were adopted. Accordingly, two complementary models were used: a modern linear baseline, ENR, and a state-of-the-art nonlinear ensemble model, XGBoost. EN retains the clarity of linear models while adding *L1/L2* regularization for feature selection and multicollinearity control, crucial when hourly lags, calendar terms, and engineered predictors are correlated [44, 45]. XGBoost learns complex nonlinear interactions and threshold effects that are typical of solar irradiance under intermittent cloud and wind-driven convection, yet remains computationally efficient on hourly series [46, 47].

This study frames forecasting as a direct multihorizon regression problem for each target variable $y \in \{GHI', T'_{air}, v'_w\}$. Using information available up to time t, the model predicts the next 24 hourly values $\{y_{t+1}, \ldots, y_{t+24}\}$. To improve robustness, separate horizon-specific models were trained instead of using recursive one-step forecasts, thereby avoiding error accumulation (1 per lead $h = 1, \ldots, 24$). Model development used the quality-controlled hourly climate series over Tabuk, including screened satellite-derived *GHI* and station-adjusted temperature and wind-speed data. The 2024 hourly dataset contained 8784-time steps because 2024 was a leap year. After constructing lagged predictors and 24-hour-ahead target horizons, the usable sample size was 8736 hourly records. The final feature vector included 42 predictors for each forecast horizon, consisting of autoregressive lags, cross-variable lags, rolling statistics, cyclical time encodings, calendar variables, and CSI-related predictors. Separate models were trained for each target variable and each forecast horizon, resulting in 3×24 fitted models for each algorithm. A chronological split was used to preserve temporal order: the first 70% of the usable samples were used for training, the next 15% for validation and hyperparameter tuning, and the final 15% for holdout testing. This corresponded approximately to 6115 training samples, 1310 validation samples, and 1311 testing samples for each target variable and forecast horizon. The same split was applied to ENR and XGBoost to ensure a fair comparison.

Because nighttime GHI' is (near) 0, the loss for irradiance is computed on daylight hours determined by a clear-sky threshold, and predictions are constrained to nonnegative values. For the EN, the estimator for horizon h minimizes a regularized least-squares objective on standardized features (Eq. 16).

$$\min_{\beta_h,\beta_{0h}} \frac{1}{2n}\sum_{i=1}^{n}\left(y^{(i)}_{t_i+h} - \beta_{0h} - \mathbf{x}^{(i)\top}\beta_h\right)^2 + \lambda\left[(1-\alpha)\,\frac{\|\beta_h\|_2^2}{2} + \alpha\|\beta_h\|_1\right] \tag{16}$$

where $\alpha \in [0, 1]$ trades L_2 (Ridge) and L_1 (Lasso) penalties, $\lambda > 0$ controls shrinkage, and $x^{(i)}$ collects the engineered predictors at time t_i [48]. The corresponding linear predictor is as follows:

$$\hat{y}_{t+h} = \beta_{0h} + \mathrm{x}_t^{\top}\beta_h \tag{17}$$

Hyperparameter tuning was performed using rolling time-series validation on the training/validation portion of the 2024 dataset to avoid look-ahead bias. For ENR, the tested α values were 0.10, 0.30, 0.50, 0.70, and 0.90, while λ was searched over logarithmically spaced values from 10^{-4} to 10^{1}. For XGBoost, the tested ranges were: number of trees = 100–600, maximum tree depth = 2–6, learning rate = 0.01–0.20, subsample ratio = 0.70–1.00, column-sampling ratio = 0.70–1.00, and minimum child weight = 1–10. The final parameter set was selected based on the lowest validation RMSE, with daylight-only MAPE used as an additional criterion for GHI forecasts.

Only ENR and XGBoost were retained for the final comparison to balance interpretability, predictive strength, and computational efficiency. ENR was selected as a transparent regularized linear benchmark, while XGBoost was selected as a nonlinear ensemble model capable of capturing threshold effects and interactions among irradiance, temperature, wind speed, and time-based predictors. RF was not retained because preliminary testing showed similar but slightly weaker accuracy than XGBoost with less effective handling of horizon-dependent nonlinear interactions. LSTM was not included because the available dataset covered only 1 year, and deep recurrent models generally require longer time series to avoid overfitting and to justify their higher computational cost.

For the XGBoost regressor, predictions are an additive ensemble of regression trees [49].

$$\hat{y}_{t+h} = \sum_{k=1}^{K} f_k(\mathbf{x}_t), f_k \in \mathcal{F} \tag{18}$$

The XGBoost model was trained by minimizing an objective function that combines the prediction error with a regularization term controlling tree complexity. The training objective is expressed as follows [50]:

$$\mathcal{L} = \sum_{i=1}^{n}\ell\left(\mathrm{y}_i, \hat{y}_i\right) + \sum_{k=1}^{K}\Omega\left(f_k\right), \Omega\left(f_k\right) = \gamma J + \frac{1}{2}\lambda_{XGB}\sum_{j=1}^{J} w_j^2 \tag{19}$$

where n is the number of training samples, $\ell\left(y_i, \hat{y}_i\right)$ is the prediction loss between the observed value y_i and the predicted value $\hat{y}_i$, and K is the total number of regression trees in the ensemble. $\Omega\ (f_k)$ penalizes the complexity of each tree f_k. J is the number of leaves in a tree, w_j is the weight assigned to leaf j, γ penalizes the addition of new leaves, and λ_{XGB} controls L_2 regularization on leaf weights. The first term of Eq. (19) improves predictive accuracy, while the second term reduces overfitting by discouraging unnecessarily complex trees. In this study, this objective was optimized for each forecast horizon to generate stable day-ahead predictions of GHI', T'_{air}, and v'_w.

The feature vector x_t is shared across both models to ensure a fair comparison and is designed to encode diurnal/seasonal structure and short-term persistence while keeping inputs operationally available. It includes autoregressive lags of each target $\{y_t, y_{t-1}, y_{t-2}, y_{t-3}, y_{t-6}, y_{t-12}, y_{t-24}\}$; cross-predictors between variables (e.g. using recent T'_{air} and v'_w when forecasting GHI', and vice-versa); rolling statistics (3- and 6-h means and standard deviations to capture mesoscale variability); diurnal and annual cycles encoded by cyclical terms sin(2π hour/24), cos(2π hour/24), and sin(2π doy/365), cos(2π doy/365); a clear-sky reference GHI_{cs} and the CSI $k^* = GHI'/GHI_{cs}$ (and their lags), which normalize seasonal elevation of the sun; and calendar flags (weekday/weekend, month). All continuous predictors for EN are standardized (zero mean, unit variance) using statistics computed on the training fold; XGBoost ingests raw-scale features, which preserves monotonic relationships and avoids unnecessary scaling. Target vectors are the 24 hourly values of GHI', T'_{air}, and v'_w for the next day; it is trained separate models per variable and per horizon, yielding 3×24 fitted models for each algorithm. For GHI, model performance was evaluated against the screened satellite-derived irradiance series rather than ground-observed irradiance, due to the absence of continuous local GHI measurements. During operational inference, the predicted climate sequences were assembled and then fed into the BIPVT simulator to estimate next-day electrical and thermal outputs, as well as the derived technoeconomic indicators $\{\widehat{GHI}_{t+1:t+24}, \widehat{T}_{air\,t+1:t+24}, \widehat{v}_{w\,t+1:t+24}\}$. Performance is summarized on the 2024 holdout using standard regression metrics. For an evaluation set $\left\{\left(y_i, \hat{y}_i\right)\right\}_{i=1}^{N}$ [51, 52].

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2} \tag{20}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|y_i - \hat{y}_i\right| \tag{21}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2} \tag{22}$$

No separate dimensionality-reduction method such as principal component analysis was applied before model training. This decision was made because the selected predictors were physically interpretable and directly related to the diurnal, seasonal, and persistence behavior of the climate variables. For ENR, feature shrinkage and implicit feature selection were handled through the L1/L2 regularization terms. For XGBoost, feature relevance was handled internally through tree splitting and gain-based feature usage.

Table 3 Characteristics of the residential building.

Parameter	**Value**	**Notes**
Plan shape	Rectangular	Typical family housing
Number of stories	3	Predominant style
Roof area	400 m^2	Benchmark PV area
Floor-to-floor height	4 m	Total ≈12 m
Small façade area	192 m^2	Gross area
Windows	30% of façade	Public Authority for Housing standards

Because GHI' contains zeros at night, percentage errors can be ill posed; therefore, it was reported a daylight-only MAPE (Eq. 23) [53].

$$\mathrm{MAPE}_{\mathrm{day}} = \frac{100}{N_d}\sum_{i\in\mathcal{D}}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{23}$$

where D indexes daylight hours and $N_d = \widehat{|D|}$. Physical consistency was assessed during post-processing hoc $GHI \geq 0$: and clipping of extreme outliers beyond climatological bounds for T'_{air} and v'_w.

