engineering proceedings

MDPI

Proceeding Paper

Low-Cost Hybrid Pyranometer with Machine Learning-Enhanced Solar Irradiance Sensing †

Peter Eze [1], Rejoice Chigbo [1], Jerry Edeani [1], Joshua Aniehe [1], Nebechi Obiora [1], Gerald Ugwu Akubue [2], Joshua Okechukwu [3], Howard Njoku [1,*] and Muyiwa Oyinlola [4]

[1] Sustainable Energy Engineering Research Group, Department of Mechanical Engineering, University of Nigeria, Nsukka 410001, Nigeria; chukwuemeka.eze.247825@unn.edu.ng (P.E.); chigborejoice16@gmail.com (R.C.); jerry.edeani.244868@unn.edu.ng (J.E.); joshua.aniehe.244269@unn.edu.ng (J.A.); nebbieobiora@gmail.com (N.O.)
[2] National Centre for Energy Research and Development, University of Nigeria, Nsukka 410001, Nigeria; gerald.akubue@unn.edu.ng
[3] Department of Mechanical Engineering, Texas A&M University, College Station, TX 77843-3123, USA; joshua.okechukwu@tamu.edu
[4] Institute of Energy and Sustainable Development, De Montfort University, Leicester LE1 9BH, UK; muyiwa.oyinlola@dmu.ac.uk
* Correspondence: howard.njoku@unn.edu.ng; Tel.: +234-7081006700
† Presented at the 13th International Electronic Conference on Sensors and Applications, 18–20 November 2026; Available online: https://sciforum.net/event/ECSA-13.

Abstract

Accurate measurements of solar radiation are essential for photovoltaic (PV) and other solar energy applications. However, the high cost of standard thermopile pyranometers limits their deployment in distributed monitoring networks. In this study, a low-cost hybrid sensor combining a BH1750 optical sensor and dual thermistors with an ESP32 was developed. Using more than 53,000 field observations validated against a Class-A reference pyranometer, 40 engineered features were generated, of which 26 were retained using LASSO selection, out of which a ratio of optical and thermal corresponding values made up 88.15% of the importance magnitude. Four machine-learning algorithms Generalised Linear Regression (GLR), Random Forest (RF), Multilayer Perceptron Neural Network (MLPNN), and Gradient Boosting (GB) were evaluated. GB achieved the best overall performance, with ($R^2 = 0.9986$), RMSE = 10.60 W/m^2, and nRMSE = 2.30%, while RF provided the lowest mean absolute error (MAE = 4.23 W/m^2) with ($R^2 = 0.9985$) and RMSE = 10.84 W/m^2). Compared with the single-sensor baseline, the ensemble models reduced prediction error by approximately 80%. The prototype cost approximately $22.8 per unit, representing a potential 96% reduction relative to the Class-A reference pyranometer. These results demonstrate that combining low-cost optical and thermal measurements with machine-learning can achieve high-accuracy irradiance estimation while substantially reducing the cost.

Keywords: solar irradiance sensing; low-cost pyranometer; hybrid optical–thermal sensor; machine learning; gradient boosting; LASSO feature selection

Check for updates

Academic Editor: Francisco Falcone

Published: 29 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Solar radiation supplies most of the energy the Earth receives, with roughly 30% scattered, 20% absorbed by the atmosphere, and the remainder reaching the Earth's surface as terrestrial irradiance [1]. The devices used for the measurement of this terrestrial irradiance are crucial for monitoring the power output, fault diagnosis, performance calculation and

energy management procedures in distributed solar-powered systems and networks [2]. Standard thermopile pyranometers, though accurate, are expensive, limiting dense spatial deployment and necessitating alternatives like low-cost pseudo-pyranometers [3]. Prior low-cost pyranometer designs, including photodiode-based, LDR-based, and silicon-cell reference cells, deviate from the output of high-accuracy thermopile pyranometers, with Karki et al. [4] reporting a deviation of 1.43% and Imoleayo et al. [5] identifying temperature and spectral effects as the cause of this error in the system. Studies have applied correction models to reduce the difference compared to standard thermopile sensor measurements, with Mubarak et al. [6] achieving a daily solar irradiation relative root mean squared difference from 10.6% to 5.4%, and Heusinkveld et al. [7] achieving an RMSE of 11.4 W/m^2 under clear and cloudy sky conditions. The feasibility of combining low-cost ambient light sensors with statistical and machine learning models, specifically linear, random forest, and support vector regressions, to accurately estimate solar irradiance has been demonstrated in a study by [8], where low-cost light sensors with cloud storage and processing capabilities for data collection and model application on an internet-of-things (IoT) device were developed. The study demonstrates that the random forest model achieved a high coefficient of determination (R^2) of 0.9922, indicating a strong correlation between sensor-measured illuminance and solar irradiance. In a different study, Ramadhan et al. [9] compared the accuracies of physical and machine learning models for solar power modelling, specifically for global horizontal irradiance, direct normal irradiance, global tilted irradiance, and photovoltaic power and concluded that machine learning models generally outperform conventional physical models for estimating solar irradiance and photovoltaic power when the input parameters are properly selected.

