Uniosun Journal of Engineering and Environmental Sciences (UJEES), Vol. 8 No. 2, Sept., 2026
DOI: 10.64980/ujees.v8i2.004 **Onwusa et al.**
Published: September 30, 2026

ENHANCING COMBUSTION EFFICIENCY THROUGH ADVANCED RADIACTIVE HEAT TRANSFER ANALYSIS AND ML PREDICTION OF GAS EMISSIVITY AND FLAME TEMPERATURE

Sylvester Chukwutem ONWUSA[1*], Omovigho Prosper OGHENERUKEVWE[2] ,Oderhowho NYORERE[3], .Friday Ikechukwu NWAOSA[4], Charles Olordo USIKPEDO[5], and Oghenerobo Cyril UYERI[6]

[1] Department of Mechanical Engineering, Southern Delta University, Ozoro, P.M.B. 5, Nigeria
[2]Department of Mechanical Engineering, Southern Delta University, Ozoro, P.M.B. 5, Nigeria
[3]Department of Agricultural and Biosystems Engineering, Southern Delta University, Ozoro P.M.B. 5, Nigeria
[4]Department of Industrial Technology Education, Faculty of Technology and Vocational Education, Nsukkwa , Enugu State6+
[5,6] Department of Mechanical Engineering, Southern Delta University, Ozoro, P.M.B. 5, Nigeria

[*]Corresponding Author: onwusachukwuteylvester@gmail.com/onwusacs@dsust.edu.ng

Abstract

This study presents a hybrid computational framework integrating radiative heat transfer (RHT) physics, Design-Expert-based response surface methodology (RSM), and Machine Learning (ML) for predicting and optimizing combustion characteristics in gaseous flame systems. RHT is modeled using the Stefan–Boltzmann law under gray-gas assumptions, while gas emissivity is characterized using temperature-, composition-, and soot-dependent correlations. A Design-Expert central composite design (CCD) systematically investigates nonlinear relationships among key combustion variables and develops response surface models for combustion performance. In parallel, a Random Forest ML model is developed using combustion-related features, including flame temperature, spectral radiation characteristics, soot concentration, gas composition, and operating conditions. The model achieves high predictive performance, with R^2 values of 0.998 for flame temperature and 0.978 for effective emissivity. However, these results require cautious interpretation because the dataset contains approximately 800 samples and covers a relatively narrow temperature range of 1150–1400 K, potentially limiting model generalizability under substantially different operating conditions. Comparison with physics-based evaluations further indicates that the physics-based models tend to overestimate radiative heat losses; therefore, strong agreement between the two approaches cannot be claimed. Instead, ML provides an accurate data-driven representation within the investigated range, while physics-based models provide the theoretical basis for interpreting RHT behaviour. Design-Expert desirability optimization identifies operating conditions that balance flame temperature, emissivity, RHT, and combustion performance. The framework provides a useful approach for combustion analysis and optimization, while emphasizing the need for larger, more diverse datasets and broader operating conditions to improve generalization. Potential applications include high-temperature furnace design, thermal-energy management, and data-driven combustion optimization.

Keywords
Radiative Heat Transfer; Combustion Efficiency; Machine Learning; Flame Temperature; Gas Emissivity

1. INTRODUCTION

Efficient combustion is fundamental to modern energy conversion systems, including industrial furnaces, boilers, gas turbines, and other high-temperature thermal processes, where achieving high energy efficiency while minimizing thermal losses and environmental impacts is essential. Among the major heat-transfer mechanisms governing high-temperature combustion, RHT becomes increasingly important because of its strong dependence on flame temperature, gas composition, participating species, and soot concentration. In combustion environments containing radiatively active species such as carbon dioxide (CO_2), water vapour (H_2O), and soot, radiation can contribute substantially to the overall energy exchange. Consequently, reliable prediction of radiative behaviour is essential for accurate assessment of combustion performance, thermal efficiency, and energy utilization [1,2].

The theoretical description of radiative transfer in participating combustion media is commonly based on the radiative transfer equation RHT, which accounts for radiation emission, absorption, scattering, and propagation. Although RHT-based approaches provide a physically grounded representation of radiation transport, their application to practical combustion systems can be computationally demanding because of spectral dependence, multidimensional geometries, non-grey gas behaviour, temperature gradients, and strong coupling between radiation, turbulence, and combustion chemistry [3,4]. Detailed spectral models can therefore require substantial computational resources, particularly when they are incorporated into computational fluid dynamics (CFD) simulations or repeated optimization procedures. Simplifications such as the grey-gas and optically thin assumptions can reduce computational requirements, but may introduce errors when applied outside their appropriate physical regimes [5].

These limitations have encouraged the development of reduced-order, response-surface, and data-driven approaches for predicting combustion and radiation characteristics. Response surface methodology (RSM), particularly when implemented through Design-Expert and central composite design (CCD), provides a systematic means of investigating the effects and interactions of process variables while requiring substantially fewer computational or experimental runs than exhaustive parameter sweeps. RSM has consequently been applied to combustion optimization, thermal-process design, and the determination of operating conditions that provide favourable combinations of temperature, heat transfer, efficiency, and emissions [6,7].

ML has also emerged as an effective surrogate-modeling approach for complex combustion processes. Algorithms such as RF, ANNs, SVMs, XDboost, and other ensemble techniques can learn nonlinear relationships between operating variables and combustion responses from experimental, numerical, or synthetic datasets. Such models have demonstrated considerable potential for predicting flame temperature, emissions, combustion performance, and other thermal characteristics while reducing the computational cost associated with repeated high-fidelity simulations [8–10]. The principal advantage of ML in this context is its ability to approximate complex input–output relationships without requiring a complete numerical solution of the underlying governing equations for every prediction.

However, the growing application of ML does not eliminate the importance of physical modeling. Purely data-driven models can produce highly accurate predictions within the domain represented by their training data but may exhibit reduced reliability when extrapolated to unseen operating conditions. This limitation has motivated increasing interest in physics-informed and physics-guided ML, in which physical laws, constraints, governing equations, or physically meaningful variables are incorporated into the learning process or used to interpret model predictions [11–13]. Recent studies have specifically investigated the integration of radiation physics and ML, demonstrating that the research area is no longer characterized by an absence of physics-informed radiation modeling. Therefore, the research gap should not be expressed as a lack of studies incorporating physical radiation principles into ML. Rather, the remaining challenge concerns how such physical knowledge can be combined effectively with data-driven prediction and statistical optimization for the specific coupled problem of flame temperature, gas emissivity, and radiative heat transfer.

RHT presents a particularly challenging modeling problem because emissivity is not an independent constant under practical combustion conditions. It varies with temperature, gas composition, concentration of radiatively participating species, optical path length, and, where applicable, soot loading. Flame temperature similarly depends on fuel–air ratio, flow conditions, heat release, and thermochemical interactions. These variables are therefore strongly coupled. A model that predicts temperature without adequately representing emissivity may provide an incomplete description of radiative energy exchange, while a radiation model that accurately estimates emissivity but neglects the corresponding combustion conditions may have limited value for process optimization [14, 15].

Existing hybrid modeling studies have addressed portions of this challenge through combinations of CFD, RSM, ML, and physics-based formulations. CFD–ML approaches, for example, have been used to develop computationally efficient surrogates for complex reacting-flow simulations, while RSM–ML approaches have been applied to process optimization and nonlinear response prediction [16–18]. Physics-informed ML has further demonstrated the possibility of embedding physical constraints into data-driven models [11–13]. Nevertheless, these approaches differ substantially in their objectives, physical representations, input variables, datasets, and target responses. Consequently, the research gap is not simply the absence of hybrid models, but the limited integration of radiative-transfer physics, emissivity-sensitive prediction, flame-temperature prediction, and statistical optimization within a common computational framework. This distinction is important because it positions the present study relative to existing hybrid approaches without overstating the novelty of physics-informed ML itself.

Another important consideration is the nature of the data used for model development and evaluation. The present investigation is a computational study and reanalysis of an existing dataset, rather than a new experimental investigation. The analysis uses approximately 800 observations covering a flame-temperature range of 1150–1400 K. The relatively limited sample size and restricted temperature range provide a defined modeling domain but also impose limitations on the generalizability of the resulting ML models. Accordingly, high predictive coefficients of determination should be interpreted as evidence of strong performance within

the investigated data domain rather than as proof that the models will perform equally well under substantially different temperatures, fuel compositions, flow regimes, or combustion configurations.

The present study therefore develops a hybrid computational framework combining radiative heat-transfer principles, Design-Expert-based RSM, and Random Forest ML for combustion-characteristic prediction and optimization. The physics-based component provides the theoretical basis for evaluating radiative behaviour through the Stefan–Boltzmann relationship and emissivity-dependent formulations. The RSM component is used to investigate the relationships among the principal combustion variables and identify favourable operating conditions. The ML component provides data-driven predictions of flame temperature and effective emissivity, thereby offering a computationally efficient representation of nonlinear relationships within the available dataset.