2.5. Building characteristics and EnergyPlus validation

In this study, a three-storey residential building representing the predominant construction style in Tabuk, Saudi Arabia, is employed. The building has a rectangular plan, a floor-to-floor height of 4 m, and a roof area of 400 m^2. The façades are distributed uniformly with windows occupying 30% of façade area. Table 3 summarizes the main geometrical data.

PV modules (Canadian Solar CS3U-380MS, 380 W, efficiency ≈ 19.1%) were mounted on roof and façades excluding windows. It simulated east, west, and south façades (tilt = 90°). Configurations are detailed in Table 4.

In the present work, these specifications are mirrored in EnergyPlus to provide a validation platform. It constructed an EnergyPlus model (v9.6.0) of a three-storey residential block with equivalent geometry and façade coverage. PV modules were represented with the object PV performance; PVWatts for the electrical side and the façade-mounted thermal coupling were simulated with solar collector, flat plate, and PV thermal. The model was forced with the quality-controlled Tabuk weather data in EnergyPlus weather file (EPW) format, including screened satellite-derived irradiance and station-adjusted temperature and wind-speed inputs, consistent with the ML-driven climate series.

2.6. Technoeconomic evaluation

For the technoeconomic assessment of the façade-integrated BIPVT system, it adopted the life-cycle cost analysis (LCCA) framework used as illustrated in Table 5.

In this study, the analysis covered a 25-yr horizon (matching PV module warranty), considered annual degradation of 0.85%, and

Table 4 PV configurations from the reference study.

Installation	Area (m^2)	Total modules	Strings	Tilt	Orientation
Roof	119	60 (15 × 4)	4 parallel strings	27°	South
Façade	135	68 (17 × 4)	4 parallel strings	90°	East, South, West

Table 5 Economic assumptions adapted for the current study.

Parameter	Value (baseline)	Notes
Analysis horizon	25 yrs	Module warranty
Degradation rate	0.85% per year	From ref. study
CAPEX (roof)	1000 USD·kWp^{-1}	Lower installation complexity
CAPEX (façade)	1400 USD·kWp^{-1}	Higher fixation/labor
Operation and maintenance	50 USD·kWp^{-1}·yr^{-1}	Cleaning, service
Discount rate	3% (interest)–2% (inflation)	Net ∼1%
Electricity tariff	0.133 USD·kWh^{-1}	To be adapted for Tabuk

distinguished between capital expenditure (CAPEX) and operation and maintenance costs. The LCCA requires identification of all costs over the lifespan of the modules, expressed as follows:

$$C = C_{\text{inst}} + C_{\text{operation and maintenance}} \tag{24}$$

where C_{inst} is the initial installation cost [USD] and $C_{\text{operation and maintenance}}$ is the annual maintenance cost [USD·kWp^{-1}]. To bring future operation and maintenance to present terms, the uniform present value (UPV) factor is applied:

$$UPV = \frac{1-(1+d)^{-n}}{d} \tag{25}$$

$$PV = C \times UPV \tag{26}$$

where d is the discount rate and n is the analysis period (25 years).

The installation costs for façade-mounted PV were taken as 1400 USD·kWp^{-1}, reflecting additional fixation and labor at height, while roof-mounted systems were 1000 USD·kWp^{-1}. Annual operation and maintenance costs were assumed at 50 USD·kWp^{-1}·yr^{-1}. For Saudi Arabia, these figures are scaled in sensitivity scenarios, but the baseline analysis preserves the façade–roof differential. The avoided conventional electricity cost is the annual energy generated minus degradation, multiplied by the unit cost of conventional electricity. In the Tabuk, Saudi Arabia case, this was valued at 0.133 USD·kWp^{-1}. ROI is then computed as follows:

$$\text{ROI} = \frac{\text{FV}}{\text{LCC}} \times 100\% \tag{27}$$

where FV is the invested savings from avoided electricity and LCC is the life-cycle cost. Payback period (PP) is determined as the year when the cumulative avoided cost equals the total investment.

3. Results and discussion

3.1. Climate-data validation and irradiance screening

Accurate climate inputs are essential for robust BIPVT simulation. In this study, station-based validation was performed for the variables available from the Tabuk Regional Airport meteorological station, namely, ambient temperature and wind speed. Because continuous ground-based GHI measurements were not available from this station, irradiance was not treated as a ground-validated variable. Instead, NASA POWER and CAMS irradiance products were compared with each other and screened using CSI-based physical plausibility checks. Accordingly, Fig. 5 summarizes two complementary steps: station-based validation and bias adjustment for temperature and wind speed, and interproduct/physical screening for GHI.

The station-based comparison showed that bias adjustment improved the agreement of ambient temperature and wind speed with the Tabuk Regional Airport records. To make this improvement transparent, Table 6 reports the pre- and postadjustment validation metrics for the two variables for which ground-station observations were available. The metrics were calculated using temporally matched hourly station–reanalysis pairs after the quality-control procedure described in Section 2.1.

As shown in Table 6, the MSS correction reduced the temperature RMSE from 3.10°C to 1.72°C and the MAE from 2.42°C to 1.31°C, while increasing R^2 from 0.89 to 0.96. This indicates that most of the temperature bias was associated with systematic mean and variance differences between the reanalysis product and the local station record. For wind speed, the Weibull-based quantile mapping reduced RMSE from 1.38 to 0.76 m·s^{-1} and MAE from 0.96 to 0.54 m·s^{-1}, while R^2 increased from 0.64 to 0.81. The larger improvement in distributional agreement for wind speed supports the use of a quantile-based method for this nonnegative and skewed variable. The reductions in RMSE and MAE were

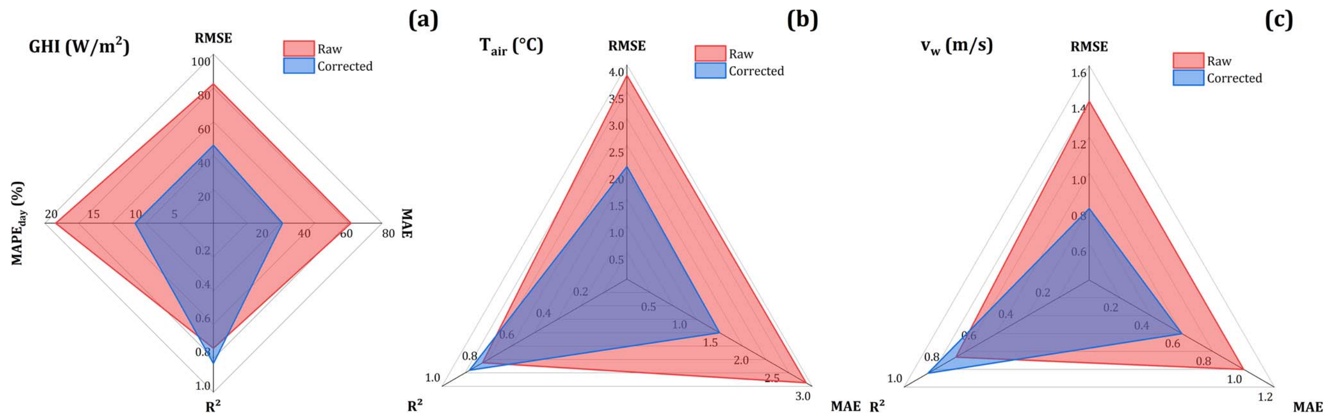

Figure 5 Climate-data evaluation for Tabuk: (a) interproduct comparison and CSI screening of GHI, (b) station-based validation of ambient temperature, and (c) station-based validation of 10-m wind speed.