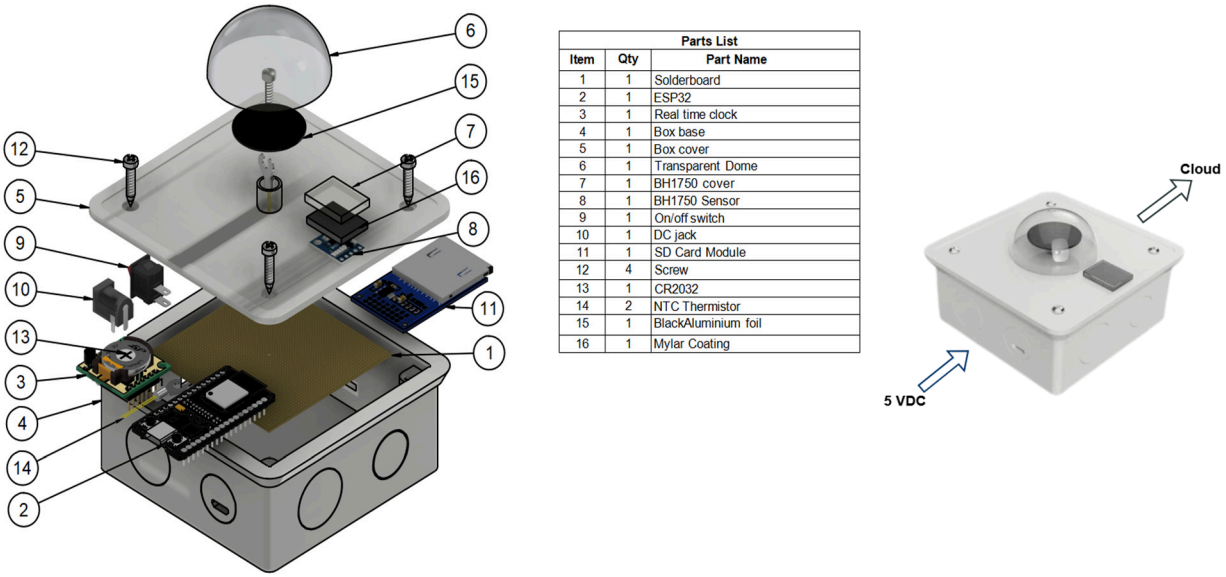
These studies confirm that low-cost sensors paired with ML are a viable alternative to expensive reference instruments, but a hybrid optical–thermal sensing approach for enhanced irradiance sensing remains unexplored, as well the application of systematic feature selection determined by ML-based learning. This study addresses that gap with a hybrid optical–thermal pyranometer using a BH1750 light sensor and two thermistors for thermal sensing integrated on an IoT-enabled ESP32 microcontroller for model implementation and cloud storage. The model is trained on least absolute shrinkage and selection operator (LASSO) selected datasets and validated directly against a Class-A reference pyranometer. This approach aims to enable the ML models to learn the thermal effect on the light sensors while also utilising the thermistor values as additional inputs, reducing inaccuracy, at a lower cost.

2. Materials and Methods

2.1. Sensor System and Data Collection

The hybrid pyranometer, shown in Figure 1a,b integrates a BH1750 digital ambient light sensor (ROHM Semiconductors (China) Co., Ltd., Shanghai, China) and two thermistors (NTC type; RPD Sensor, Dongguan, China), mounted respectively on a black anodised aluminium absorber and an ambient-reference position and interfaced with an ESP32 microcontroller (Espressif Systems, Shanghai, China) for synchronous data acquisition. The BH1750 is enclosed in a translucent Mylar-coated housing to mitigate saturation under high irradiance. The system was deployed outdoors alongside a Class-A Veinasa XS-TBQ02A thermopile pyranometer (Mianyang Xiangshu Electronic Technology Co., Ltd., Mianyang, China), as the reference. The experimental setup is shown in Figure 2.

https://doi.org/10.3390/engproc2026163054



Parts List		
Item	Qty	Part Name
1	1	Solderboard
2	1	ESP32
3	1	Real time clock
4	1	Box base
5	1	Box cover
6	1	Transparent Dome
7	1	BH1750 cover
8	1	BH1750 Sensor
9	1	On/off switch
10	1	DC jack
11	1	SD Card Module
12	4	Screw
13	1	CR2032
14	2	NTC Thermistor
15	1	BlackAluminium foil
16	1	Mylar Coating

(**a**) (**b**)

Figure 1. (**a**,**b**) Detailed description of hybrid optical–thermal pyranometer.

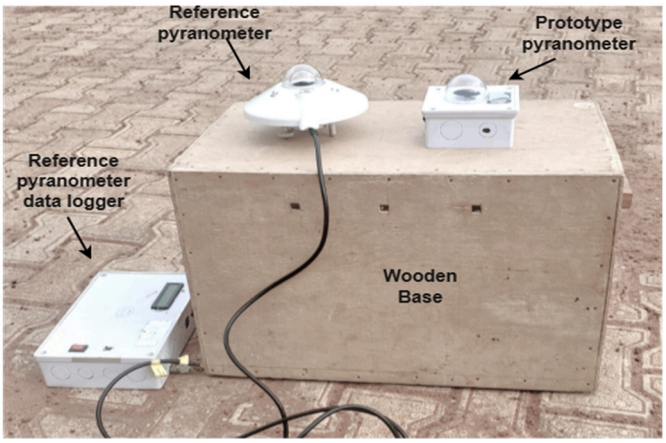


Figure 2. Experimental setup (National Centre for Energy Research and Development, University of Nigeria, Nsukka, [6.86696, 7.40965]).

Both instruments were time-synchronised and logged at 1 s intervals, yielding a dataset of over 53,000 paired observations spanning clear-sky, partly cloudy, and overcast conditions.

2.2. Data Preprocessing and Validation

The raw analogue-to-digital converter (ADC) values from the hybrid pyranometer and the global horizontal irradiance values from the Class-A reference pyranometer were first synchronised and examined for transient irradiance conditions. The periods exhibiting rapid changes in irradiance were excluded because abrupt irradiance transitions will introduce a temporal mismatch between the thermal response of the hybrid sensor and the reference instrument. The resulting dataset contained more than 53,000 paired observations of between 0 and 1333 W/m^2.