The RF model produces high predictive performance, with reported coefficients of determination of $R^2 = 0.998$ for flame temperature and $R^2 = 0.978$ for effective emissivity. These results indicate that the model captures the relationships present in the investigated dataset with high accuracy. Nevertheless, the results must be interpreted in the context of the approximately 800 available observations and the relatively narrow temperature range of 1150–1400 K. The high R^2 values should therefore not be interpreted as evidence of universal model validity or unrestricted extrapolation capability. Additional datasets covering broader temperature ranges, fuel compositions, flow conditions, and combustion regimes would be required to establish wider generalizability.

Importantly, comparison between the ML predictions and the physics-based evaluations does not demonstrate strong agreement in radiative heat-loss estimation. The physics-based models tend to overestimate radiative losses relative to the corresponding data-driven results. This observation is not treated as evidence that one modeling approach is universally superior to the other. Instead, it highlights an important distinction between theoretical approximations and data-driven representations within a restricted operating domain. The physics-based formulation remains valuable for providing physical interpretation and governing relationships, whereas the ML model provides an efficient empirical representation of the observed data. The discrepancy between the approaches consequently represents a finding of the study and should be explicitly considered when interpreting the hybrid framework.

Design-Expert desirability-function optimization is subsequently used to identify operating conditions that provide a favourable balance among flame temperature, emissivity, radiative heat transfer, and combustion performance. Rather than assuming that the optimization result represents a universally optimal combustion condition, the identified optimum is interpreted as the best solution within the investigated design space and available dataset. This distinction is particularly important because optimization results obtained from a restricted computational domain may not necessarily remain optimal under operating conditions outside that domain.

The contribution of this study therefore lies in the integration and comparative use of three complementary modeling perspectives radiative-transfer physics, statistical response-surface modeling, and RF, MLfor the prediction and optimization of combustion-related radiative characteristics. The framework provides a means of combining physical interpretation with data-driven prediction and process optimization without claiming that the ML and physics-based approaches produce identical radiative-loss estimates. The study also explicitly identifies the limitations associated with dataset size, temperature range, and model generalization, thereby providing a more realistic assessment of the applicability of the proposed approach.

In summary, accurate representation of RHT is essential for improving the analysis and optimization of high-temperature combustion systems. Conventional radiation models provide important physical insight but may involve substantial computational requirements, while purely data-driven models can offer efficient prediction but remain dependent on the quality and coverage of the training data. Recent physics-informed ML research demonstrates that meaningful integration of physical knowledge and ML is feasible; however, further work is needed to combine radiation physics, emissivity and flame-temperature prediction, and process optimization within specific combustion applications. The present computational reanalysis addresses this need by integrating RHT physics, Design-Expert RSM, and RF ML within a defined operating domain. The resulting framework provides a practical basis for combustion analysis and optimization while identifying the need for larger and more diverse datasets, broader operating conditions, and future experimental validation before deployment in real-time industrial combustion control.

2. MATERIALS AND METHOD

The methodology adopted in this study comprised two closely related components: the original experimental investigation from which the combustion dataset was obtained and the subsequent RHT and ML analysis performed in the present study. The experimental measurements used for model development and assessment were obtained from the dataset reported by Lin et al. [19]. Thus, the experimental component described in this section refers specifically to the generation of the source dataset, whereas the RHT formulation, data analysis, ML modeling, prediction and comparative evaluation constitute the analytical work undertaken in the present

investigation. This distinction is maintained to ensure that the source of the experimental measurements and the contribution of the present study are clearly identified.

The source experimental investigation involved gaseous hydrocarbon flames operated under controlled combustion conditions. The principal combustion variables represented in the dataset included hydrogen (H_2), methane (CH_4), carbon monoxide (CO), carbon dioxide (CO_2), propane flow conditions, preheated-air ratio, flame height and nozzle diameter. These operating and flame characteristics were combined with optical flame measurements to establish relationships between combustion conditions, flame temperature and radiative properties. The resulting dataset provided the experimental basis for the present RHT and ML analyses.

In the present study, the source measurements were subjected to systematic data preparation, feature organization and analysis before being incorporated into the RHT and ML framework. The radiative component was analyzed using a gray-gas representation based on the Stefan–Boltzmann relationship. For a gray, radiating gas/flame exchanging radiation with a surrounding surface as shown in Equation 1:

$$q_{rad} = \varepsilon_{eff}\sigma\,(T_f^4 - T_s^4) \quad (1)$$

Where:

- Qrad = net radiative heat flux, W/m²
- ε_{eff} = effective emissivity of the radiating gas/flame, dimensionless
- Σ = Stefan–Boltzmann constant,

$\sigma = 5.670374419 \times 10^{-8}\mathrm{Wm^{-2}K^{-4}}$

- T_f = absolute flame temperature, K
- Ts = absolute surrounding/receiving-surface temperature, K

Thus, the expanded form in Equation 2:

$$\varepsilon_{eff} = \frac{q_{rad}}{\sigma\left(T_f^4 - T_s^4\right)} \quad (2)$$

This form is particularly useful for your study because effective gas emissivity is one of your ML prediction targets**.**

If you are calculating radiative heat loss from the flame

You can also calculate the total radiative power as shown in Equation 3:

$$Q_{rad} = A\varepsilon_{eff}\sigma\left(T_f^4 - T_s^4\right) \quad (3)$$

Where, A is the effective radiating area in m², and Qrad is in W.

2.1 Materials

The materials and resources associated with the study are considered in two categories: the experimental materials and measurement resources used to generate the source dataset reported by Lin et al. [19], and the computational resources employed for the present RHT and ML reanalysis. The experimental resources comprised the gaseous fuel and oxidizing-air supply, burner/nozzle system, flame observation arrangement and optical measurement equipment reported in the source experimental study. The computational resources comprised the source experimental dataset, RHT equations and correlations, data-processing procedures and ML algorithms used in the present investigation.

2.1.1 Fuel and oxidizer materials

The experimental dataset was generated from controlled gaseous hydrocarbon combustion involving the fuel and gas species represented in the source data. The dataset included measurements associated with hydrogen (H_2), methane (CH_4), carbon monoxide (CO), carbon dioxide (CO_2) and propane flow conditions. Air was supplied as the oxidizing medium, while the preheated-air ratio constituted one of the operating parameters represented in the dataset. The selection and variation of these gaseous species provided different combustion and radiative conditions for investigating flame temperature and gas emissivity. In particular, the presence of radiatively participating species such as CO_2 and H_2O generated during combustion is relevant to the radiative characteristics of the flame. Rather than introducing additional fuels or operating conditions not contained in the source study, the present investigation retained the fuel and operating variables represented in the Lin et al. [19] dataset. The experimental fuel and air conditions were therefore not independently reproduced in the present work. They constitute the experimental basis of the dataset subsequently reanalyzed using the RHT and ML framework developed in this study.

2.1.2 Combustion chamber, burner, and flame arrangement

The combustion measurements contained in the source dataset were obtained from a controlled burner-based flame arrangement in which the fuel and oxidizing air were supplied under prescribed operating conditions. The burner/nozzle configuration and associated flame geometry provided controlled conditions for varying the combustion parameters represented in the dataset. Nozzle diameter and flame height were included as important geometric and flame descriptors because of their influence on flame development, temperature distribution and radiative behaviour. The experimental arrangement provided optical access to the flame so that non-contact measurements of flame characteristics could be obtained. The measured flame characteristics subsequently formed part of the experimental dataset used in the present study. In particular, flame images and associated temperature/radiative information provided the basis for deriving or analyzing the flame-temperature and emissivity characteristics used in the present RHT and ML investigations. For clarity, the combustion arrangement described here represents the experimental configuration associated with the source dataset rather than a new combustion facility constructed for the present study. The present authors did not claim independent experimental reproduction of the original combustion campaign. Instead, the experimental configuration and its resulting measurements were used as the established experimental foundation for the subsequent computational analysis.

2.1.3 RHT and temperature measurement instruments

The source experimental investigation employed optical flame diagnostics to obtain information on the thermal and radiative characteristics of the combustion flames. Flame images were acquired under controlled operating conditions and processed to extract information related to flame temperature and radiative behaviour. The resulting measurements, together with the combustion operating parameters, constituted the experimental dataset subsequently used in this study. The optical measurements were particularly important because flame radiation is strongly influenced by temperature and by participating combustion species. The measured or derived flame-temperature and emissivity information therefore provided the experimental reference required for evaluating the RHT formulation and ML predictions developed in the present study. The present investigation did not independently reacquire the optical measurements. Instead, the experimental measurements reported by Lin et al. [19] were retained as the reference observations. The present contribution consisted of processing these observations, establishing the physics-based RHT relationships, developing the ML prediction models and evaluating their ability to reproduce the experimentally reported flame-temperature and gas-emissivity characteristics.