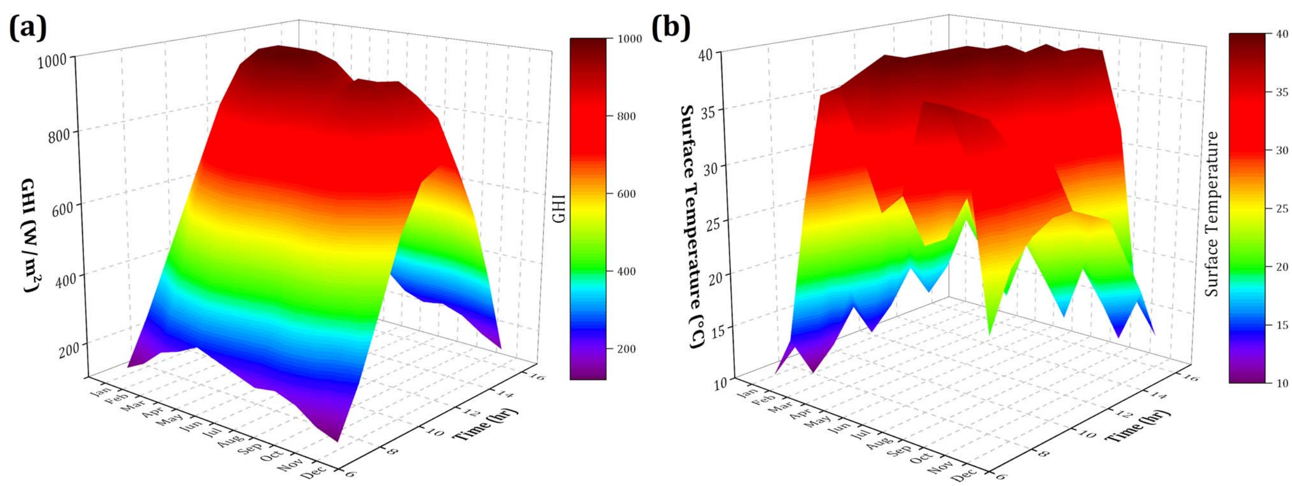

Figure 6 Annual-diurnal distribution of (**a**) screened satellite-derived GHI and (**b**) adjusted ambient temperature in Tabuk (2024).

∼44%–45%, which is consistent with the 35%–50% improvement range discussed above.

Because continuous ground-based GHI observations were not available from the Tabuk Regional Airport station, station-based pre- and postcorrection metrics were not calculated for irradiance. Therefore, GHI is not included in Table 6 as a ground-validated variable. Instead, irradiance was evaluated through NASA POWER–CAMS interproduct consistency, removal of physically unrealistic values, and daylight-only plausibility checks based on the CSI. The resulting screened GHI series was then used as the solar input for ML forecasting and BIPVT simulation.

It should be mentioned that Fig. 5 provides the visual comparison of the screened and adjusted climate series, while Table 6 reports the annual aggregated statistical metrics calculated from the matched hourly station–reanalysis pairs. Therefore, the figure and table are complementary: Fig. 5 illustrates the temporal behavior and improvement pattern, whereas Table 6 quantifies the overall pre- and postadjustment accuracy.

Following the climate-data validation, screening, and adjustment procedure, the annual and diurnal profiles of irradiance and ambient temperature were extracted as baseline features for ML. Figure 6 shows the three-dimensional distribution of screened satellite-derived GHI and adjusted ambient temperature across months and hours of the day.

The irradiance profile (Fig. 6a) highlights maximum intensities exceeding 950 $W{\cdot}m^{-2}$ around noon during the summer months (May–July), with shorter day lengths and reduced peaks in winter (December–January). This strong seasonality underscores the need for a façade-integrated solution to mitigate cooling loads while harnessing electricity and hot water production. The ambient temperature profile (Fig. 6b) shows daytime averages rising above 40°C in summer and remaining above 25°C for most of the year. This thermal environment accelerates PV cell heating and could degrade electrical efficiency if unmanaged. The integration of a PCM layer in the BIPVT façade is thus justified, as it stabilizes module temperature and enhances both thermal and electrical yields. Together, these climatic patterns explain the motivation for deploying ML-based day-ahead forecasts: the strong daily-seasonal cycles coupled with abrupt variations (clouds, wind) require predictive tools beyond static averages. Moreover, they emphasize the suitability of Tabuk's hot-arid climate for façade-integrated BIPVT, where both high-irradiance and high-cooling demand coexist.

3.2. Machine learning forecasting

To evaluate the predictive capacity of the proposed ML framework, day-ahead hourly forecasts were generated for GHI', T_{air}', and v_w'

Table 6 Pre- and postadjustment validation metrics for station-validated meteorological variables in Tabuk during 2024.

Variable	Reference station variable	Satellite/reanalysis input	Adjustment method	RMSE before	RMSE after	MAE before	MAE after	R^2 before	R^2 after	RMSE reduction
T_{air} (°C)	Station air temperature	ERA5 Tair	MSS	3.10	1.72	2.42	1.31	0.89	0.96	44.5%
v_w ($m \cdot s^{-1}$)	Station wind speed	ERA5 10-m wind speed	Weibull-based quantile mapping	1.38	0.76	0.96	0.54	0.64	0.81	44.9%

in Tabuk. The models were trained on the quality-controlled 2024 climate dataset, in which GHI' represents the screened satellite-derived irradiance series, while T_{air}' and v_w' represent station-adjusted meteorological variables. Two fundamentally different approaches were compared: ENR as a linear baseline and XGBoost as a nonlinear ensemble method (Fig. 7).

As Fig. 7 shows, for the screened satellite-derived irradiance series, XGBoost achieved an RMSE of 41.2 $W \cdot m^{-2}$ compared to 58.7 $W \cdot m^{-2}$ for ENR, and a daylight-only MAPE of 8.6% versus 12.4%. These metrics indicate the ability of the ML models to reproduce the screened satellite-based GHI target, rather than an independently measured ground GHI series. For temperature, both models performed comparably, with RMSE $\approx$ 1.9°C–2.3°C and R^2 exceeding 0.94. Wind speed forecasts remained more challenging, with XGBoost outperforming ENR (RMSE = 0.74 vs 1.12 $m \cdot s^{-1}$). Given its superior performance, XGBoost was selected as the final forecasting model, and the predicted hourly profiles of irradiance and temperature were extracted for subsequent BIPVT simulation. Figure 8 shows the comparison between reference and forecasted values for screened satellite-derived GHI and adjusted ambient temperature.

In Fig. 8a, the model successfully reproduces the diurnal bell-shaped curve of the screened satellite-derived solar irradiance, with good agreement at both the morning ramp-up and evening decline. Slight underestimation is observed near peak irradiance around noon, but the error remains within 10%. This confirms the model's suitability for capturing solar resource dynamics essential for façade-integrated BIPVT applications. In Fig. 8b, the forecasted ambient temperature follows the reference pattern closely, with small deviations throughout the day. The model correctly reproduces the gradual increase from morning to noon and the afternoon cooling, which are critical for estimating PV cell temperature and thermal output. Together, these results highlight that a robust nonlinear ML framework (XGBoost) can provide accurate day-ahead forecasts of the climate drivers most relevant to BIPVT performance. This improves the reliability of subsequent technoeconomic analysis by ensuring realistic hourly inputs for both electrical and thermal simulations.

3.3. Technical performance of building-integrated photovoltaic-thermal

The façade-integrated BIPVT model was driven with both the quality-controlled reference climate inputs and the day-ahead forecasts generated by the XGBoost model to evaluate its thermoelectrical performance in Tabuk.

To isolate the role of PCM, the BIPVT façade was simulated under two configurations: a reference case without PCM and a PCM-assisted case using the paraffin layer described in Section 2.3. The comparison was performed under the same climate inputs, façade orientation, and water-flow conditions. Table 7 summarizes the direct PCM versus non-PCM comparison under representative peak-summer operating conditions.

As shown in Table 7, PCM integration reduced the peak PV cell temperature by 16.7°C and the mean daytime PV temperature by 14.7°C under peak-summer conditions. This temperature reduction explains the improvement in electrical efficiency because PV conversion efficiency decreases with increasing cell temperature. The PCM layer absorbs part of the heat transferred from

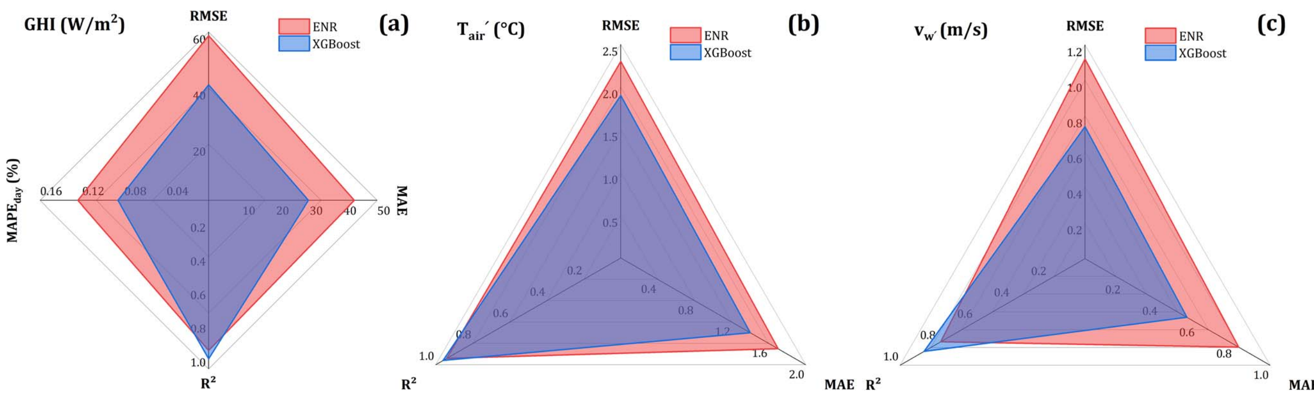

Figure 7 Performance of ML models: (**a**) ENR and (**b**) XGBoost for day-ahead forecasts in Tabuk.