Data partitioning, feature standardisation, feature selection, and model optimisation were performed using the training data. The test dataset was retained as an independent dataset and was not used during feature selection or hyperparameter optimisation. The data were partitioned chronologically to ensure that observations in the test period represented

conditions not directly available during model training. The candidate predictors were then standardised using the mean and standard deviation calculated from the training data.

2.3. Feature Engineering and Selection

The measured illuminance, absorber thermistor ADC values, and ambient thermistor ADC values were used to construct a set of candidate predictors representing instantaneous sensor states, temporal behaviour, variability, and nonlinear interactions. In addition to the original signals, moving averages, moving standard deviations, first- and second-order derivatives, interaction terms, ratios, and polynomial terms were generated.

A total of 40 candidate features were considered. LASSO regression reduced this set to 26 predictors. The retained variables included the instantaneous illuminance and temperature signals, moving statistics, temporal derivatives, optical–thermal interaction terms, temperature ratios, and polynomial terms and include L, T_2, $\mathrm{MA}_{30}(L)$, $\mathrm{MA}_{60}(L)$, $\mathrm{MA}_{30}(T_1)$, $\mathrm{MA}_{30}(\Delta V)$, $\mathrm{MSD}_{30}(L)$, $\mathrm{MSD}_{60}(L)$, $\mathrm{MSD}_{30}(T_1)$, $\mathrm{MSD}_{30}(T_2)$, $\mathrm{MSD}_{30}(\Delta V)$, L', ${T_1}'$, ${T_2}'$, L'', ${T_1}''$, ${T_2}''$, $L\Delta V$, T_1/T_2, T_2/T_1, L/T_2, L^2, ${T_1}^2$, ${T_2}^2$, ΔV^2, and ΔV^3. The candidate predictors were standardised before coefficient-based selection so that the resulting LASSO coefficients can be used to examine the contribution of the retained predictors.

2.4. Model Evaluation

Four regression models were evaluated: generalised linear regression (GLR), random forest (RF), multilayer perceptron neural network (MLPNN), and gradient boosting (GB). All models were implemented in MATLAB (version 2024b), The MathWorks, Inc., MA, USA). Model performance was evaluated on the test set based on the following performance metrics: root mean squared error (RMSE), the coefficient of determination (R^2), mean absolute error (MAE), mean bias error (MBE), and normalised root mean squared error (nRMSE), as described in Equations (1)–(5).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}, \quad (1)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i|, \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}, \quad (3)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i), \quad (4)$$

$$nRMSE = 100\frac{RMSE}{\overline{y}}, \quad (5)$$

where $\hat{y}_i$ is the *ith* predicted irradiance value, y_i is the corresponding irradiance value, $\overline{y}$ is the mean of the actual values, and n is the number of observations.

The error was additionally identified by the irradiance regime for clear sky ($G > 700$ W/m^2), partly cloudy ($300 < G \leq 700$ W/m^2), and overcast ($G \leq 300$ W/m^2) to demonstrate conditional performance. The error reduction was benchmarked against a single-sensor (illuminance) linear regression baseline.

A summary of the methodology is shown in Figure 3.

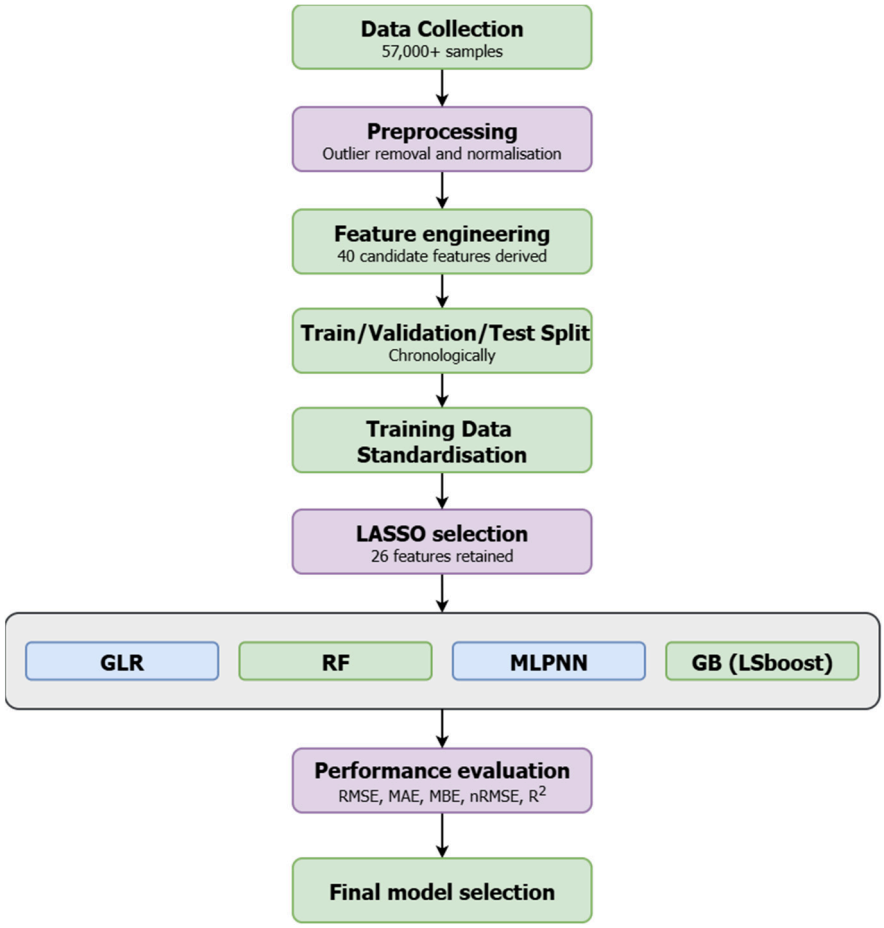


Figure 3. Flowchart of the methodology of the study.