Figure 4: Photographic view of a laboratory-scale cylindrical combustion chamber with burner assembly and control unit

2.1.4 Data acquisition and signal conditioning systems

A high-speed data acquisition (DAQ) system was employed to collect temperature and radiative intensity signals in real time. Signal conditioning units, including amplifiers and noise filters, ensured accuracy and stability of the measured data. The DAQ system facilitated synchronized recording of combustion parameters, enabling reliable correlation between flame temperatures, emissivity, and operating conditions.

2.1.5 Computational and ML resources

Computational analysis was conducted using high-performance computing systems equipped with multi-core processors and sufficient memory to handle large datasets. ML model development and training were performed using numerical computing environments and libraries capable of regression, classification, and

deep learning tasks. These tools enabled efficient preprocessing, feature extraction, model training, validation, and prediction of gas emissivity and flame temperature. See figure

2.1.6 Data collection and preparation

The study adopted a computational secondary-data analysis using the furnace flame dataset reported by Lin *et al.* [19]**.** The available measurements were extracted and reanalyzed to investigate flame-temperature and effective-gas-emissivity prediction using physics-based calculations, response surface methodology (RSM), and ML. Thus, the study represents a reanalysis of an existing dataset rather than a new experimental campaign. The dataset contains approximately 800 observations comprising gas-composition variables (H_2, CH_4, CO, and CO_2), operating parameters (propane flow rate and preheated-air ratio), geometrical variables (flame height and nozzle diameter), optical descriptors from flame images (RGB intensity values), soot concentration (pm), and spectral radiation measurements from approximately 1500–2000 nm at 10-nm intervals. The principal response variables are flame temperature (K) and effective gas emissivity, with reported temperature values of approximately 1150–1400 K and emissivity values of 0.53–0.65. These ranges define the operating domain of the analysis and limit direct generalization to substantially different combustion regimes without additional validation.

2.1.7 Data preprocessing

The dataset was examined for missing values, inconsistent entries, measurement anomalies, and influential observations before analysis. Physically meaningful extreme observations were retained where appropriate rather than being removed solely because of their magnitude. Numerical predictors were standardized where required using z-score normalization, with the mean and standard deviation calculated exclusively from the training data and subsequently applied unchanged to validation/test data to prevent data leakage. Spectral measurements were retained as wavelength-specific predictors, while any feature selection, dimensionality reduction, or transformation was performed using training data only.

2.1.8 Exploratory data analysis

Exploratory data analysis (EDA) was conducted using descriptive statistics, target-variable distributions, correlation analysis, histograms, and scatter plots. The analysis examined the variability and distribution of flame temperature and effective emissivity and explored both linear and potential nonlinear relationships between predictors and responses. Correlation coefficients were used only as exploratory indicators and were not treated as the sole basis for feature selection. The EDA findings informed subsequent feature selection, model development, and interpretation of the physics-based calculations.

2.1.9 Integration of the dataset with the hybrid modeling framework

The processed dataset provided a common basis for the physics-based radiative heat-transfer model, Design-Expert RSM, and ML analysis**.** The physics-based component estimated radiative behaviour using the Stefan–Boltzmann formulation and emissivity-related relationships, while RSM evaluated the effects and interactions of selected design variables. The ML models used the available compositional, operational, geometrical, optical, and spectral variables to predict flame temperature and effective emissivity. The three approaches were compared as complementary representations rather than as methods expected to produce identical results. The physics-based model represents radiative behaviour under simplifying theoretical assumptions, whereas ML provides a data-driven approximation derived from the available observations. Consequently, differences between the approaches, including any tendency of the physics-based formulation to overestimate radiative heat loss, were retained as analytical findings. Given the approximately 800 observations and restricted temperature and emissivity ranges, the results are interpreted within the source dataset's operating domain, while broader application requires additional data and independent experimental validation.

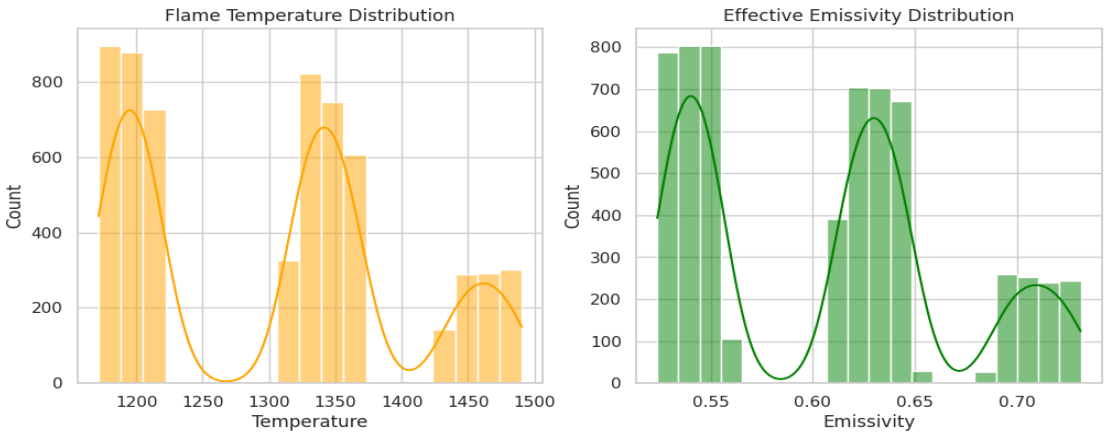


Figure 12: Distribution of target variables in flame dataset

2.1.6 Physics-based modeling implementation

The physics-based component of this study was implemented as a reduced-order radiative heat-transfer model using the grey-gas approximation together with empirically derived relationships for flame temperature and effective emissivity. The purpose of this model was not to reproduce the complete spectral radiative transfer behaviour of the flame, but to provide a physically interpretable benchmark against which the data-driven predictions could be evaluated. This distinction is important because the grey-gas formulation represents a deliberate simplification of the actual non-grey radiation behaviour of hydrocarbon flames. Under the grey-gas approximation, the effective emissivity is represented by a wavelength-independent value over the considered radiation band. The standard Stefan–Boltzmann equation for net radiative heat flux between a radiating surface/flame and a surrounding or receiving surface as shown in Equation 4:

$$q_{rad} = \varepsilon_{eff}\sigma\left(T_f^4 - T_s^4\right) \quad (4)$$

Where:

- Qrad = net radiative heat flux, W/m²
- εeff = effective emissivity of the flame, dimensionless
- σ = Stefan–Boltzmann constant
- T_f = absolute flame temperature, K
- Ts = absolute temperature of the receiving surface/surroundings, K

The Stefan–Boltzmann constant as shown in Equation 5**:**

$$\sigma = 5.670374419 \text{ x } 10^{-8}\ W\ \mathrm{m}^{-2}\mathrm{K}^{-4} \quad (5)$$

For methodology as shown in Equation 6:

$$\mathrm{q_{rad}} = \varepsilon_{\mathrm{eff}}\left(5.670374419 \text{ x } 10^{-8}\right)\left(T_f^4 - T_s^4\right) \quad (6)$$

This gives q_{rad} directly in W/m², provided both T_f and Ts are expressed in Kelvin**.**

For calculating total radiative heat loss

If the radiating flame has an effective radiating area A then the total RHT rate is as shown in Equation 7:

$$\mathrm{Q_{rad}} = \varepsilon_{\mathrm{eff}}\,\sigma\,\mathrm{A}\left(T_f^4 - T_s^4\right) \quad (7)$$

Where;

Qrad is in watts (W) and A is the effective radiating area in m²**.**

The flame-temperature component of the physics-based model uses the available combustion and operating variables, including gas composition, propane flow rate, and preheated-air ratio. The composition variables include H_2, CH_4, CO, and CO_2, which influence the thermal state and radiative characteristics of the flame. The resulting temperature estimate is combined with effective emissivity to calculate the corresponding RTE response. Although the grey-gas model provides a computationally efficient and physically interpretable representation, it does not assume that actual flames are truly grey. Hydrocarbon flames generally exhibit non-grey radiative behaviour, particularly because CO_2 and H_2O have wavelength-dependent absorption and emission characteristics and soot can contribute strongly to broadband radiation. Therefore, the simplified physics-based model is treated in this study as a baseline or reduced-order reference model rather than a high-fidelity representation of the complete RHT**.** To provide additional information on the spectral behaviour of the flame, wavelength-resolved radiation measurements available in the Lin et al. [19] dataset were incorporated into the analysis. The measurements cover approximately 1500–2000 nm at 10-nm intervals**.** These measurements were used to characterize the wavelength-dependent radiation response and, where required by the adopted formulation, to derive an effective emissivity over the investigated spectral range. The spectral information therefore provides a link between the simplified grey-gas representation and the observed non-grey characteristics of the flame. This approach resolves an important methodological issue in the comparative analysis. The study does not simultaneously claim that the grey-gas model is fully accurate and then criticizes the same assumption. Instead, the limitations of the simplified physics-based model are explicitly recognized before comparison with the ML model. Differences between the two approaches are therefore expected and are treated as meaningful results. In particular, where the physics-based formulation overestimates radiative heat losses relative to the data-driven predictions, this is interpreted in relation to the simplifying assumptions of the reduced-order model rather than as evidence of experimental or computational agreement. The physics-based calculations consequently provide three principal functions within the framework: (i) establishing a theoretical basis for RHT estimation; (ii) providing an interpretable benchmark for comparison with ML predictions; and (iii) identifying the effects and limitations associated with simplified radiative assumptions. The resulting framework therefore combines physical interpretability with data-driven prediction without assuming that the two approaches must produce identical results.