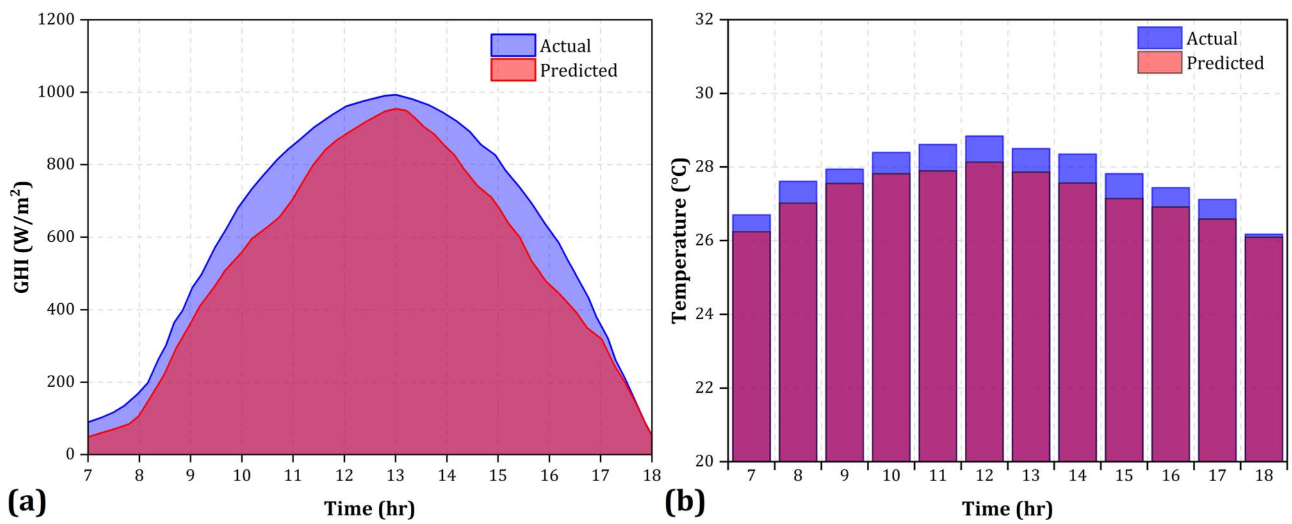

Figure 8 Comparison of screened/reference and forecasted hourly values for (**a**) satellite-derived GHI and (**b**) ambient temperature for Tabuk during daylight hours.

the absorber during high-irradiance hours through latent heat storage, delaying the temperature rise of the PV laminate. When irradiance decreases later in the day, the stored heat is gradually released to the absorber–water loop, which helps stabilize the useful thermal output.

A sensitivity check was also performed for key design parameters, including PCM thickness, water mass flow rate, and PV reference efficiency. Table 8 summarizes the effect of these parameters on PV temperature and total efficiency.

Table 7 summarizes the direct PCM versus non-PCM comparison under representative peak-summer operating conditions at the selected reference flow setting. This comparison isolates the effect of PCM integration, whereas the subsequent Fig. 9 evaluates the separate sensitivity of system performance to water mass flow rate. In the following, Fig. 9 presents the influence of water mass flow rate on four coupled performance indicators of the BIPVT façade: PV cell temperature, electrical efficiency, thermal efficiency, and total first-law efficiency. The figure compares the reference-input simulation with the XGBoost-driven simulation to evaluate whether day-ahead climate forecasts can reproduce the same operational trends.

Figure 9a shows that increasing the water mass flow rate from 0.02 to 0.10 kg·s^{-1} reduced the PV cell temperature from ∼52°C to 33°C. This occurs because a higher flow rate increases the heat-removal capacity of the water loop and reduces heat accumulation behind the PV laminate. Since PV electrical efficiency is strongly temperature dependent, this cooling effect directly improves electrical conversion performance. The difference between the reference-input and XGBoost-driven curves remains below 1°C, indicating that the forecasted climate inputs reproduce the same cooling trend.

Figure 9b shows the corresponding increase in electrical efficiency from ∼9% at the lowest flow rate to nearly 16% at the highest flow rate. This trend follows the PV temperature coefficient: as the water loop removes more heat from the absorber-PV assembly, the PV cell temperature decreases and electrical efficiency increases. The XGBoost-driven curve closely follows the reference-input curve, with deviations below 0.5% points, confirming that ML-based day-ahead climate forecasts can support reliable estimation of electrical output.

Figure 9c shows that thermal efficiency decreases from ∼44% at 0.02 kg·s^{-1} to ∼28% at 0.10 kg·s^{-1}. This does not mean that heat removal becomes ineffective. Rather, at low flow rates, the water remains longer inside the absorber tubes and experiences a larger temperature rise, which increases the apparent useful thermal gain. At higher flow rates, heat is removed more rapidly, but the outlet temperature rise per unit mass becomes smaller; therefore, the calculated thermal efficiency decreases even though PV

Table 7 Direct comparison between BIPVT configurations with and without PCM under representative peak-summer conditions.

Performance indicator	Without PCM	With PCM	Change due to PCM
Peak PV cell temperature (°C)	68.5	51.8	−16.7
Mean daytime PV cell temperature (°C)	55.6	40.9	−14.7
Mean electrical efficiency (%)	10.2	14.8	+4.6 percentage points
Mean thermal efficiency (%)	34.5	38.9	+4.4 percentage points
Total efficiency (%)	44.7	53.7	+9.0 percentage points
Daily useful thermal output (kWh·day^{-1})	22.8	25.6	+12.3%

Table 8 Sensitivity of BIPVT performance to selected design parameters.

Parameter	Tested range	Main effect on system performance
PCM thickness	5–15 mm	Increasing PCM thickness reduced peak PV temperature by ∼7°C–19°C, but gains became smaller beyond 10 mm because of added thermal resistance.
Water mass flow rate	0.02–0.10 kg·s^{-1}	Higher flow rate reduced PV temperature from ∼52°C to 33°C and increased electrical efficiency from ∼9% to 16%, but reduced outlet temperature rises and thermal efficiency.
PV reference efficiency	17%–21%	Higher PV efficiency increased electrical yield almost proportionally, while slightly reducing the fraction of absorbed energy available for thermal recovery.

cooling improves. This confirms the trade-off between maximizing thermal output temperature and maximizing PV cooling.

Figure 9d combines the electrical and thermal contributions and shows that the total efficiency remains within ∼42%–55%, with the best overall performance occurring at intermediate mass flow rates around 0.05–0.06 kg·s^{-1}. At very low flow rates, electrical efficiency is penalized by high PV temperature, whereas at very high flow rates, thermal efficiency decreases because the outlet temperature rise becomes smaller. Therefore, the intermediate region represents the best compromise between electrical gain from PV cooling and useful heat recovery in the water loop. The close agreement between the reference-input and XGBoost-driven curves across all subplots confirms that forecast-driven simulations preserve the main thermoelectrical behavior of the BIPVT façade.

The close agreement between the reference-input and XGBoost-driven performance curves indicates that the forecasting errors were not large enough to substantially alter the simulated operating trends. However, some uncertainty remains because forecast errors in GHI, ambient temperature, and wind speed propagate into PV temperature, thermal output, and efficiency estimates. In this study, uncertainty propagation was evaluated by comparing the variability in BIPVT simulation outputs obtained using reference inputs with those generated using ML-based predictions, rather than through a fully probabilistic uncertainty quantification framework such as Monte Carlo simulation. The resulting differences were small: PV temperature deviations were generally below 1°C, electrical-efficiency deviations were below 0.5 percentage points, and annual energy-yield differences remained below ∼3%–4%. Therefore, the smoothness and close alignment of the curves reflect the strong deterministic diurnal structure of the climate inputs and the relatively low forecast error of XGBoost, but the results should still be interpreted as deterministic scenario outputs rather than probabilistic uncertainty bounds.