3. Results

3.1. LASSO Feature Selection and Optical–Thermal Relationships

Figure 4 shows the relative contribution of the selected predictors based on the magnitude of their standardised LASSO coefficients.

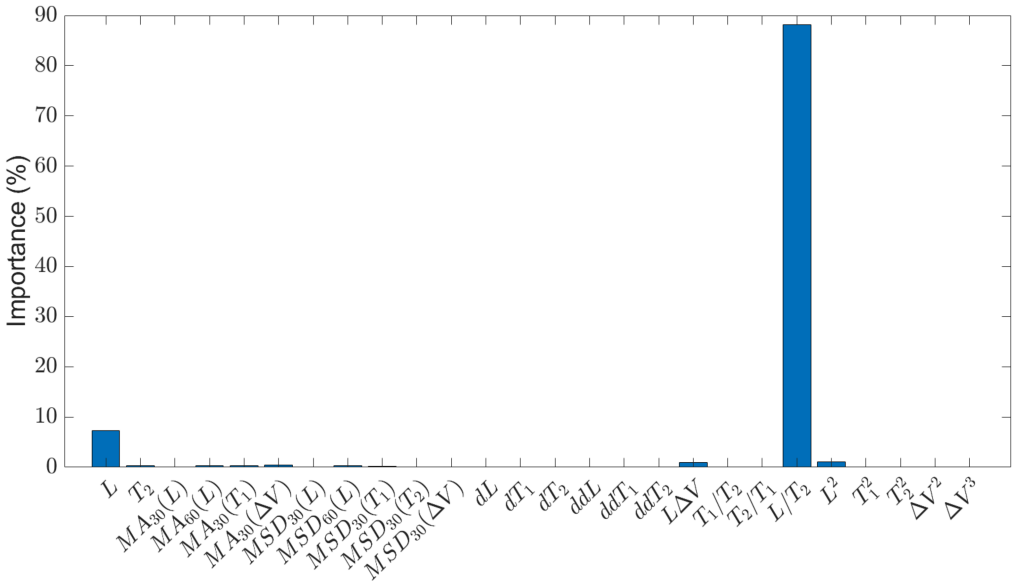


Figure 4. Importance of selected features.

The L/T_2 feature exhibited the largest importance of 88.15%, followed by L with 7.27%, among the retained predictors, indicating that the combination of optical and thermal information provides substantially more predictive information than instantaneous illuminance alone. This result is significant because the objective of the proposed architecture is not simply to measure illuminance but to compensate for the environmental and thermal influences that affect the relationship between the low-cost optical sensor and broadband solar irradiance.

The prominence of L/T_2 also suggests that the ambient thermal state of the sensor is associated with variations in the optical measurement. This indicates that incorporating thermal information provides additional predictive information that is not contained in the instantaneous optical signal alone.

The remaining selected predictors include moving statistics, derivatives, interaction terms, ratios, and polynomial terms make up the remaining 4.58% of combined importance. Their selection indicates that the relationship between the low-cost sensor outputs and broadband irradiance is not adequately represented by a simple instantaneous linear relationship. Temporal variation and nonlinear combinations of the optical and thermal measurements contribute additional information that can be exploited by the regression models.

3.2. Model Performance Comparison

Table 1 compares the performance of GLR, MLPNN, GB, and RF using the 26 LASSO-selected predictors. The models exhibited substantial differences in prediction accuracy.

Table 1. Comparison of different ML models used in this study.

Model	RMSE (W/m^2)	MAE (W/m^2)	R^2	MBE (W/m^2)	nRMSE (%)	Error Reduction vs. Single Optical Model (%)
GLR	44.529	32.726	0.9750	−0.799	9.675	17.12
MLPNN	20.363	13.398	0.9948	0.155	4.424	62.10
GB (LSBoost)	10.597	5.162	0.9986	0.241	2.300	80.28
RF	10.840	4.233	0.9985	0.244	2.355	79.82

GLR produced the largest prediction error, with an RMSE of 44.529 W/m^2, MAE of 32.726 W/m^2, R^2 of 0.9750, and nRMSE of 9.675%. MLPNN substantially reduced the error, achieving an RMSE of 20.363 W/m^2, MAE of 13.398 W/m^2, R^2 of 0.9948, and nRMSE of 4.424%. The two ensemble methods provided the strongest overall performance. GB achieved an RMSE of 10.597 W/m^2, MAE of 5.162 W/m^2, R^2 of 0.9986, MBE of 0.241 W/m^2, and nRMSE of 2.300%. RF produced a comparable RMSE of 10.840 W/m^2 and R^2 of 0.9985, while achieving the lowest MAE of 4.233 W/m^2.

This result shows that under this dataset, feature representation, network architecture, and optimisation procedure, the tree-based models (GB and RF) provide a more effective representation of the relationship between the engineered optical–thermal predictors and the reference irradiance.

The optical-only baseline produced an RMSE of 53.73 W/m^2. Relative to this baseline, GB reduced the RMSE by 80.28%, while RF reduced it by 79.82%.

3.3. Performance Under Different Irradiance Conditions

Figure 5 presents the RMSE obtained under clear-sky, partly cloudy, and overcast conditions to assess how the models maintain their predictive performance when the irradiance environment changes.