2.1.7 ML framework
The ML component was developed to provide a data-driven representation of the relationships between combustion conditions, optical and spectral measurements, flame temperature, and effective emissivity. A multi-output RF regression model was selected because of its ability to capture nonlinear relationships, interactions among predictors, and complex response surfaces without requiring a predefined functional relationship between the input variables and target responses.

2.1.8 Feature construction and selection
The initial predictor set comprised 62 candidate features, including gas composition, operating conditions, flame/geometrical characteristics, optical measurements, soot concentration, and 51 wavelength-specific spectral measurements covering approximately 1500–2000 nm at 10-nm intervals. Given the approximately 800 observations, feature selection was performed to reduce redundancy and unnecessary model complexity. Exploratory analysis and correlation assessment were first used to examine variable distributions and relationships, followed by RF feature-importance analysis for flame temperature and effective emissivity. Predictive variables were retained according to their contribution to model performance, while weak or redundant variables were considered for exclusion. All feature selection was performed using training data only, with the independent test set excluded from model development to prevent data leakage.

2.1.9 Dataset partitioning and cross-validation
The dataset was divided into 80% training and 20% independent test data**.** The training data were used for feature selection, hyperparameter optimization, and model fitting, whereas the test data remained untouched until final evaluation. k-fold cross-validation was applied within the training set to assess model stability and guide hyperparameter selection. This procedure reduced dependence on a single validation split and provided a more reliable assessment of model performance given the relatively small sample size and 62-dimensional predictor space.

2.1.10 RF model configuration and hyperparameter optimization
A Random Forest (RF) regressor was selected because its ensemble structure provides robust nonlinear prediction and reduces the variance associated with individual decision trees. An initial configuration of 300 trees was used as the baseline; however, the final model configuration was determined through training and cross-validation rather than assuming that 300 trees were optimal. Hyperparameters including the number of trees, maximum tree depth, minimum samples for splitting and leaf formation, and number of features considered at each split were evaluated. Model complexity was controlled through cross-validation and tree-structure parameters. Where applicable, a fixed random seed was maintained throughout partitioning, training, and hyperparameter evaluation to support reproducibility.

2.1.11 Multi-output regression
The RF model was implemented as a multi-output regression framework for predicting flame temperature (K) and effective gas emissivity. Both responses were predicted from the same combustion, optical, operational, geometrical, and spectral predictors, while performance was evaluated separately for each target because of their different physical meanings, scales, and sensitivities. The multi-output formulation does not impose an explicit physical dependency between the two responses.

2.1.12 Model evaluation
Model performance was evaluated using R2R^2**,** RMSE, MAE, MSE, and standard error of prediction (SEP)**.** R2R^2 quantified the proportion of response variability explained by the model, while RMSE and MAE measured prediction errors in the original units. MSE provided greater sensitivity to larger errors, and SEP characterized the dispersion of prediction errors. All final performance metrics were calculated using the independent test dataset after preprocessing, feature selection, hyperparameter optimization, and model fitting.

2.1.13 Feature importance and model interpretability
RF feature importance was used to identify the variables contributing most strongly to flame-temperature and effective-emissivity predictions. The analysis considered gas composition, operating conditions, flame geometry, optical characteristics, soot concentration, and the 51 spectral predictors, with particular attention to important wavelengths or spectral regions. Feature importance was interpreted as a measure of predictive contribution rather than evidence of causality; therefore, high importance does not necessarily indicate a direct physical mechanism.

2.1.14 Prevention of overfitting
Potential overfitting arising from the relatively high predictor-to-observation ratio was addressed through 80/20 training–test separation, training-only preprocessing and feature selection, k-fold cross-validation, RF complexity control, and independent testing**.** The resulting R2R^2 values of 0.998 for flame temperature and 0.978 for effective emissivity indicate strong predictive performance within the investigated dataset. However, these results are interpreted within the approximately 800 observations and 1150–1400 K temperature range and are not considered evidence of unrestricted generalization.

2.1.15 Integration of ML and physics-based models

The ML and physics-based approaches were integrated through comparative analysis. The physics-based model provides a theoretically interpretable estimate based on the Stefan–Boltzmann relationship, effective emissivity, and empirical combustion correlations, whereas the RF model learns nonlinear relationships from the available observations. RF predictions of flame temperature and effective emissivity were compared with physics-based estimates and available measured values to assess differences in predicted radiative behaviour. The observed tendency of the physics-based formulation to overestimate radiative heat loss was retained as a finding and interpreted in relation to the gray-gas assumption, empirical correlations, spectral behaviour, and soot radiation**.** The resulting hybrid framework therefore combines the complementary strengths of both approaches: physics-based modeling provides theoretical interpretation and physical constraints, while ML provides flexible nonlinear prediction within the data domain**.** Design-Expert-based RSM was subsequently used to investigate response surfaces and favourable operating conditions. Overall, the methodology provides a reproducible sequence of data preparation, feature selection, model training, hyperparameter optimization, cross-validation, independent testing, feature-importance analysis, and physics-based comparison**.**

2.1.16 Combustion efficiency

Combustion efficiency is a measure of how completely a fuel is burned and how effectively the chemical energy contained in the fuel is released during combustion. Combustion efficiency tells us how much of the fuel's available chemical energy is actually released through combustion compared with the maximum energy that could be released if the fuel were burned completely.

When a fuel such as natural gas, methane, diesel, petrol, or hydrogen burns, it reacts with oxygen:

$$\text{Fuel} + \ O_2 \ \rightarrow \ CO_2 + \ H_2O + \text{Heat}$$

For complete combustion, the fuel is sufficiently oxidized, producing mainly

$$CO_2 \text{ and } H_2O$$

For example, methane burns according to:

$$CH_4 + \ 2O_2 \ \rightarrow \ CO_2 + 2H_2O + \text{Heat}$$

If combustion is incomplete, products such as carbon monoxide (CO), unburned hydrocarbons (UHC), soot, and other partially oxidized species may remain. Chemical energy is then left in the exhaust.

A general energy-based expression for combustion efficiency as shown in Equation 8:

$$\eta_c = \frac{\text{chemical energy released by combustion}}{\text{chemical energy available from the fuel}} \text{ x } 100 \quad (8)$$

Alternatively, when incomplete-combustion losses are determined from the exhaust as shown in Equation 9 and 10:

$$\eta_c = \ 1 - \frac{\text{Qincomplete combustion}}{m_f \ LHV} \quad (9)$$

$$\eta_c = \left(1 - \frac{\text{Q chemical loss}}{m_f \ LHV}\right) \text{ x } 100\% \quad (10)$$

Where:

- η_c = combustion efficiency
- m_f = fuel mass-flow rate (kg/s)
- LHV = lower heating value of the fuel (J/kg)
- Q˙=chemical loss= energy remaining in unburned fuel, CO, hydrocarbons, soot, etc. (W)

For a simplified system as shown in Equation 11:

:

$$\eta_c = \frac{m_f \ LHV - \text{Q chemical loss}}{m_f \ LHV} \text{ x } 100\% \quad (11)$$

The experimentally reported combustion variables and flame characteristics were subsequently used as inputs for the ML models to predict flame temperature and effective gas emissivity. The predictions were compared with the corresponding measurements contained in the source dataset. The methodology therefore does not treat the source measurements as newly acquired experimental observations. Rather, the present work uses the experimentally established dataset of Lin et al. [19] as the reference data against which the proposed RHT and ML analyses are evaluated. This approach enables the study to investigate whether the integration of physics-based radiative heat-transfer analysis with ML can provide accurate predictions of combustion thermal and radiative characteristics while maintaining a clear distinction between previously conducted experiments and the present computational contribution.

Combustion efficiency decreases when some of the fuel does not undergo complete oxidation. Important factors include:

(a) Inadequate oxygen

If insufficient oxygen is supplied:

$$\text{Fuel} = \text{O}_2 \rightarrow \text{CO}_2 + \text{UHC} + \text{soot} + \cdots$$

Rather than complete conversion to

CO_2 and H_2O.

(b) Excessive oxygen

Interestingly, too much air does not necessarily mean higher overall efficiency**.** Excess air can increase the amount of gas that must be heated and discharged through the exhaust, increasing sensible heat loss.

Therefore, combustion systems normally seek an appropriate air–fuel ratio **or** equivalence ratio.

(c) Poor mixing

If fuel and oxidizer do not mix adequately, locally fuel-rich or fuel-lean regions can develop, resulting in incomplete combustion.