3.4. Cross-validation with EnergyPlus

To confirm the robustness of the proposed BIPVT façade model and its coupling with ML-based climate drivers, an independent simulation was performed using EnergyPlus (v9.6.0). The residential building configuration used for validation consisted of a three-storey rectangular block with a gross façade area of 327 m^2 and 30% window fraction. PV modules were assigned to the opaque façade surfaces using the PV Performance: PVWatts object, while the thermal coupling to a water loop for DHW was represented with the Solar Collector: Flat Plate: PV Thermal object. The same hourly climate inputs, including the quality-controlled reference series and the ML-predicted series for 2024, were provided in EPW format to ensure consistent forcing across the two modeling platforms. To ensure comparability between the in-house model and EnergyPlus, the same façade area, PV module efficiency, tilt angle, orientation, inlet water temperature, mass flow rate, and hourly climate inputs were used in both models. No empirical calibration factor was applied to force agreement between the two models. The only harmonization steps involved matching the physical configuration, using equivalent PV and PVT object settings, and applying the same annual operating schedule. Therefore, the validation tested whether the independently implemented EnergyPlus model could reproduce the annual and temporal trends predicted by the layer-resolved in-house model under equivalent boundary conditions. Table 9 compares the annual energy outputs from the in-house BIPVT model and EnergyPlus for both reference-input and ML-driven simulations. The metrics include annual AC electricity (MWh), useful thermal energy delivered to the DHW tank (MWh$_{th}$), and combined efficiency (%).

The results show strong agreement between the two models. For the reference-input case, the difference in annual electrical yield is less than 4% and thermal yield differs by 3%. Similar deviations were observed in the ML-driven case, with EnergyPlus

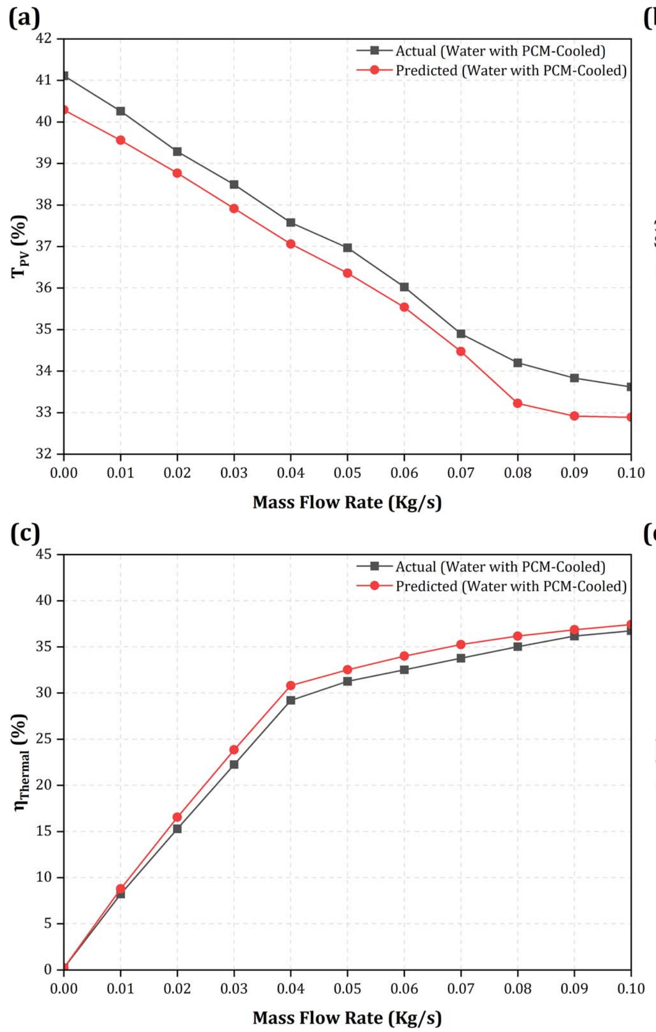

Figure 9 Effect of water mass flow rate on BIPVT performance under reference-input and XGBoost-driven simulations: (**a**) PV cell temperature, (**b**) electrical efficiency, (**c**) thermal efficiency, and (**d**) total first-law efficiency.

Table 9 Comparison of BIPVT results between in-house model and EnergyPlus.

Input type	Model	Electrical yield (MWh)	Thermal yield (MWh_{th})	Total efficiency (%)
Reference	In-house BIPVT	5.6	9.3	51.1
	EnergyPlus	5.4	9.0	50.2
ML	In-house BIPVT	5.5	9.2	50.8
	EnergyPlus	5.3	8.8	49.7

slightly underpredicting both electrical and thermal outputs relative to the in-house model. These discrepancies can be attributed to simplified electrical submodels and default heat transfer correlations in EnergyPlus, whereas the in-house model applies layer-resolved energy balances tailored for façade integration.

To further examine temporal agreement, seasonal comparisons were also calculated for the reference-input case. Table 10 summarizes the seasonal deviations between the in-house BIPVT model and EnergyPlus.

The seasonal comparison shows that the two models remained in close agreement throughout the year. The largest deviations occurred during summer, when irradiance intensity and façade surface temperatures were highest. These differences can be attributed to differences in sky-temperature treatment, PV electrical submodels, and default heat-transfer correlations in EnergyPlus, whereas the in-house model applies layer-resolved energy balances tailored for façade-integrated BIPVT operation.

3.5. Comparison with previous studies

To further verify the efficiency and relevance of the proposed framework, the present results were compared with recent studies on PCM-assisted façades, BIPV/T-PCM systems, ML-supported

Table 10 Seasonal agreement between the in-house BIPVT model and EnergyPlus for the reference-input case.

Season	Electrical-yield deviation (%)	Thermal-yield deviation (%)
Winter	2.1	2.4
Spring	3.3	3.6
Summer	4.8	5.2
Autumn	3.5	3.9

performance prediction, and technoeconomic BIPV evaluation. Table 11 summarizes the comparison.

The comparison shows that the present results are consistent with previous studies in three main aspects. First, the observed PCM-induced PV temperature reduction agrees with recent BIPV/BIPV/T-PCM studies, where PCM was shown to reduce operating temperature and improve electrical performance. For example, Mangherini *et al.* reported a PV-window temperature reduction from 42°C to 36°C, while Alsagri and Alrobaian reported a PV temperature reduction of up to 25.47°C in a Saudi BIPV/T-PCM system. The 15°C–20°C reduction obtained in the present study therefore falls within the expected range for PCM-assisted PV thermal regulation, especially under high irradiance.

Second, the thermal-efficiency range obtained in this study is consistent with previous BIPV/T-PCM findings. Alsagri and Alrobaian reported thermal efficiencies of 29.65%–59.40%, while the present model produced total efficiency values of ∼42%–55% depending on the water mass flow rate. This agreement supports the physical reliability of the layer-resolved BIPVT model.

Third, the present study extends previous work by integrating several components that are usually treated separately. Previous studies have generally focused on passive PCM façade energy savings, BIPV/BIPVT component performance, PCM thermal regulation, ML prediction, or technoeconomic evaluation as separate tasks. In contrast, the proposed framework combines climate-data screening, day-ahead ML forecasting, PCM-stabilized water-based BIPVT simulation, EnergyPlus cross-validation, façade-orientation comparison, and 25-year economic assessment. Therefore, the efficiency of the proposed method is supported by the numerical results and by its broader operational and technoeconomic coverage.

3.6. Technoeconomic evaluation

To complement the technical results, an economic assessment was carried out based on a 25-year LCCA for the façade-integrated BIPVT system in Tabuk. The analysis considered both capital expenditure (CAPEX) and operation and maintenance costs, discounted to present value using the UPV factor. The results are presented separately for the east-, west-, and south-facing façades of the reference residential building. Table 12 presents the baseline LCC of the façade-integrated BIPVT installation used for the economic comparison.

The façade installation has higher initial and operation and maintenance costs compared to roof-mounted PV due to additional structural and integration requirements. Nevertheless, the LCC remains within a competitive range for long-term investment in Tabuk, where solar potential is high and electricity tariffs are gradually increasing. The energy performance of the BIPVT façades over the 25-yr horizon is summarized in Table 13.

It should be mentioned that the invested savings reported in Table 13 represent the accumulated avoided-cost benefits under the reinvestment scenario and therefore are not calculated as a simple multiplication of net electrical generation by the baseline electricity tariff. East- and west-facing façades produced significantly higher net power generation compared to the south-facing façade, which is consistent with the daily solar trajectory in Tabuk, where morning and afternoon irradiance is strong but the midday sun strikes the vertical south façade less effectively. This translated into larger monetary savings for the east and west façades compared to the south façade. The long-term financial indicators are visualized in Fig. 10.

The PP was calculated as the first year in which the cumulative discounted avoided electricity cost became equal to or greater than the LCC of the façade installation. ROI was calculated as the ratio between the net discounted benefit and the total LCC over the 25-yr analysis period. In the reinvestment scenario shown in Fig. 10a, annual savings were assumed to be reinvested using the same net discount framework applied in the LCCA, whereas Fig. 10b represents cumulative savings without reinvestment. Therefore, the two curves represent different financial interpretations of the same annual energy-saving stream.