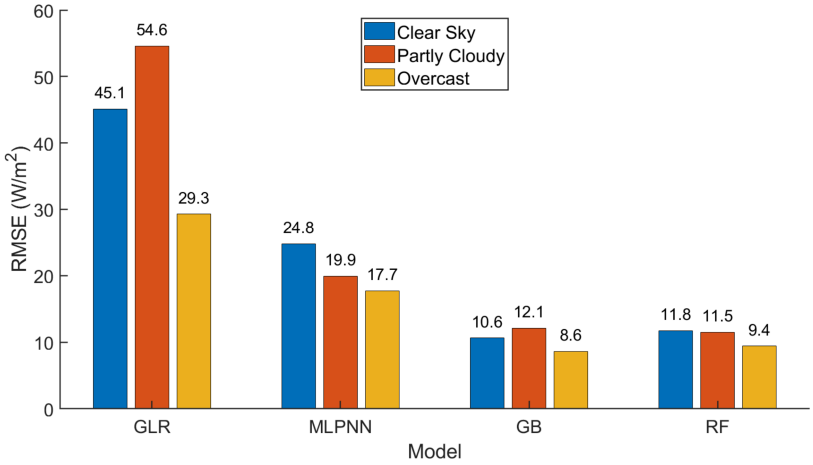


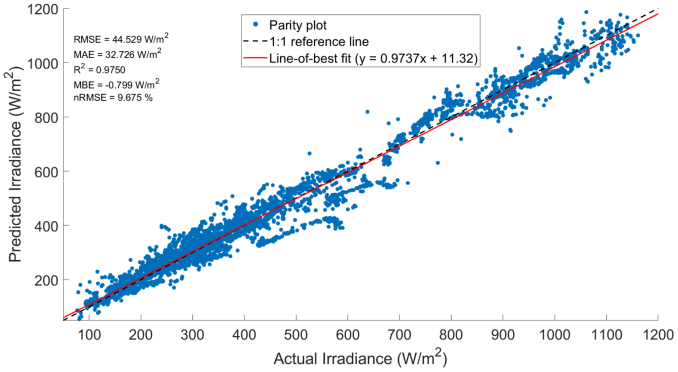
Figure 5. Model performance across different irradiance regime.

For GB, the RMSE values were 10.638, 12.112, and 8.605 W/m^2 for clear-sky, partly cloudy, and overcast conditions, respectively. RF produced corresponding RMSE values of 11.753, 11.506, and 9.422 W/m^2.

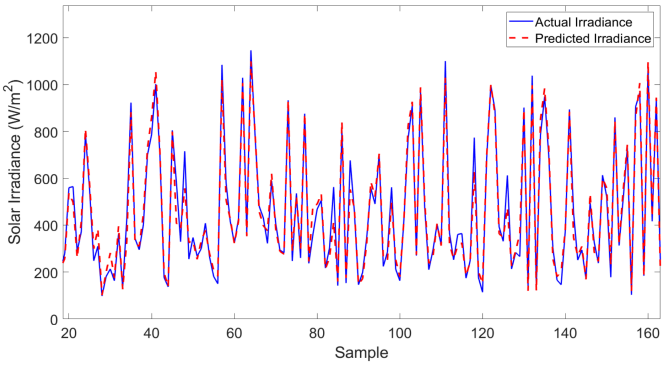
The relatively low RMSE obtained by both ensemble models across all three regimes demonstrates that their performance was not restricted to a single irradiance condition. For GB, the RMSE remained between 8.605 and 12.112 W/m^2 across the three regimes. Overall, the results demonstrate that the ensemble models retain relatively consistent absolute accuracy under different irradiance conditions, with partly cloudy conditions representing the most challenging regime for GB.

3.4. Prediction and Error Characteristics

Figure 6a,c,e,g presents the parity plot for GLR, MLPNN, GB and RF with the corresponding predicted and actual irradiance line plots for GLR, MLPNN, GB and RF in Figure 6b,d,f,h, respectively. The predictions are seen to become more concentrated around the 1:1 line, as the line of best fit also closely overlaps the 1:1 line from GLR to MLPNN and to the ensemble models of GB and RF. The line of best fit provides an additional indication of the systematic error. A slope close to unity indicates that the model reproduces the scale of the reference measurements with limited proportional bias, while an intercept close to zero indicates limited offset bias.