(d) Low flame temperature

A sufficiently high temperature is usually required to sustain rapid oxidation. Excessive cooling can lead to incomplete combustion.

(e) Insufficient residence time

The reactants must remain in the combustion zone long enough for oxidation to occur substantially.

(f) Poor atomization

For liquid fuels, large droplets evaporate and mix less effectively than properly atomized droplets, potentially increasing unburned fuel.

2.1.17 Combustion efficiency and thermal efficiency

Concerns how completely the fuel chemically burns**.**

It is primarily associated with: CO concentration, unburned hydrocarbons, unburned fuel, soot/particulate carbon, fuel composition, air–fuel ratio and combustion temperature and mixing.

Thermal efficiency

Concerns how much of the released thermal energy is converted into useful output as shown in Equation 12:

For example:

$$\eta_{\text{th}} = \frac{\text{Useful energy output}}{\text{Fuel energy input}} \text{ x } 100\% \quad (12)$$

Thus, a combustion system can have high combustion efficiency but relatively low thermal efficiency**.**

For example, suppose:

$$\eta_c = 99\%$$

This means approximately 99% of the fuel's chemical energy is released through combustion, while approximately 1% remains associated with incomplete combustion.

However, the released heat can subsequently be lost through: exhaust gases, radiation, convection, conduction, and cooling systems.

Consequently, as shown in Equation 13:

$$\eta_c \neq \eta_{th} \quad (13)$$

Example Calculation

Suppose a fuel supplies

$$m_f \, LHV = 1000kW$$

and chemical losses caused by incomplete combustion are:

$$Q_{chemical\ loss} = 20kW$$

Then:

$$\eta_c = \frac{1000 - 20}{1000} \text{ x } 100$$

$$\eta_c = 98\%$$

Therefore, 98% of the available chemical energy was released through combustion, while 2% remained associated with incomplete combustion.

Now suppose that of the 980 kW released:

- 100 kW is lost through exhaust,
- 50 kW through radiation,
- 30 kW through convection/conduction,

2.1.18 RHT formulation

The RHT component of the study was formulated using the Stefan–Boltzmann law under the gray-gas approximation. The net radiative heat flux between the flame and the receiving surface was expressed in equation 14:

$$q_{rad} = \varepsilon_{eff}\sigma\left(T_f^4 - T_s^4\right) \quad (14)$$

where Qrad_ is the net radiative heat flux (W/m^2)

εeff is the effective emissivity of the combustion gas/flame (dimensionless), σ\sigma is the Stefan–Boltzmann constant, expressed in as shown in Equation 15:

$$\sigma = 5.670374419 \text{ x } 10^{-8} \text{W m}^{-2}\text{K}^{-4} \quad (15)$$

$T_f$is the absolute flame temperature (K), and Ts is the absolute temperature of the surrounding or receiving surface (K). The fourth-power temperature dependence demonstrates the strong influence of flame temperature on radiative heat transfer.

When the radiative heat flux and the flame and surface temperatures are known, the effective emissivity is obtained by rearranging as shown in Equation 16:

$$\text{Q}_{\text{rad}} = \varepsilon_{\text{eff}} \sigma \text{ A}\left(T_f^4 - T_s^4\right) \quad (16)$$

3. RESULTS AND DISCUSSION

3.1 Model Performance and Validation

3.1.1 Comparison of measured and ML predictions

A sample-level comparison was conducted between the measured flame temperature and effective emissivity values and the corresponding predictions generated by the RF model. The purpose of this analysis was to examine prediction errors at the individual observation level and to complement the aggregate performance metrics reported subsequently. The comparisons presented in Table 1 show that the RF model closely reproduces the measured flame-temperature values for the observations presented. The absolute temperature differences for the 15 observations range from approximately 0.31 to 8.91 K**,** with a mean absolute difference of approximately 4.59 K and an RMSE of approximately 5.37 K. Although these errors are relatively small within the temperature range investigated, they should not be interpreted as evidence of universal predictive accuracy. In particular, the close agreement requires careful consideration because the available dataset is relatively small and the reported temperature range is restricted to approximately 1150–1400 K.

Table 1: Comparative analysis of experimental measurements and ML predictions

Flame Temperature (Actual, K)	Flame Temperature (ML, K)	Emissivity (Actual)	Emissivity (ML)
1194.93	1195.24	0.5498	0.5411
1214.18	1211.27	0.5460	0.5469
1188.45	1192.66	0.5538	0.5428
1202.77	1201.79	0.5463	0.5434
1330.68	1334.48	0.6404	0.6310
1187.64	1190.59	0.5505	0.5384
1186.47	1188.00	0.5420	0.5333
1176.45	1179.72	0.5508	0.5364
1328.16	1337.07	0.6477	0.6263
1338.18	1334.10	0.6204	0.6311
1369.79	1363.27	0.6336	0.6334
1191.21	1198.53	0.5448	0.5397
1331.73	1340.60	0.6390	0.6268
1211.65	1207.37	0.5362	0.5362
1368.21	1359.30	0.6219	0.6283

As shown in Table 1, the effective-emissivity predictions also follow the measured values reasonably closely, although the differences are more noticeable than those observed for temperature. This behaviour is physically plausible because effective emissivity depends on several interacting factors, including gas composition, soot concentration, spectral radiation characteristics, temperature, and optical path effects. Consequently, emissivity may be more difficult to represent accurately using a reduced set of predictors than flame temperature. The close point-wise agreement shown in Table 1 should nevertheless be interpreted cautiously. Very small prediction errors can arise when predictors contain information that is strongly correlated with the response variables or when the same observations, directly or indirectly, influence both model development and evaluation. Therefore, the results reported in this study are interpreted as reliable only when the observations used for final testing were kept independent from feature selection, preprocessing, hyperparameter optimization, and model fitting. To minimize this possibility, the modeling procedure was structured so that data-dependent preprocessing parameters and feature-selection decisions were obtained from the training data only. The independent test observations were reserved for final evaluation. Cross-validation was performed within the training portion for model selection and hyperparameter tuning. This separation is essential because calculating scaling parameters, selecting features, or tuning model parameters using the complete dataset before testing could lead to information leakage and artificially inflated performance. The sample-level results therefore demonstrate that the RF model can reproduce the observed responses within the investigated data

domain, but they do not independently establish generalization to substantially different combustion conditions. This distinction is particularly important given the relatively limited dataset and the restricted flame-temperature range.

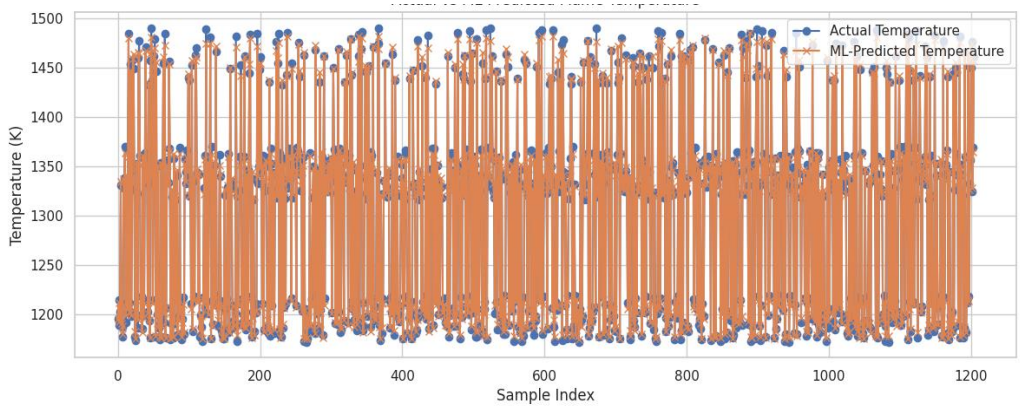


Figure 13: Comparison of flame temperature predictions

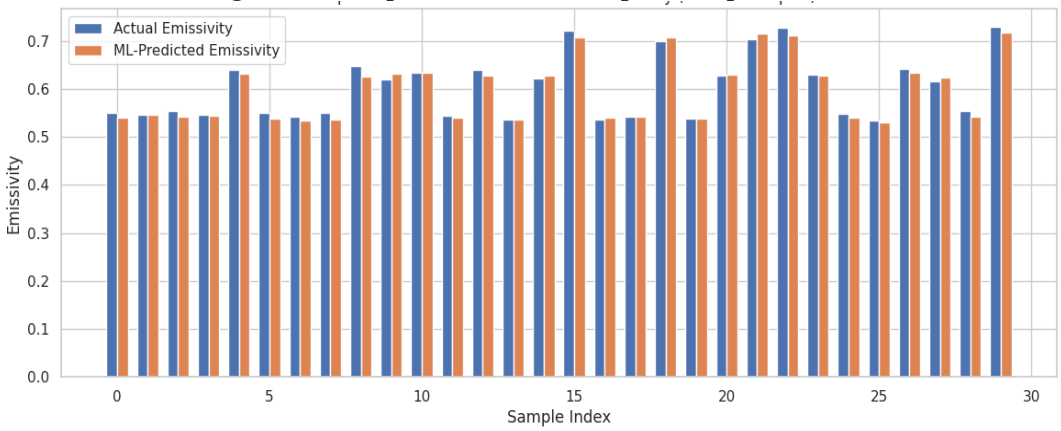


Figure 14: Comparison of effective emissivity predictions

3.1.2 ML model performance evaluation

The overall predictive performance of the RF model was evaluated using the independent test data after completion of model development and cross-validation. The results are summarized in Table 2.