The higher performance of the east and west façades compared with the south façade is explained by the solar geometry of vertical surfaces in Tabuk. For vertical façades, the south-facing plane receives reduced effective irradiance around solar noon when the sun is high in the sky, especially during summer. In contrast, east- and west-facing façades receive stronger morning and afternoon irradiance at lower solar-altitude angles, which better match the incidence angle of vertical surfaces. Since Tabuk experiences strong morning and afternoon solar availability during much of the year, east and west façades can generate higher annual useful energy than the vertical south façade, despite the general expectation that south-facing tilted surfaces are optimal for rooftop PV.

The economic results are sensitive to electricity tariffs, installation cost, and policy incentives. A higher electricity tariff or the presence of feed-in tariffs, net-metering credits, or capital subsidies would shorten the PP and increase ROI. Conversely, higher façade installation costs or the absence of policy support would delay payback. A ± 20% sensitivity check showed that increasing the electricity tariff by 20% reduced the east/west PP from ∼9–10 yrs to ∼7–8 yrs, while a 20% reduction in tariff increased it to ∼11–12 yrs. Similarly, a 20% reduction in façade CAPEX reduced the PP by ∼1.5–2 yrs, whereas a 20% CAPEX increase delayed payback by ∼2 yrs.

Table 11 Comparison of the present study with recent related works.

Study	System/method	Climate/location	Main reported results	Comparison with the present study
Alqaed [30]	Simple façade, DSF, and DSF with PCM	Jeddah, Abha, and Tabuk, Saudi Arabia	In Tabuk, DSF with PCM reduced heating energy by 40% and cooling energy by 25% compared with a simple façade.	Confirms the suitability of PCM-assisted façade strategies for Tabuk, but focused on passive façades and did not quantify active electricity generation, thermal recovery, ML forecasting, or life-cycle economics.
Mangherini *et al.* [14]	Semitransparent BIPV window with paraffin PCM	Building-integrated PV-window system	PCM reduced nominal operating cell temperature from 42°C to 36°C and increased PV-window power output by 11.80%.	Supports the role of PCM in PV thermal regulation. The present study extends this concept to an active water-based BIPVT façade and reports a larger PV temperature reduction of ∼15°C–20°C under hot-arid Tabuk conditions.
González Gallero *et al.* [31]	Façade-based BIPVT-PCM system for DHW	Façade-integrated DHW application	The proposed modular BIPVT-PCM system was able to provide DHW at ∼45°C for a 5-min shower demand.	Closely related to the present BIPVT-PCM concept, but did not include Saudi hot-arid boundary conditions, ML-based day-ahead forecasting, façade-orientation comparison, or life-cycle economic analysis.
Alsagri and Alrobaian [32]	BIPV/T system with and without PCM combined with RF prediction	Sharurah, Saudi Arabia	PCM reduced PV cell temperature by up to 25.47°C; thermal efficiency ranged from 29.65% to 59.40%; RF prediction achieved R^2 values around 96%–98%.	Confirms that PCM and ML prediction are effective for Saudi BIPV/T systems. The present study provides comparable thermal-efficiency behavior and ML accuracy, while additionally incorporating satellite-data screening, EnergyPlus cross-validation, façade orientation, and economic payback.
Shahverdian *et al.* [17]	Comparative technoeconomic assessment of rooftop PV, BIPV, and hybrid systems	Residential building in Iran	Rooftop PV had the lowest LCOE at 0.023 USD·kWh^{-1}, BIPV had higher LCOE at 0.077 USD·kWh^{-1}, and the hybrid system reached 16.2 MWh annual generation.	Supports the economic challenge of BIPV due to installation complexity. The present study addresses this issue for a BIPVT façade by combining electrical generation with useful thermal output and showing east/west payback of ∼9–10 yrs.
Present study	PCM-stabilized façade-integrated BIPVT with ML forecasting and EnergyPlus validation	Tabuk, Saudi Arabia	XGBoost achieved RMSE = 41.2 W·m^{-2} and daylight MAPE = 8.6% for screened GHI forecasting; increasing the water mass flow rate from 0.02 to 0.10 kg·s^{-1} reduced PV temperature from ∼52°C to 33°C, while PCM combination reduced peak PV temperature by ∼15°C–20°C under high solar loads; electrical efficiency increased from ∼9% to nearly 16%; EnergyPlus deviations remained below 5%; east/west façades reached payback in 9–10 yrs.	Provides a combined framework linking climate-data screening, ML forecasting, layer-resolved BIPVT-PCM modeling, EnergyPlus validation, and life-cycle economic assessment for a hot-arid Saudi residential case study.

Table 12 LCC of façade configurations over 25 yrs.

Studied configuration	Initial cost (US$)	Operation and maintenance cost (US$/base date)	UPV factor	LCC (US$)
Facade	25 850	1300	22.10	54 360

Table 13 Net power generation, compensated power, and invested savings over 25 yrs.

Installation area	Net power generation (MWh)	Compensated power (MWh)	Invested savings (US$)
East	932.20	7.77	165 950
South	508.11	5.16	90 520
West	932.30	8.65	165 950

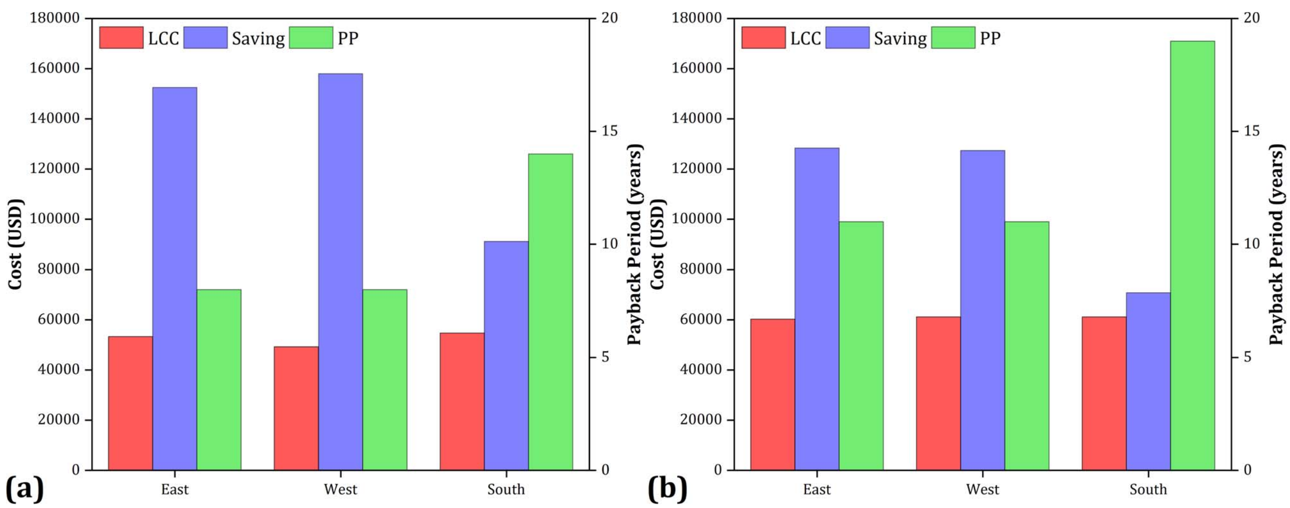

Figure 10 Economic evaluation of façade-integrated BIPVT: (**a**) with investment and (**b**) without investment.

The results highlight that, although façade-integrated PV systems require higher initial investment than conventional rooftop installations, their long-term performance in hot-arid climates such as Tabuk makes them financially attractive. The superior performance of east and west orientations confirms their suitability for vertical BIPVT applications in Saudi Arabia. The inclusion of PCM in the façade module, while not explicitly monetized in this analysis, is expected to provide additional value by stabilizing thermal performance, enhancing DHW output, and potentially reducing cooling loads in the building.

4. Conclusions

In this study, a façade-integrated BIPVT system with PCM stabilization was investigated for residential buildings in Tabuk, Saudi Arabia. The proposed framework combined quality-controlled satellite/reanalysis climate inputs, station-based validation of temperature and wind speed, CSI-based screening of satellite-derived irradiance, day-ahead ML forecasting, layer-resolved thermoelectrical BIPVT modeling, EnergyPlus cross-validation, and life-cycle economic assessment. This integrated workflow was developed to evaluate whether an ML-assisted, PCM-stabilized BIPVT façade can provide technically reliable and economically feasible performance under hot-arid and data-limited conditions. The main findings of the study can be summarized as follows:

I. The climate-data preprocessing procedure improved the reliability of the simulation inputs: station-based bias adjustment improved temperature and wind-speed inputs, while GHI was treated as a screened satellite-derived variable because continuous ground-based irradiance observations were not available for Tabuk.
II. XGBoost provided the best forecasting performance among the tested models, achieving an RMSE of 41.2 $W{\cdot}m^{-2}$ and daylight MAPE of 8.6% for screened GHI forecasting, and it also outperformed ENR for wind-speed prediction.
III. Increasing the water mass flow rate from 0.02 to 0.10 $kg{\cdot}s^{-1}$ reduced PV cell temperature from ~52°C to ~33°C and increased electrical efficiency from ~9% to nearly 16%.
IV. PCM combination reduced PV operating temperature by ~15°C–20°C under high solar loads, confirming its role as a thermal-buffering layer that absorbs excess heat during

peak irradiance and stabilizes the thermal response of the façade.