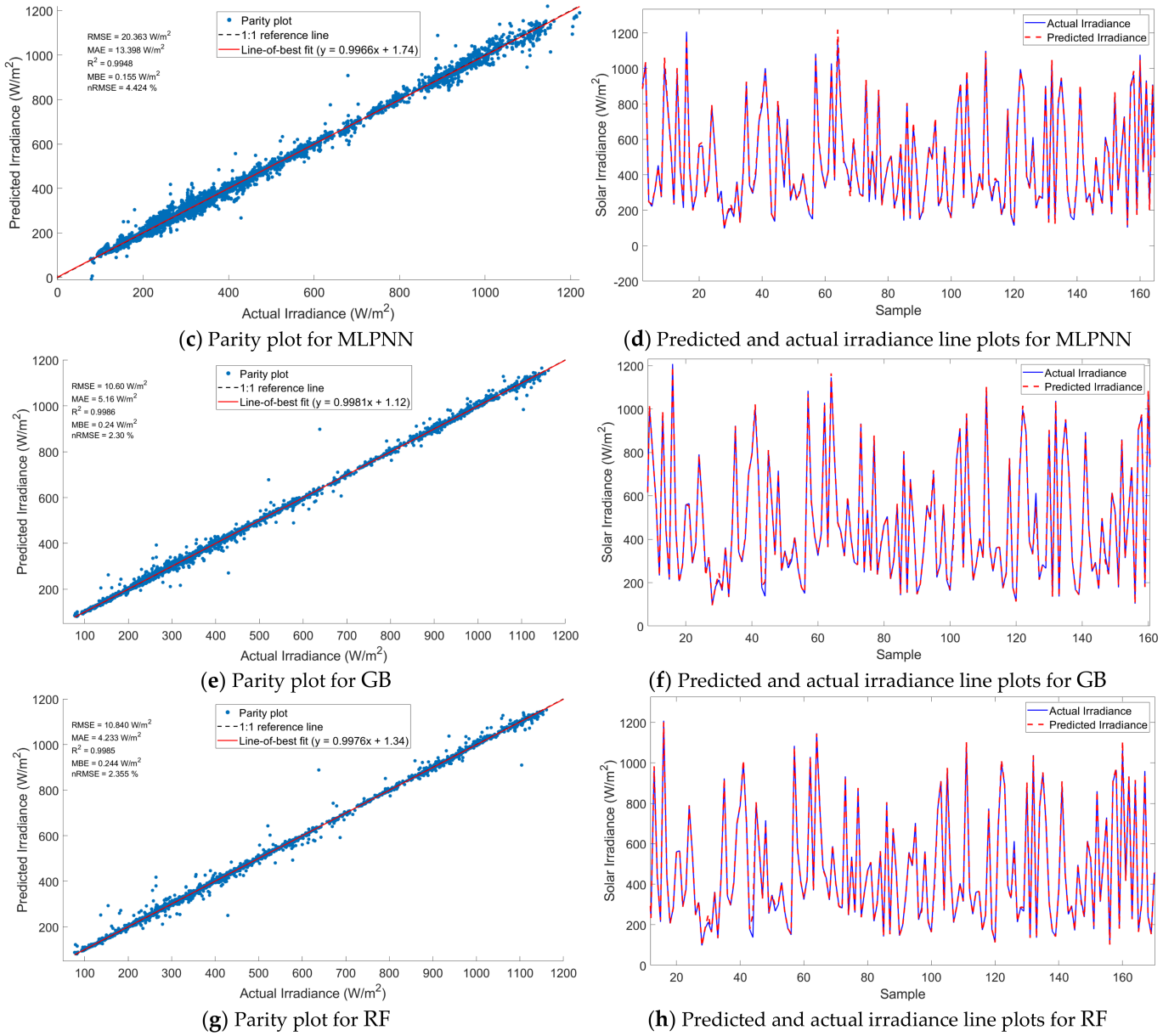


(**a**) Parity plot for GLR



(**b**) Predicted and actual irradiance line plots for GLR

Figure 6. *Cont.*



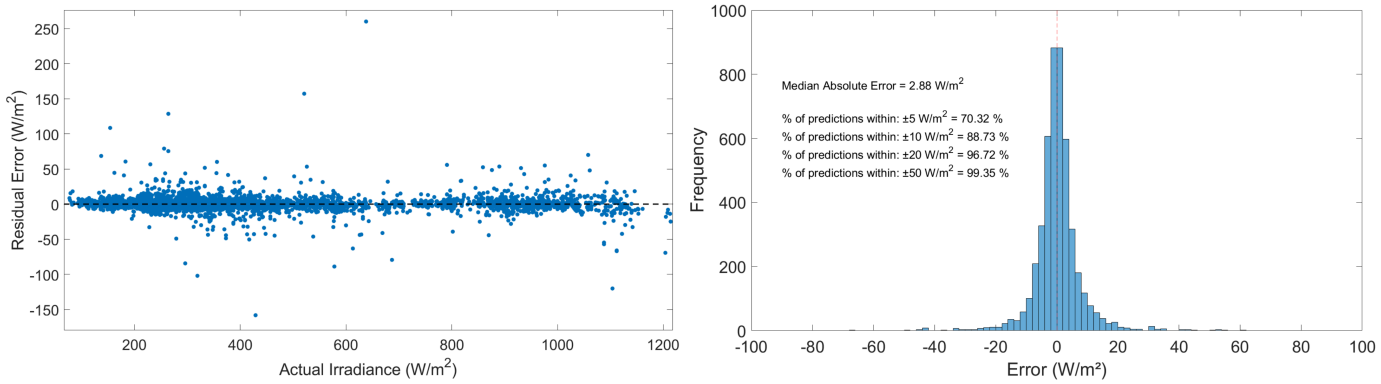
(**c**) Parity plot for MLPNN (**d**) Predicted and actual irradiance line plots for MLPNN

(**e**) Parity plot for GB (**f**) Predicted and actual irradiance line plots for GB

(**g**) Parity plot for RF (**h**) Predicted and actual irradiance line plots for RF

Figure 6. Visual representation of test results: parity plot for GLR, MLPNN, GB and RF in (**a**), parity plot for GLR, (**c**) parity plot for MLPNN, (**e**) parity plot for GB, and (**g**) parity plot for RF. Corresponding line plots of GLR, MLPNN, GB and RF in (**b**), corresponding line plots of GLR, (**d**) corresponding line plots of MLPNN (**f**), corresponding line plots of GB, and (**h**) corresponding line plots of RF.

Figure 7a,b presents the residual error analysis together with the quantitative error metrics for GB, selected based on its lower RMSE. The residuals are centred close to zero, consistent with the MBE of 0.241 W/m^2 obtained for the GB model. The concentration of the residuals around zero indicates that the model does not exhibit a substantial systematic tendency to overestimate or underestimate irradiance. As shown in Figure 7, the median absolute error of 2.88 W/m^2 and the percentage of predictions within ±5 W/m^2 of 70.32%, indicate that most errors are relatively small.