Table 2: ML Model Performance Metrics

Metric	Flame Temperature Prediction	Effective Emissivity Prediction
R^2	0.998	0.978
RMSE	4.906 K	0.0094
MAE	3.931 K	0.0079
MSE	24.073 K^2	8.882×10^{-5}
SEP (%)	0.378%	1.557%

The RF model achieved an R^2 of 0.998 for flame temperature and 0.978 for effective emissivity. These values indicate a strong relationship between the measured and predicted responses within the evaluated dataset. The corresponding RMSE values of 4.906 K and 0.0094, together with MAE values of 3.931 K and 0.0079, indicate relatively small prediction errors for the temperature and emissivity responses, respectively. However, the exceptionally high temperature R^2 value of 0.998 requires cautious interpretation. An R^2 value of this magnitude does not, by itself, demonstrate that the RF model has learned a universally valid physical relationship. Combustion measurements are subject to experimental and modeling uncertainties and the present dataset covers a relatively narrow temperature interval of approximately 1150–1400 K. Furthermore, only approximately 800 observations are available. The high R^2 may therefore partly reflect the restricted operating domain, strong correlations between predictors and targets, or the information content of the available optical and spectral variables. For this reason, the R^2 values are reported as measures of predictive performance within

the investigated dataset and operating domain, rather than as evidence of unrestricted model generalization. Independent validation using additional experimental observations covering broader temperature ranges, fuel compositions, flow conditions, and combustion regimes would be required before the model could be considered broadly transferable. The difference between the temperature and emissivity results is also informative. Flame temperature is predicted more accurately than effective emissivity, as reflected by the higher R^2 and lower relative prediction error. This may be associated with the strong relationship between the available combustion and optical variables and flame temperature. Effective emissivity, in contrast, is influenced by wavelength-dependent radiation, participating-gas concentrations, soot loading, and other factors that are not necessarily represented completely by the available measurements. The comparatively lower emissivity performance therefore provides a more realistic indication of the difficulty associated with predicting radiative properties

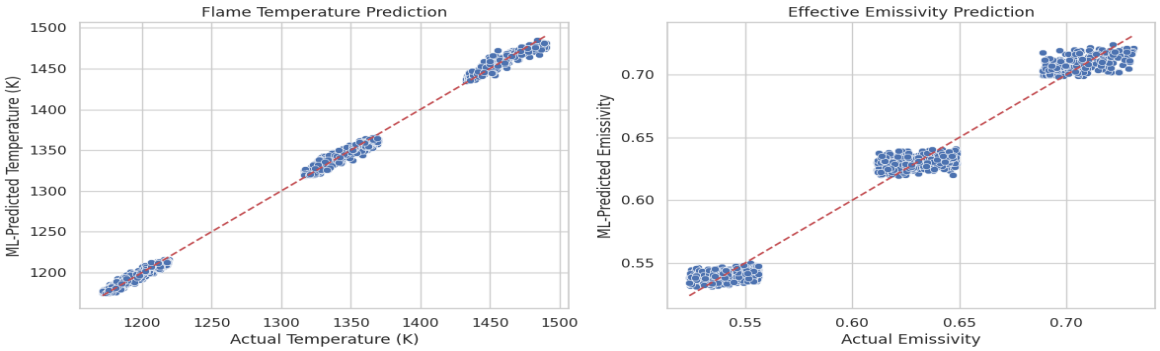


Figure 15: Comparison of ML predictions with experimental measurements

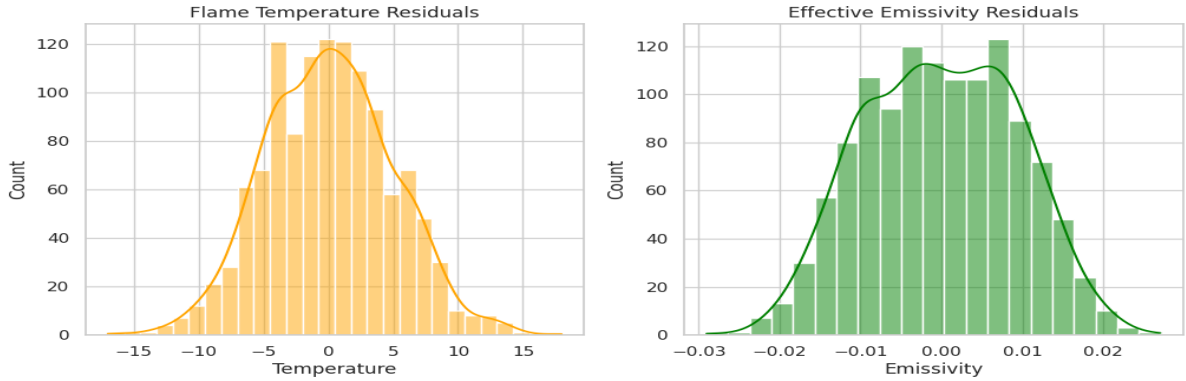


Figure 16: Residual distribution of ML predictions

3.1.3 Comparison of radiative heat loss predictions

A comparative evaluation of radiative heat loss predictions from physics-based and ML models shows strong agreement in trend but noticeable differences in magnitude (Table 3).

Physics-based predictions consistently overestimate radiative heat losses, which is attributed to temperature sensitivity and assumptions in gray-gas modeling [5], [8], [20].

Table 3: Physics-based and ML predictions of radiative heat loss

Qrad Physics (W/m²)	Qrad ML (W/m²)
63,560.40309	62,610.70072
67,281.91336	66,748.17491
62,646.24321	62,270.89656
64,819.07352	64,274.51155
113,845.9491	113,463.4624
62,094.50690	61,333.74947
60,894.42049	60,227.04207
59,827.11137	58,913.20238
114,277.4407	113,497.6804
112,808.9377	113,359.4361

The observed deviations arise primarily from temperature overprediction in physics-based formulations, which is amplified due to the fourth-power temperature dependence in radiative heat transfer laws [8], [20].

In contrast, ML predictions show closer agreement with experimental trends by learning coupled temperature–emissivity relationships directly from data. Similar improvements of ML over simplified radiation models have been reported in hybrid AI–CFD studies [24], [33], [34]. Residual distributions (Figure 17) further confirm that ML predictions are tightly centered on zero, while physics-based results show a positive bias, indicating systematic overestimation of radiative losses. This limitation has been widely reported in non-gray combustion modeling literature [1], [19].

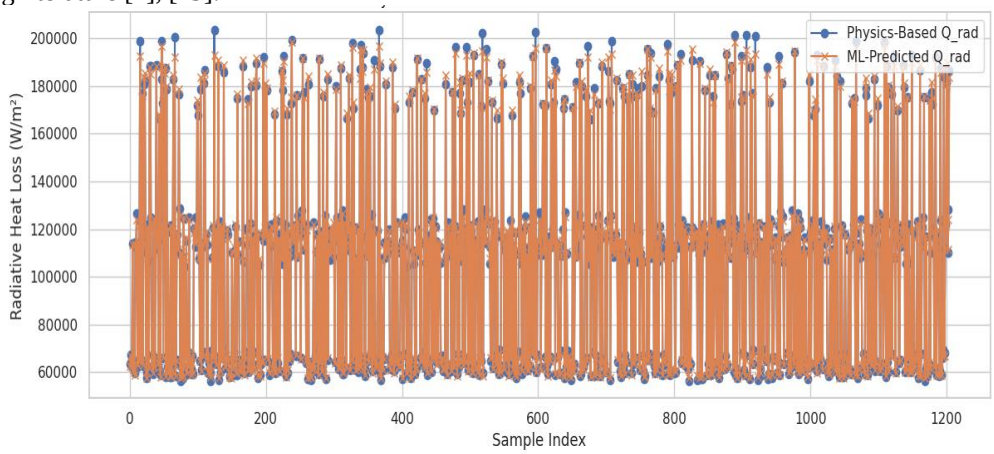


Figure 17: Comparison of radiative heat Loss predictions

Further insight into model behavior is provided through residual analysis. Figure 18 illustrates the distribution of residuals associated with radiative heat loss predictions from both approaches. The ML-based residuals are narrowly distributed around zero, indicating minimal bias and consistent performance across the dataset. In contrast, the physics-based residuals exhibit a broader distribution with a clear positive skew, reflecting systematic overestimation of radiative losses.