V. Cross-validation with EnergyPlus showed deviations within ∼5%, supporting the robustness of the in-house layer-resolved BIPVT model.

VI. Technoeconomic assessment showed that east- and west-facing façades achieved shorter PP of ∼9–10 yrs, while the south façade required ∼15 yrs, mainly because vertical east and west façades receive more favorable morning and afternoon irradiance under Tabuk's solar geometry.

In addition to these main findings, the results showed that the proposed ML-driven simulation workflow can reproduce the main thermoelectrical trends obtained from reference climate inputs. The close agreement between reference-input and XGBoost-driven simulations indicates that day-ahead forecasts can be used for operational planning of BIPVT façade performance. The EnergyPlus comparison further confirmed that the proposed physical model is not strongly dependent on a single implementation platform. The economic results also suggest that façade-integrated BIPVT systems can become more attractive when their dual electrical and thermal outputs are considered together rather than evaluating electricity generation alone.

The study has several limitations. Continuous ground-based GHI observations were not available for Tabuk; therefore, irradiance validation relied on satellite-product consistency and CSI plausibility screening rather than direct station-based irradiance measurements. The PCM was modeled using simplified thermophysical properties and an effective melting range, without considering long-term degradation, phase segregation, or repeated cycling effects. The economic assessment was based on fixed baseline assumptions for electricity tariff, installation cost, operation and maintenance cost, and discount rate, although these parameters may change with future market conditions and policy incentives. In addition, the analysis was simulation based and therefore requires long-term experimental monitoring for full field validation.

Future studies should address these limitations by collecting local ground-based irradiance measurements in Tabuk, experimentally monitoring façade-integrated BIPVT-PCM prototypes, and validating PCM melting/solidification behavior under real operating conditions. Further work should also investigate advanced PCM composites with higher thermal conductivity, variable-flow control strategies, hybrid electrical–thermal storage, uncertainty propagation using probabilistic methods, and broader economic scenarios including carbon pricing, demand-response participation, net metering, and policy incentives. Extending the framework to other Saudi and Middle Eastern climates would also help identify the most suitable façade orientations and control strategies under different solar and thermal conditions.

The contribution of this study is the development of an integrated, operation-oriented assessment framework for PCM-stabilized BIPVT façades in hot-arid residential buildings. Unlike previous studies that often examined climate-data processing, solar forecasting, PCM thermal buffering, BIPVT simulation, model validation, or economic feasibility separately, this work links these components within a single workflow. The results demonstrate that combining screened climate inputs, XGBoost forecasting, layer-resolved BIPVT-PCM modeling, EnergyPlus cross-validation, and life-cycle economic evaluation can provide a more reliable basis for designing and assessing low-carbon solar façades in data-limited hot-arid regions.

Acknowledgements

During the preparation of this manuscript/study, the author(s) used ChatGPT and Quillbot tools for the purposes of the article language enhancement. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Author contributions

Khaled Almazam (Data curation [equal], Formal analysis [equal], Investigation [equal], Resources [equal], Software [equal]), Eman Nasrallah (Formal analysis [equal], Investigation [equal], Validation [equal], Visualization [equal]), Faizah Mohammed Bashir (Conceptualization [equal], Investigation [equal], Methodology [equal], Software [equal], Validation [equal]), Abubakar Sadiq Mahmoud (Investigation [equal], Resources [equal], Visualization [equal], Writing—original draft [equal]), Yakubu Aminu Dodo (Conceptualization [equal], Formal analysis [equal], Supervision [equal], Writing—original draft [equal], Writing—review & editing [equal]), and Yohannes Mehari Andiye (Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Software [equal], Writing—review & editing [equal])

Conflicts of interest

The authors declare they have no conflict of interests.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Wang G, Ji X, Zou T. *et al.* Effect evaluation of frame perforation on reducing photovoltaic panel temperature with passive air cooling. *Case Stud Therm Eng* 2025;**76**:107296. https://doi.org/10.1016/j.csite.2025.107296.
2. Zhang Z, Wang S, Ye B. *et al.* A feature prediction-based method for energy consumption prediction of electric buses. *Energy* 2025;**314**:134345. https://doi.org/10.1016/j.energy.2024.134345.
3. Mata É, Wanemark J, Cheng SH. *et al.* Systematic map of determinants of buildings' energy demand and CO_2 emissions shows need for decoupling. *Environ Res Lett* 2021;**16**:055011. https://doi.org/10.1088/1748-9326/abe5d7.
4. Röck M, Saade MRM, Balouktsi M. *et al.* Embodied GHG emissions of buildings—the hidden challenge for effective climate change mitigation. *Appl Energy* 2020;**258**:114107. https://doi.org/10.1016/j.apenergy.2019.114107.