(**a**) Residual Analysis for GB (**b**) Residual error distribution for GB

Figure 7. (**a**) Residual Analysis for (GB), (**b**) residual error distribution for GB.

3.5. Computational Cost and Reproducibility

Table 2 presents the model sizes and selected hyperparameters. GLR had the smallest model size, followed by MLPNN, whereas GB and RF required substantially larger model representations.

Table 2. Model sizes and hyperparameters.

Model	Model Size (kB)	Hyperparameters
GLR	7.6289	Distribution = normal, link = identity, terms = linear (OLS)
MLPNN	34.841	LayerSizes = 149, activations = "relu"
GB (LSBoost)	26,536	NumLearningCycles = 2219, MinLeafSize = 3, MaxNumSplits = 150, LearnRate = 0.00379
RF	28,443	NumLearningCycles = 340, MinLeafSize = 1, MaxNumSplits = 10,810, NumVariablesToSample = 12

The results show that although GB and RF produced the best prediction accuracy, their reported model sizes indicate that direct deployment of the complete trained models on resource-constrained microcontrollers will require additional optimisation.

3.6. Comparison Against Prior Low-Cost Approaches

The proposed prototype was constructed at a component-level cost of approximately $22.78. Relative to the $595.19 purchase price of the Class-A reference instrument used in this study, this corresponds to an approximately 96% reduction in prototype component cost. This comparison highlights the economic potential of the proposed architecture for applications requiring multiple irradiance measurement points.

Table 3 shows that the RMSE obtained in the present study is lower than the values reported for the following recent low-cost irradiance-sensing studies. It is also seen that the error reduction compared to those studies is similar to the error reduction also found in this study for the optical-only baseline. The proposed system therefore demonstrates competitive prediction accuracy while using inexpensive optical and thermal sensing components.

Table 3. Comparison of this study against recent studies.

Studies	RMSE (W/m^2)	Error Reduction
This Study	10.60	-
[3]	55.82	81.01%
[8]	44.46	76.17%
[10]	27.80	61.88%

4. Discussion

4.1. Contribution of Hybrid Optical–Thermal Sensing

The main contribution of this study is the combination of low-cost illuminance measurements with thermal measurements rather than relying on illuminance alone. The LASSO analysis has provided evidence that thermal information contributes significantly to irradiance prediction, with L/T_2 emerging as the dominant retained predictor with 88.15% of the total importance magnitude.

4.2. Why Ensemble Models Outperform

The superior performance of GB and RF relative to GLR and MLPNN can be interpreted in terms of the structure of the engineered predictor space. The 26 selected predictors contain temporal statistics, derivatives, ratios, polynomial terms, and interaction variables. These predictors introduce nonlinear relationships between the optical and thermal sensor outputs and the reference irradiance.

GLR is constrained to a linear functional form and therefore cannot fully exploit complex threshold and interaction effects. Although polynomial and interaction features partially compensate for this limitation, the resulting linear model remains less flexible than a tree-based ensemble, hence, the substantially lower errors of RF and GB; this also indicates that the nonlinear structure of the optical–thermal relationship is well suited to tree-based regression.

4.3. Robustness Across Irradiance Conditions

A major consideration for an irradiance sensor is whether its performance remains reliable under changing atmospheric conditions. The analysis shows that both GB and RF maintained a relatively low RMSE across clear-sky, partly cloudy, and overcast conditions, necessary for practical outdoor deployment.

4.4. Error Characteristics and Practical Accuracy

The error statistics indicate both high agreement and low systematic bias. The GB model produced an RMSE of 10.597 W/m^2, MAE of 5.162 W/m^2, and MBE of only 0.241 W/m^2. The near-zero MBE indicates that positive and negative errors largely cancel over the test dataset, suggesting almost no bias, given that the reference pyranometer has some inherent error.

The difference in results between GB and RF, where GB achieved the lowest RMSE of 10.597 W/m^2, whereas RF achieved the lowest MAE of 4.233 W/m^2 indicates that GB has better control of relatively large errors, while the lower MAE of RF indicates smaller average absolute deviations. Thus, the two models exhibit similar overall performance but different error characteristics. The residual analysis also indicates that there is no persistent systematic bias.

4.5. Comparison with Previous Low-Cost Irradiance-Sensing Approaches

The RMSE achieved by the proposed system achieves competitive accuracy relative to the recently published low-cost irradiance-sensing studies presented in Table 3. Compared with the reported values of 55.82, 44.46, and 27.8 W/m^2 for the selected recent studies, this study's idea that combining optical and thermal information can reduce the limitations associated with single-sensor irradiance estimation and results of LASSO selected features is confirmed.

4.6. Cost, Deployment Potential, and Limitations

The low component cost of the proposed system is particularly relevant for applications requiring spatially distributed irradiance measurements. A conventional Class-A pyranometer is expensive when a large number of measurement points are required. A low-cost machine-learning enhanced sensor system could therefore provide a less expensive alternative for applications such as distributed solar-resource assessment, academic research and photovoltaic monitoring in micro- and nano- grid applications.

It is noted, however, that the experimental validation conducted under the environmental conditions represented by the collected dataset does not by itself establish generalisation across different geographical locations, seasons, and climates. Therefore, the present results do not imply formal compliance with the complete performance requirements for a Class-A pyranometer. It is also noted that the relatively large model size representations reported for GB and RF indicate that model deployment requires consideration of computational resources.

Future work will, therefore, prioritise the investigation of model compression and feature reduction to enable direct deployment on resource-constrained platforms.