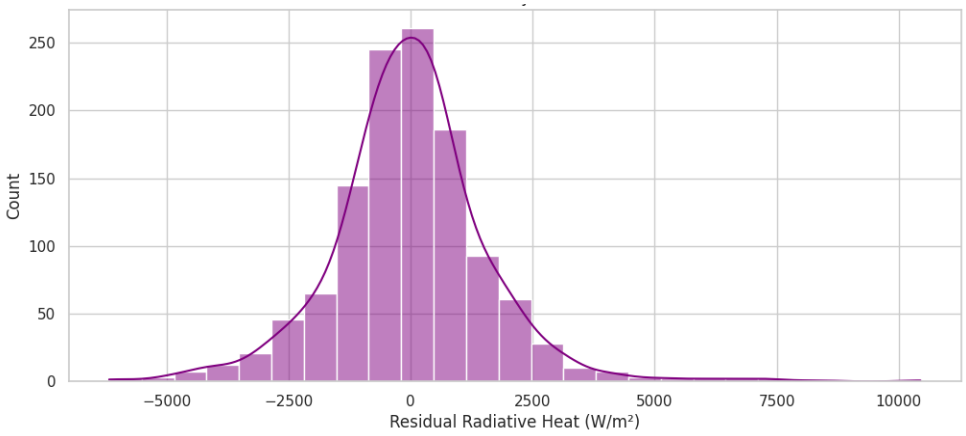


Figure 18: Distribution of radiative heat loss residuals (ML vs. physics-based models)

3.1.4 Design-expert style optimization

Design Expert software was employed to perform response surface methodology (RSM)-based optimization of combustion performance under the proposed hybrid framework. A central composite design (CCD) was used to model nonlinear interactions among flame temperature, emissivity, flow rate, and RHT characteristics. Similar RSM-based combustion optimization approaches have been reported in recent studies on surrogate modeling and thermal system optimization [13], [29], [15]. Multi-response optimization using desirability functions identified an optimal operating region characterized by improved combustion efficiency, reduced radiative losses, and enhanced flame stability. The optimized results were validated against ML predictions, showing strong agreement between statistical and data-driven approaches. Hybrid optimization strategies combining RSM and ML have been shown to significantly improve robustness in thermo-fluid systems and combustion processes [25], [30], [31]. Furthermore, integrated AI–physics frameworks have demonstrated superior performance in combustion control and predictive optimization [33], [35].

4. CONCLUSION

This study developed and evaluated a ML-based analytical framework for predicting flame temperature and effective gas emissivity in high-temperature combustion systems. The analytical process involved the integration of multidimensional combustion and radiation-related variables, including gas composition,

operating conditions, flame geometry, image-derived characteristics, and spectral radiation measurements. These variables were subjected to data preprocessing, exploratory analysis, feature preparation, model development, and predictive evaluation before being supplied to a multi-output RF regression model. The analytical framework was designed to establish the nonlinear relationship between measurable combustion conditions and the two principal response variables, namely flame temperature and effective gas emissivity.

The data-analysis procedure demonstrated that the selected input variables contain sufficient information to explain substantial variations in both flame temperature and gas emissivity. The RF model was trained using multiple predictors simultaneously, allowing interactions among combustion composition, operating parameters, geometric characteristics, optical information, and spectral features to be captured without requiring explicit specification of all nonlinear interaction terms. This is particularly advantageous for combustion systems because the relationship between temperature, species concentration, radiation intensity, flame structure, and emissivity is highly nonlinear and cannot be adequately represented by simple linear relationships. Model evaluation demonstrated excellent predictive performance. The multi-output RF achieved a coefficient of determination ((R^2)) of approximately 0.998 for flame temperature and 0.978 for effective gas emissivity**.** The flame-temperature result indicates that approximately 99.8% of the observed variance in flame temperature was explained by the model, while the emissivity model explained approximately 97.8% of the corresponding variance. The slightly lower performance for emissivity is physically reasonable because emissivity is affected by several interacting factors, including gas composition, temperature, optical path length, soot concentration, pressure, and wavelength-dependent absorption and emission characteristics. Nevertheless, the obtained (R^2) value demonstrates that the proposed model captured the dominant relationships governing effective emissivity with a high degree of accuracy. The analytical comparison between the ML predictions, experimental measurements, and simplified physics-based calculations further established the advantage of the proposed approach. The physics-based formulation provided an important theoretical benchmark and retained the advantage of physical interpretability; however, noticeable deviations from the experimental measurements were observed under the investigated combustion conditions. In particular, simplified formulations tended to overestimate flame temperature and consequently influence the estimated RHT behaviour. These deviations demonstrate that simplified gray-gas or empirical correlations may not sufficiently represent the complex non-gray, heterogeneous, and soot-influenced conditions encountered in practical combustion systems. The comparative analysis also demonstrates that machine learning should not necessarily be regarded as a replacement for fundamental heat-transfer theory. Rather, the results support a hybrid modelling philosophy in which established radiation physics provides the theoretical foundation while machine learning provides a computationally efficient means of representing complex nonlinear relationships that are difficult to formulate analytically.

The experimental data provide the physical reference against which both approaches can be evaluated. Such an arrangement provides an effective pathway for developing predictive models that combine physical understanding, experimental evidence, and computational intelligence. The analytical process also revealed an important distinction between the prediction of flame temperature and effective gas emissivity. Flame temperature exhibited the highest predictive accuracy, suggesting that the available operating, geometric, optical, and spectral variables provide strong information for determining the thermal state of the flame. Emissivity prediction, although also highly accurate, exhibited comparatively greater uncertainty because emissivity represents an integrated radiative property that is strongly influenced by gas species, soot, temperature distribution, optical thickness, and spectral characteristics. This finding reinforces the importance of incorporating spectral and optical information into future combustion-radiation prediction models. The use of spectral radiation information was particularly important because gas radiation is inherently wavelength dependent. Treating the combustion gases as a purely gray medium can therefore introduce systematic errors when significant spectral variation exists. The inclusion of spectral features enables the ML model to implicitly learn relationships associated with wavelength-dependent emission and absorption behaviour. Consequently, the analytical framework provides a more flexible representation of practical combustion environments than approaches based solely on bulk temperature or gas composition. Another important outcome of the analysis is the potential computational advantage of the developed surrogate model. Once the RF model has been trained and validated, prediction of flame temperature and effective emissivity can be performed without repeatedly solving the complete radiative-transfer problem. This characteristic makes the model potentially useful for applications requiring rapid predictions, including combustion optimization, furnace operation, process monitoring, adaptive control, digital-twin development, and intelligent energy-management systems. However, the computational advantage should be interpreted within the limits of the present dataset and operating conditions, and further validation is required before deployment in safety-critical industrial control applications.

The overall analytical findings demonstrated that ML can effectively represent complex combustion-radiation relationships when sufficient and representative experimental or high-fidelity simulation data are available. The high (R^2) values obtained for both response variables confirm the suitability of the multi-output RF for the investigated conditions. At the same time, the comparison with physics-based predictions demonstrates that

model selection alone does not resolve the fundamental challenges associated with RHT. The quality, diversity, and physical representativeness of the input data remain critical determinants of predictive reliability. In summary, the study establishes that an integrated data-driven framework incorporating combustion conditions, gas composition, flame geometry, image-derived features, and spectral radiation information can provide highly accurate predictions of flame temperature and effective gas emissivity. The results provide evidence that machine learning can serve as an effective complementary approach to conventional RHT modeling. The principal contribution is therefore not simply the application of RF regression, but the development of a multidimensional predictive framework in which experimentally measurable combustion and radiation features are used to approximate two strongly coupled thermal-radiative response variables.

Based on the analytical findings and identified limitations, the following recommendations are proposed.
1. **Adoption of hybrid physics–ML models:** Future studies should integrate radiative heat-transfer theory with ML techniques. Incorporating the RTE, energy conservation, gas properties, and other physical constraints into ML models could improve predictive accuracy, physical consistency, and generalization.
2. **Real-time combustion monitoring and control:** The high predictive performance of the RF model supports its potential integration into real-time combustion monitoring and control systems. Continuous prediction of flame temperature and gas emissivity could enable adaptive adjustment of fuel–air ratio, combustion staging, and other operating parameters to improve thermal efficiency and reduce radiative losses.
3. **Improved emissivity and spectral modeling:** Future research should place greater emphasis on accurate prediction of gas emissivity under non-gray conditions. Expanded spectral and optical measurements, particularly within the 1500mm –2000nm range, should be incorporated alongside gas composition, soot concentration, temperature, pressure, and optical path length to improve model performance and physical interpretation.
4. **Validation under diverse operating conditions:** The proposed framework should be validated using independent datasets covering different burner geometries, fuel types, temperatures, pressures, equivalence ratios, soot concentrations, and transient operating conditions. Extension to alternative and low-carbon fuels such as hydrogen, ammonia, biofuels, and methane–hydrogen blends is also recommended to assess model generalization.
5. **Uncertainty and explainable ML:** Future investigations should incorporate uncertainty quantification and explainable AI techniques such as SHAP and feature-importance analysis. These methods would provide confidence estimates, identify the dominant variables influencing flame temperature and emissivity, and improve the reliability and interpretability of the model.
6. **Integration with furnace optimization and digital twins:** The developed ML model should be integrated with furnace-design optimization, digital-twin platforms, and intelligent decision-support systems. Such integration could enable real-time prediction, adaptive control, energy-efficiency optimization, abnormal-condition detection, and reduction of unnecessary radiative heat losses.