5. Aldersoni A, Ibrahim AO, Aldamady AAH. *et al.* Investigating the impact of low-carbon building materials on energy consumption and carbon emissions in construction projects. *Int J Low-Carbon Technol* 2025;**20**:1581–92. https://doi.org/10.1093/ijlct/ctaf096.
6. Guo Y, Uhde H, Wen W. Uncertainty of energy consumption and CO_2 emissions in the building sector in China. *Sustain Cities Soc* 2023;**97**:104728. https://doi.org/10.1016/j.scs.2023.104728.
7. Balaras CA. *et al.* Carbon and greenhouse gas emissions from electricity consumption in European Union buildings. *Buildings* 2023;**14**:71.
8. Atmaca N, Atmaca A, Özçetin Aİ. The impacts of restoration and reconstruction of a heritage building on life cycle energy consumption and related carbon dioxide emissions. *Energ Buildings* 2021;**253**:111507. https://doi.org/10.1016/j.enbuild.2021.111507.
9. Abdelrazik AS, Shboul B, Elwardany M. *et al.* The recent advancements in the building integrated photovoltaic/thermal (BIPV/T) systems: an updated review. *Renew Sustain Energy Rev* 2022;**170**:112988. https://doi.org/10.1016/j.rser.2022.112988.
10. Tiwari AK, Chatterjee K, Agrawal S. *et al.* A comprehensive review of photovoltaic-thermal (PVT) technology: performance evaluation and contemporary development. *Energy Rep* 2023;**10**:2655–79. https://doi.org/10.1016/j.egyr.2023.09.043.
11. Alhammadi N, Rodriguez-Ubinas E, Alzarouni S. *et al.* Building-integrated photovoltaics in hot climates: experimental study of CIGS and c-Si modules in BIPV ventilated facades. *Energ Conver Manage* 2022;**274**:116408. https://doi.org/10.1016/j.enconman.2022.116408.
12. Yu G, Yang H, Yan Z. *et al.* A review of designs and performance of façade-based building integrated photovoltaic-thermal (BIPVT) systems. *Appl Therm Eng* 2021;**182**:116081. https://doi.org/10.1016/j.applthermaleng.2020.116081.
13. Şirin C, Goggins J, Hajdukiewicz M. A review on building-integrated photovoltaic/thermal systems for green buildings. *Appl Therm Eng* 2023;**229**:120607. https://doi.org/10.1016/j.applthermaleng.2023.120607.
14. Mangherini G, Baccega E, Diolaiti V. *et al.* Integration of phase change material into PV windows to improve the efficiency of semi-transparent panels based on luminescent solar concentrator technology. *Sustainability* 2024;**16**:11148. https://doi.org/10.3390/su162411148.
15. Čurpek J, Čekon M. Climate response of a BiPV façade system enhanced with latent PCM-based thermal energy storage. *Renew Energy* 2020;**152**:368–84. https://doi.org/10.1016/j.renene.2020.01.070.
16. Edmund JWT, Go YI. Enhancement and validation of building integrated PV system: 3D modelling, techno-economics and environmental assessment. *Energy Built Environ* 2022;**3**:444–66. https://doi.org/10.1016/j.enbenv.2021.05.001.
17. Shahverdian MH, Najaftomaraei M, Fassadi Chimeh A. *et al.* Towards zero-energy buildings: a comparative techno-economic and environmental analysis of rooftop PV and BIPV systems. *Buildings* 2025;**15**:999. https://doi.org/10.3390/buildings15070999.
18. Niyazi B, Hussain S, Elfeki AM. *et al.* Comparative evaluation of techniques for missing rainfall data estimation in arid regions: case study of Al-Madinah Al-Munawarah, *Saudi Arabia*. *Theor Appl Climatol* 2024;**155**:2195–214. https://doi.org/10.1007/s00704-023-04752-2.
19. Helmi AM, Farouk MI, Hassan R. *et al.* Comparing remote sensing and geostatistical techniques in filling gaps in rain gauge records and generating multi-return period isohyetal maps in arid regions—case study: Kingdom of Saudi Arabia. *Water* 2024;**16**:925. https://doi.org/10.3390/w16070925.
20. Adem E, Elfeki A, Chaabani A. *et al.* Impact of satellite precipitation estimation methods on the hydrological response: case study Wadi Nu'man basin, *Saudi Arabia*. *Theor Appl Climatol* 2024;**155**:3907–25. https://doi.org/10.1007/s00704-024-04855-4.
21. Odnoletkova N, Patzek TW. Data-driven analysis of climate change in Saudi Arabia: trends in temperature extremes and human comfort indicators. *J Appl Meteorol Climatol* 2021;**60**:1055–70. https://doi.org/10.1175/JAMC-D-20-0273.1.
22. Papacharalampous G, Tyralis H, Doulamis A. *et al.* Comparison of machine learning algorithms for merging gridded satellite and earth-observed precipitation data. *Water* 2023;**15**:634. https://doi.org/10.3390/w15040634.
23. Imam AA, Abusorrah A, Seedahmed MMA. *et al.* Accurate forecasting of global horizontal irradiance in Saudi Arabia: a comparative study of machine learning predictive models and feature selection techniques. *Mathematics* 2024;**12**:2600. https://doi.org/10.3390/math12162600.
24. Bot K, Aelenei L, Gomes MG. *et al.* Performance assessment of a building integrated photovoltaic thermal system in mediterranean climate—a numerical simulation approach. *Energies* 2020;**13**:2887. https://doi.org/10.3390/en13112887.
25. Uddin MM, Ji J, Wang C. *et al.* Building energy conservation potentials of semi-transparent CdTe integrated photovoltaic window systems in Bangladesh context. *Renew Energy* 2023;**207**:512–30. https://doi.org/10.1016/j.renene.2023.03.049.
26. Magni M, Ochs F, de Vries S. *et al.* Detailed cross comparison of building energy simulation tools results using a reference office building as a case study. *Energ Buildings* 2021;**250**:111260. https://doi.org/10.1016/j.enbuild.2021.111260.
27. Vuong E, Kamel RS, Fung AS. Modelling and simulation of BIPV/T in EnergyPlus and TRNSYS. *Energy Procedia* 2015;**78**:1883–8. https://doi.org/10.1016/j.egypro.2015.11.354.
28. Alyami SH, Ashraf N, Alyami FM. *et al.* Energy assessment of BIPV system in code-compliant residential building in cooling-dominated climates. *PLoS One* 2025;**20**:e0318786. https://doi.org/10.1371/journal.pone.0318786.
29. Al-Shammari S, Ko W, Al Ammar EA. *et al.* Optimal decision-making in photovoltaic system selection in Saudi Arabia. *Energies* 2021;**14**:357. https://doi.org/10.3390/en14020357.
30. Alqaed S. Effect of annual solar radiation on simple façade, double-skin facade and double-skin facade filled with phase change materials for saving energy. *Sustain Energy Techn and Assess* 2022;**51**:101928. https://doi.org/10.1016/j.seta.2021.101928.
31. González Gallero FJ, González Siles G, Rodríguez Maestre I. *et al.* Performance analysis of a novel façade-based building integrated photovoltaic-thermal system with phase change material for domestic hot water. *Appl Therm Eng* 2024;**240**:122321. https://doi.org/10.1016/j.applthermaleng.2023.122321.

32. Alsagri AS, Alrobaian AA. Analysis and performance prediction of a building integrated photovoltaic thermal system with and without phase change material. *Energy* 2024;**310**:133249. https://doi.org/10.1016/j.energy.2024.133249.
33. Ismaeil EMH, Sobaih AEE. Evaluating BIPV façades in a building envelope in hot districts for enhancing sustainable ranking: a Saudi Arabian perspective. *Buildings* 2023;**13**:1110. https://doi.org/10.3390/buildings13051110.
34. Awan A, Zubair M, P PR. *et al.* Solar energy resource analysis and evaluation of photovoltaic system performance in various regions of Saudi Arabia. *Sustainability* 2018;**10**:1129. https://doi.org/10.3390/su10041129.
35. Zell E, Gasim S, Wilcox S. *et al.* Assessment of solar radiation resources in Saudi Arabia. *Sol Energy* 2015;**119**:422–38. https://doi.org/10.1016/j.solener.2015.06.031.
36. Perez R, Ineichen P, Seals R. *et al.* Modeling daylight availability and irradiance components from direct and global irradiance. *Sol Energy* 1990;**44**:271–89. https://doi.org/10.1016/0038-092X(90)90055-H.
37. Ineichen P, Perez R. A new airmass independent formulation for the Linke turbidity coefficient. *Sol Energy* 2002;**73**:151–7. https://doi.org/10.1016/S0038-092X(02)00045-2.
38. Jakob Themeßl M, Gobiet A, Leuprecht A. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *Int J Climatol* 2011;**31**:1530–44.
39. Cannon AJ, Sobie SR, Murdock TQ. Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes? *J Climate* 2015;**28**:6938–59. https://doi.org/10.1175/JCLI-D-14-00754.1.
40. Li H, Sheffield J, Wood EF. Bias correction of monthly precipitation and temperature fields from intergovernmental panel on climate change AR4 models using equidistant quantile matching. *J Geophys Res Atmos* 2010;**115**:1–20. https://doi.org/10.1029/2009JD012882.
41. Maraun D. Bias correcting climate change simulations—a critical review. *Curr Clim Chang Rep* 2016;**2**:211–20. https://doi.org/10.1007/s40641-016-0050-x.
42. Justus CG, Hargraves WR, Mikhail A. *et al.* Methods for estimating wind speed frequency distributions. *J Appl Meteorol* 1978;**17**:350–3. https://doi.org/10.1175/1520-0450(1978)017<0350:MFEWSF>2.0.CO;2.
43. Diwania S, Kumar M, Kumar R. *et al.* Machine learning-based thermo-electrical performance improvement of nanofluid-cooled photovoltaic–thermal system. *Energy Environ* 2024;**35**:1793–817. https://doi.org/10.1177/0958305X221146947.
44. Hamada AA. Elastic net principal component regression with an application. *Int J Sci Math Educ* 2024;**1**:13–23. https://doi.org/10.62951/ijsme.v1i3.25.
45. Saputro DRS, Wahyu NL, Widyaningsih Y. Performance of ridge regression, least absolute shrinkage and selection operator, and elastic net in overcoming multicollinearity. *J Multidiscip Appl Nat Sci* 2025;**5**:370–82. https://doi.org/10.47352/jmans.2774-3047.251.
46. Obiora CN, Ali A, Hasan AN. 2021. Implementing extreme gradient boosting (XGBoost) algorithm in predicting solar irradiance. In *2021 IEEE PES/IAS PowerAfrica*. New Jersey: IEEE. 1–5.
47. Çarklı Yavuz B, Aksoy Tırmıkçı C. Comparative analysis of machine learning and deep learning models for solar energy forecasting: a meteorological data-driven approach. *J Math Sci Model* 2025;**8**:153–66.
48. Tay JK, Narasimhan B, Hastie T. Elastic net regularization paths for all generalized linear models. *J Stat Softw* 2023;**106**:1–31. https://doi.org/10.18637/jss.v106.i01.
49. Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms. *Artif Intell Rev* 2021;**54**:1937–67. https://doi.org/10.1007/s10462-020-09896-5.
50. Boldini D, Grisoni F, Kuhn D. *et al.* Practical guidelines for the use of gradient boosting for molecular property prediction. *J Chem* 2023;**15**:73. https://doi.org/10.1186/s13321-023-00743-7.
51. Siahaan AA, Asrol M. Development of a machine learning model for predicting hardness in the water treatment pharmaceutical industry. *Int J Ind Eng Manag* 2023;**14**:138–46. https://doi.org/10.24867/IJIEM-2023-2-329.
52. Cárdenas Escorcia Y. *et al.* Privacy-preserving federated learning for predictive maintenance in smart manufacturing networks. *Int J Ind Eng Manag* 2025;**16**:296–315.
53. Mamyrbayev O, Akhmediyarova A, Oralbekova D. *et al.* Optimizing renewable energy integration using IoT and machine learning algorithms. *Int J Ind Eng Manag* 2025;**16**:101–12.

International Journal of Low-Carbon Technologies, 2026, 21, 1–25, ctag108
https://doi.org/10.1093/ijlct/ctag108
Original Article