4.7. Conclusions

In summary, this study developed a low-cost hybrid optical–thermal pyranometer based on a BH1750 optical sensor and dual thermistors on an ESP32 platform. Forty engineered predictors were created from the optical and thermal measurements, of which 26 were retained through LASSO feature selection.

Among the evaluated models, the tree-based ensemble methods provided the strongest overall performance. Gradient Boosting achieved an RMSE of 10.597 W/m^2, MAE of 5.162 W/m^2, R^2 of 0.9986, MBE of 0.241 W/m^2, and nRMSE of 2.300%, while RF produced a comparable RMSE of 10.840 W/m^2 and R^2 of 0.9985. The near-zero MBE indicates negligible overall systematic bias.

The feature-selection results indicate that optical–thermal combinations, particularly L/T_2 with 88.15% of the importance magnitude, provide important predictive information beyond instantaneous illuminance. The condition-based analysis further showed that the ensemble models maintained relatively stable errors under different sky conditions.

The prototype component cost of approximately $22.78 also demonstrates the economic potential of the proposed architecture for applications requiring distributed irradiance measurements. Future work in this study will focus on multi-site, multi-season, and long-term field validation, model compression, ablation studies, and stand-alone embedded system deployment. Overall, the results demonstrate that combining inexpensive optical and thermal sensing with machine-learning-based models is a promising approach for improving the accuracy of low-cost solar irradiance monitoring systems.

https://doi.org/10.3390/engproc2026163054

Author Contributions: Conceptualisation, H.N.; methodology, P.E.; software, P.E. and J.E.; validation, P.E.; formal analysis, P.E.; resources, H.N., M.O. and G.U.A.; data curation, P.E., N.O. and J.A.; writing—original draft preparation, P.E. and J.O.; writing—review and editing, H.N.; visualisation, P.E. and R.C.; supervision, H.N.; project administration, H.N. and M.O.; funding acquisition, M.O. and H.N. All authors have read and agreed to the published version of the manuscript.

Funding: This work was conducted under the Circular Economy Powered Renewable Energy Centre (CEPREC), supported by the UK Government's Ayrton Fund and the International Science Partnerships Fund (ISPF), awarded by UK Research and Innovation (UKRI) under grant UKRI/EP/B000320/1.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data supporting the reported results can be found at https://drive.google.com/file/d/1CgWka1tJrQKW0jjOgSVh0qSZxLummyhw/view?usp=drive_link (accessed on 28 October 2026).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations were used in this manuscript:

ADC	Analog-to-Digital Converter
GB	Gradient Boosting
GLR	Generalised Linear Regression
I2C	Inter-Integrated Circuit
IoT	Internet of Things
LASSO	Least Absolute Shrinkage and Selection Operator
MAE	Mean Absolute Error
MBE	Mean Bias Error
MLPNN	Multilayer Perceptron Neural Network
nRMSE	Normalised Root Mean Square Error
R^2	Coefficient of Determination
RF	Random Forest
RMSE	Root Mean Square Error
RTC	Real-Time Clock

References

1. Rottman, G. Measurement of total and spectral solar irradiance. *Space Sci. Rev.* **2006**, *125*, 39–51. [CrossRef]
2. Zhu, H.; Jiang, T.; Zhou, H.; Sun, Y.; Ma, W.; Zhang, X. A novel soft sensing method using intelligent modeling method for solar irradiance and temperature in distributed PV power plant. *Energy Sci. Eng.* **2024**, *12*, 1041–1058. [CrossRef]
3. Voicu, I.V.; Petreuș, D. Global Horizontal Irradiance Estimation Using a Hybrid Physical-Machine Learning Soft Sensor Based on a Low-Cost Photovoltaic Measurement Platform. *Appl. Sci.* **2026**, *16*, 4507. [CrossRef]
4. Karki, S.; Ziar, H.; Korevaar, M.; Bergmans, T.; Mes, J.; Isabella, O. Performance evaluation of silicon-based irradiance sensors versus thermopile pyranometer. *IEEE J. Photovolt.* **2020**, *11*, 144–149. [CrossRef]
5. Imoleayo, A.J.; Aduramigba, A.D.; Melodi, A.O. Development of a Low-Cost Irradiance Meter with Remote Data Logger. *Adv. Eng. Forum* **2025**, *54*, 91–103. [CrossRef]
6. Mubarak, R.; Schilke, H.; Seckmeyer, G. Improving the irradiance data measured by silicon-based sensors. *Energies* **2021**, *14*, 2766. [CrossRef]
7. Heusinkveld, B.G.; Mol, W.B.; van Heerwaarden, C.C. A new accurate low-cost instrument for fast synchronized spatial measurements of light spectra. *Atmos. Meas. Tech.* **2023**, *16*, 3767–3785. [CrossRef]
8. Nava-Pintor, J.A.; Alcalá-Rodríguez, U.E.; Guerrero-Osuna, H.A.; Mata-Romero, M.E.; Lopez-Neri, E.; García-Vázquez, F.; Solís-Sánchez, L.O.; Carrasco-Navarro, R.; Luque-Vega, L.F. Development and evaluation of solar radiation sensor using cost-effective light sensors and machine learning techniques. *Technologies* **2025**, *13*, 182. [CrossRef]

9. Ramadhan, R.A.; Heatubun, Y.R.; Tan, S.F.; Lee, H.J. Comparison of physical and machine learning models for estimating solar irradiance and photovoltaic power. *Renew. Energy* **2021**, *178*, 1006–1019. [CrossRef]
10. Lim, M.T.; Tan, R.H.; Tan, G.A.; Hew, C.H.; Lee, W.K.; Mercha, M. Development of low cost pyranometer IoT based solar irradiance measurement station. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2023; Volume 1281, p. 012012.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.