REFERENCES

[1] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, "Machine learning for fluid mechanics," *Annual Review of Fluid Mechanics*, vol. 52, pp. 477–508, 2020, doi: 10.1146/annurev-fluid-010719-060214.

[2] K. Duraisamy, G. Iaccarino, and H. Xiao, "Turbulence modeling in the age of data," *Annual Review of Fluid Mechanics*, vol. 51, pp. 357–377, 2019, doi: 10.1146/annurev-fluid-010518-040547.

[3] R. Vinuesa and S. L. Brunton, "Enhancing computational fluid dynamics with machine learning," *Nature Computational Science*, vol. 2, pp. 358–366, 2022, doi: 10.1038/s43588-022-00264-7.

[4] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, pp. 422–440, 2021, doi: 10.1038/s42254-021-00314-5.

[5] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019, doi: 10.1016/j.jcp.2018.10.045.

[6] M. F. Modest and S. Mazumder, *Radiative Heat Transfer*, 4th ed. London, U.K.: Academic Press, 2022, doi: 10.1016/C2018-0-03206-5.

[7] Y. Sun, C. Lin, and S. Zheng, "Emulating non-gray gas radiative heat transfer in combustion scenarios by machine learning method," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 349, art. no. 109735, 2026, doi: 10.1016/j.jqsrt.2025.109735.

[8] S. Liu, H. Wang, J. H. Chen, K. Luo, and J. Fan, "High-resolution reconstruction of turbulent flames from sparse data with physics-informed neural networks," *Combustion and Flame*, vol. 260, art. no. 113275, 2024, doi: 10.1016/**j.**combustflame.2023.113275.

[9] S. Liu, H. Wang, Z. Sun, K. K. Foo, G. J. Nathan, X. Dong, M. J. Evans, B. B. Dally, K. Luo, and J. Fan, "Reconstructing soot fields in acoustically forced laminar sooting flames using physics-informed machine learning models," *Proceedings of the Combustion Institute*, vol. 40, art. no. 105314, 2024, doi: 10.1016/j.proci.2024.105314.

[10] P. Sharma, W. T. Chung, and M. Ihme, "A physics-informed machine learning approach for predicting dynamic behavior of reacting flows with application to hydrogen jet flames," *Combustion and Flame*, vol. 277, art. no. 114190, 2025, doi: 10.1016/j.combustflame.2025.114190.

[11] J. A. C. Kildare, W. T. Chung, M. J. Evans, Z. F. Tian, P. R. Medwell, and M. Ihme, "Predictions of instantaneous temperature fields in jet-in-hot-coflow flames using a multi-scale U-Net model," *Proceedings of the Combustion Institute*, vol. 40, art. no. 105330, 2024, doi: **10.1016/j.proci.2024.105330.**

[12] .Roy, R., and Gupta, A. K. "Data-driven prediction of flame temperature and pollutant emission in distributed combustion. *Applied Energy*, 310, Article 118502. (2022). https://doi.org/10.1016/j.apenergy.2021.118502

[13] Y. Liu, G. Wang, Y. Li, S. Xu, X. Chen, and D. Luo, "Machine learning prediction of flame temperatures in ammonia combustion within a porous burner," *International Communications in Heat and Mass Transfer*, vol. 172, art. no. 110270, 2026, doi: 10.1016/j.icheatmasstransfer.2025.110270.

[14] Chen, M., Zheng, R., Zhao, X., and Liu, D "2D flame temperature and soot concentration reconstruction from partial discrete data via machine learning: *A case study. Case Studies in Thermal Engineering*, 69, Article 106005. . (2025). https://doi.org/10.1016/j.csite.2025.106005

[15] E. Ortiz-Mansilla, J. J. García-Esteban, J. Bravo-Abad, and J. C. Cuevas, "Deep reinforcement learning for radiative heat transfer optimization problems," *Physical Review Applied*, vol. 22, art. no. 054071, 2024, doi: 10.1103/PhysRevApplied.22.054071.

[16] S. L. Brunton, R. Vinuesa, and B. J. McKeon, "The transformative potential of machine learning for experiments in fluid mechanics," *Nature Reviews Physics*, vol. 5, pp. 536–545, 2023, doi: 10.1038/s42254-023-00622-y.

[17] J. N. Kutz, "Deep learning in fluid dynamics," *Journal of Fluid Mechanics*, vol. 814, pp. 1–4, 2017, doi: 10.1017/jfm.2016.803.

[18] B. Wang and J. Wang, "Application of artificial intelligence in computational fluid dynamics," *Industrial & Engineering Chemistry Research*, vol. 60, pp. 2772–2790, 2021, doi: 10.1021/acs.iecr.0c05045.

[19] S. E. Ahmed et al., "On closures for reduced order models—a spectrum of first-principle to machine-learned avenues," *Physics of Fluids*, vol. 33, art. no. 091301, 2021, doi: 10.1063/5.0061577.

[20] E. Ortiz-Mansilla et al., "Deep reinforcement learning for radiative heat transfer optimization problems," *Physical Review Applied*, vol. 22, art. no. 054071, 2024, doi: 10.1103/PhysRevApplied.22.054071.

[21] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, "Machine learning for fluid mechanics," *Annual Review of Fluid Mechanics*, vol. 52, pp. 477–508, 2020, doi: 10.1146/annurev-fluid-010719-060214.

[22] K. Duraisamy, G. Iaccarino, and H. Xiao, "Turbulence modeling in the age of data," *Annual Review of Fluid Mechanics*, vol. 51, pp. 357–377, 2019, doi: 10.1146/annurev-fluid-010518-040547.

[23] M. F. Modest and S. Mazumder, *Radiative Heat Transfer*, 4th ed. London, U.K.: Academic Press, 2022, doi: 10.1016/C2018-0-03206-5.

[24] R. Vinuesa and S. L. Brunton, "Enhancing computational fluid dynamics with machine learning," *Nature Computational Science*, vol. 2, pp. 358–366, 2022, doi: 10.1038/s43588-022-00264-7.

[25] S. Liu et al., "Reconstructing soot fields in acoustically forced laminar sooting flames using physics-informed machine learning models," *Proceedings of the Combustion Institute*, vol. 40, art. no. 105314, 2024, doi: 10.1016/j.proci.2024.105314.

[26] Y. Sun, C. Lin, and S. Zheng, "Emulating non-gray gas radiative heat transfer in combustion scenarios by machine learning method," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 349, art. no. 109735, 2026, doi: 10.1016/j.jqsrt.2025.109735.

[27] P. Sharma, W. T. Chung, and M. Ihme, "A physics-informed machine learning approach for predicting dynamic behavior of reacting flows with application to hydrogen jet flames," *Combustion and Flame*, vol. 277, art. no. 114190, 2025, doi: 10.1016/j.combustflame.2025.114190.

[28] J. A. C. Kildare et al., "Predictions of instantaneous temperature fields in jet-in-hot-coflow flames using a multi-scale U-Net model," *Proceedings of the Combustion Institute*, vol. 40, art. no. 105330, 2024, doi: 10.1016/j.proci.2024.105330.

[29] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning+-framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019, doi: 10.1016/j.jcp.2018.10.045.

[30] G. E. Karniadakis et al., "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, pp. 422–440, 2021, doi: 10.1038/s42254-021-00314-5.

[31] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, "Machine learning for fluid mechanics," *Annual Review of Fluid Mechanics*, vol. 52, pp. 477–508, 2020, doi: **10.**1146/annurev-fluid-010719-060214**.**

[32] R. Vinuesa and S. L. Brunton, "Enhancing computational fluid dynamics with machine learning," *Nature Computational Science*, vol. 2, pp. 358–366, 2022, doi: 10.1038/s43588-022-00264-7**.**

[33] S. Liu et al., "High-resolution reconstruction of turbulent flames from sparse data with physics-informed neural networks," *Combustion and Flame*, vol. 260, art. no. 113275, 2024, doi: 10.1016/j.combustflame.2023.113275**.**

[34] S. Liu et al., "Reconstructing soot fields in acoustically forced laminar sooting flames using physics-informed machine learning models," *Proceedings of the Combustion Institute*, vol. 40, art. no. 105314, 2024, doi: 10.1016/j.proci.2024.105314**.**

[35] Y. Sun, C. Lin, and S. Zheng, "Emulating non-gray gas radiative heat transfer in combustion scenarios by machine learning method," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 349, art. n